The effects of changes in macroeconomic factors for risk parameters on the bank's mortgage portfolio

Master thesis at the University of Twente
Financial Engineering & Management

UNIVERSITEIT TWENTE.

D.J. Meel BSc (Daniël)
s0141208

Amersfoort, July 15th, 2011

Graduation committee:

Financial institution:	Drs. D. Linker (Daniël)
	Drs. N.J.T. Munting (Nico)
University of Twente	Dr. B. Roorda (Berend)
	Ir. Drs. A.C.M. de Bakker (Toon)

Management summary

The financial institution within which this study was conducted, translated the drive to get more insight in future developments of the mortgage portfolio in the main study objective of this thesis. The objective is to create models for credit risk management based on macroeconomic factors to predict the expected credit losses on the managed mortgage portfolios with regard to several macroeconomic scenarios. In current forecasting models that are used in the financial institution to estimate provisions for the mortgage portfolios, a normal market development is assumed, of course extended with imposed stress test scenarios. This experimental research analyses potential changes in the mortgage portfolio due to non-normal scenarios.

Study

The goal of this study is to forecast the risk parameters of the portfolios by linking them to macroeconomic factors, such as unemployment and interest rates, to estimate changes in credit losses under macroeconomic scenarios. The advantage of this approach is the possibility to investigate expected portfolio consequences of changes in the macroeconomic environment for the near and middle long future. In case of successful model building, macroeconomic scenarios can be used as input to predict the default fraction of the portfolio and potential losses (called risk parameters). The main research question reflects this goal:

What is the influence of macroeconomic factors on the risk parameters for the mortgage portfolio?

Methodology

Macroeconomic factors that might influence the risk parameters of the mortgage (default) portfolio are derived from a brief literature study and as a starting point scenarios are selected. The relation between macroeconomic factors and the default rates are observed by correlation studies to determine the best time lags. By use of logistic linear regression, with regard to time lags of the macroeconomic factors, the best combination of factors that estimates the number of defaults in a financial period of a month is observed and corresponding parameters are calculated. This is called the default rate, the probability of getting in default.

The loss rate (LR) is connected to macroeconomic factors by using microeconomic factors, such as Loan-to-Value (LTV) and Loan-to-Income (LTI) ratios, as intermediate step. A cross-table with LTV- and LTI-classes shows the relationship between the loss rate and both explanatory variables; higher classes correspond with higher losses. The LTV and LTI are linked to the house prices and unemployment, respectively. Because of the lack of information about the applicants, especially about their employment status, the loss rate is directly derived from the LTV-ratio. This approach is time dependent and therefore favored to the cross-table.

The only step that has to be taken to collect all information for predicting the future credit losses is to estimate the portfolio value. There are several ways to make a useful estimation, but a macroeconomic link is hard to defend. Therefore the current trend is extrapolated to complete the credit loss estimation. The total credit losses for the portfolios in scope, (1) Intermediary Channel, (2) White Label and (3) a Consolidated Portfolio including (1) and (2) and two more small passive labels, is the multiplication of the probability to get into default (default rate), the

fraction the financial institution will lose in case of default (loss rate) and the total value of the default portfolio (exposure value). Because the default rate and the loss rate could co-operate, especially by including the same input factors, it can be assumed that a correlation between the default rate and loss rate is included. Therefore a covariance analysis was performed to eventually correct the multiplication for over- or underestimation of the credit losses.

The study gives insight in future default rates, loss rates and credit losses based on selected scenarios and extended with a time series scenario. The Time Series Scenario is constructed by developing time series models for each underlying macroeconomic factor and forecasts of the risk parameters are made by using these time series forecasts as input. In other scenarios the end value of the input factor is known and a straight line from now till the end value over the forecast period is assumed. The individual factors are brought together with a regression analysis for each rate. The observed parameters are used for the forecasts.

Results and conclusions

All default rates models are calculated based on macroeconomic factors and an autoregressive term, sometimes extended with a constant value. See Figure I for the default rate (DR) of the Intermediary Channel. Loss rates are based on house prices and a constant by deriving from the LTV-ratios. Covariance between the default rate and loss rate is estimated on the aggregated level and a corrected multiplication is used to estimate the expected losses on a loan as presented in Figure II for the Consolidated Portfolio.

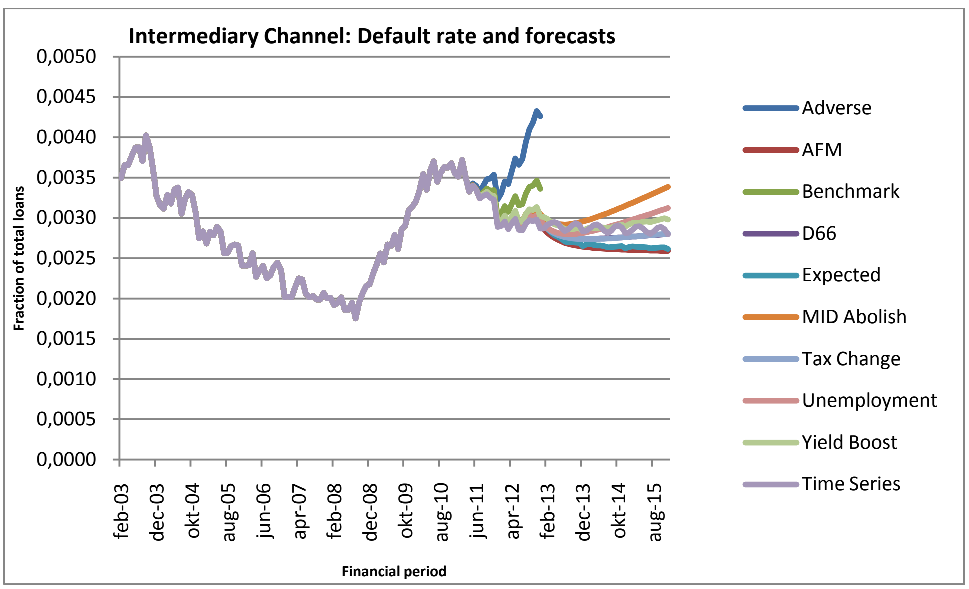


Figure I: Intermediary Channel default time series and forecasts based on applied scenarios

A result of the analyses is the huge impact on the default rate of eliminating the mortgage interest deduction (MID). Most of the scenarios are estimating the default rate on the middle long run between 0,25 and 0,30 percent. Increases of yield, unemployment or the abolishing of the MID are affecting the DR clearly. Stress scenarios (Adverse and Benchmark) are obviously resulting in worse default rates (and calculated on a shorter time horizon).

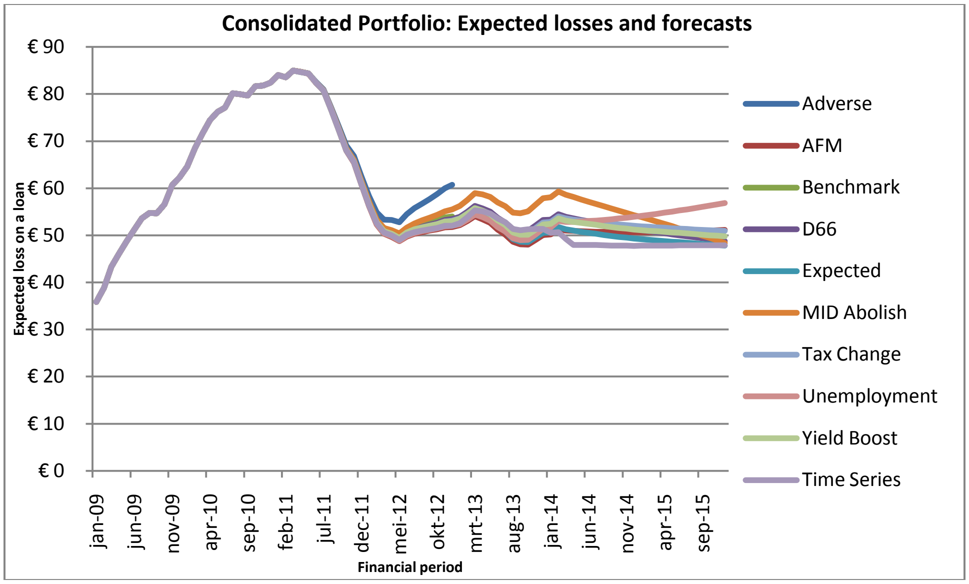


Figure II: Consolidated Portfolio expected loss and forecasts based on applied scenarios

The covariance between the DR and LR is negligible and therefore hardly not affecting the results. The expected loss on a loan is expected to stabilize around 50 Euros. On the short term the elimination of the MID will increase the loss, but in the last forecasted year the unemployment scenario is performing worse.

Although most rates are hard to predict by macroeconomic input, this approach is favored for the DR in the Intermediary Channel and Consolidated Portfolio compared to an approach only depending on the history of the rate. For the White Label, the macroeconomic inputs are not improving the model.

Index

Glossary

ADF = Augmented Dickey-Fuller test, a unit root test for time series

AR = Autoregressive term, a term referring to a value in the history of the own time series

ARIMA = Autoregressive integrated moving average model, a time series model

DR = Default rate, the fraction of defaulted loans (default definition used = three or more months in arrears) in the mortgage portfolio:

$$DR(t) = \frac{Number\ of\ defaulted\ loans\ in\ financial\ period\ t}{Number\ of\ loans\ in\ specified\ portfolio\ in\ financial\ period\ t}$$

EAD = Exposure At Default, Basel-term of the exposure value (EV)

EV = Exposure value, the outstanding loan value of a defaulted loan

FR = Foreclosure rate, the fraction of loans ending up in foreclosure in the mortgage portfolio

$$FR(t) = \frac{Number\ of\ loans\ that\ were\ in\ default\ on\ t{-}1\ and\ out\ of\ the\ portfolio\ on\ t\ in\ financial\ period\ t}{Number\ of\ loans\ in\ specified\ portfolio\ in\ financial\ period\ t}$$

IFRS = International Financial Reporting Standards, a standard for financial reporting

IR = Inflow rate, the fraction of new defaulted loans in the mortgage portfolio:

$$IR(t) = \frac{Number\ of\ new\ defaults\ in\ financial\ period\ t}{Number\ of\ loans\ in\ specified\ portfolio\ in\ financial\ period\ t}$$

LGD = Loss Given Default, Basel-term of the loss rate (LR)

LTI = Loan-to-Income, the value of the loan divided by the income of the applicant

LTV = Loan-to-Value, the value of the loan divided by the value of the security

LR = Loss rate, the fraction of the total loan value that is lost due to a default:

$$LR(t) = average\left(\frac{Loss\ on\ a\ default\ in\ financial\ period\ t}{Value\ of\ a\ defaulted\ loan\ in\ financial\ period\ t}\right)$$

MID = Mortgage interest deduction

PD = Probability of Default, Basel-term of the default rate (DR)

RR = Recovery rate, the fraction of loans out of default in the mortgage portfolio:

$$RR(t) = \frac{Number\ of\ loans\ that\ were\ in\ default\ on\ t{-}1\ and\ not\ on\ t\ in\ financial\ period\ t}{Number\ of\ loans\ in\ specified\ portfolio\ in\ financial\ period\ t}$$

SIC = Schwarz info criterion, an algorithm for model selection

1. Introduction

This experimental research thesis describes an investigation of the macroeconomic influences on the mortgage portfolio of a large Dutch bank and incorporated labels. To introduce the study, the relevance of the research will be emphasized in *Section 1.1*. The study objectives are described in *Section 1.2*, a broad overview of the scope is written in *Section 1.3* and an outline of the report is added in *Section 1.4*.

1.1 Relevance

Predicting the provisions is important business for financial institutions, normally performed by the risk management department. Supervision stimulates financial institutions to provide relevant information about the risk position to assess the current healthiness of the organization. It is not only the external inquiry that drives the need; especially the financial institution itself has a strong interest in expected changes in the mortgage portfolio under different circumstances. Therefore, the development of the default portfolio should be predicted under several scenarios to estimate the provisions necessary. Provisions are often determined in a meeting, based on the expected losses, which is the objective of this experimental study.

Many organizations are driving by the rear-view mirror, very often without respect to time developments, but particularly in a dynamic environment (time dependent) forecasting based on the actual data, involving time developments, is very important. The financial crisis taught the importance of reliable forecasts, but it takes time to implement model changes. In this thesis a possible approach to translate expected changes in macroeconomic factors into default portfolio forecasts is elaborated.

The direct motive for performing this study is the lack of forecasting models within the financial institution based on expected changes in the (general) economy to get insight in default portfolio changes and developments. Current models are assuming normal market developments, except the stress test scenarios. Therefore, an increasing interest in scenario analyses arises within the financial institution, primarily due to recent events. The objective is to develop models predicting the default probabilities and losses of the mortgage portfolio under certain circumstances (i.e. scenarios).

1.2 Study objectives

The focus of the study is on the aggregated level of the mortgage portfolio of the financial institution, one of the largest banks in The Netherlands, including several labels. In this study, macroeconomic factors are linked to the mortgage portfolio to illustrate the effects of several common scenarios, primarily concentrating on defaults.

Conventional models are disregarded in the model developing phase, except the traditional building blocks. More specifically, traditional models are built up according to the Basel guidelines. In this thesis other definitions are chosen, but the building blocks to estimate the expected losses are more or less the same. The bank's interpretation of the IFRS-definitions of being in default is used. Specifically, a loan that is three or more financial periods (measured in months) in arrears is considered as a default.

Financial institutions are used to quantify the risks in terms of Probability of Default (PD), Loss Given Default (LGD) and Exposure at Default (EAD). The latter term is a value expressed in a certain currency, unlike both other risk parameters that are fractions. These terms are assigned to each loan separately by models that are backtested frequently, resulting in general PD-, LGD- and EAD-estimates, basically known as the Basel-parameters.

In this thesis, the PD is called default rate (DR) and is defined as the fraction of total loans in a financial period (i.e. a month) that is in default (three or more financial periods in arrears). The LGD is called the loss rate (LR) and is defined as the loss fraction of the total loan value on a loan in default. The EAD is the exposure value (EV) of a loan in default in a specific financial period. Because this thesis includes other definitions for the Basel-parameters, the overall term risk parameters will be used throughout this study when referring to this kind of portfolio parameters.

To illustrate the effects in the mortgage (default) portfolio due to macroeconomic changes, macroeconomic factors are linked to the default rate and the loss rate. In fact, the default rate and loss rate over time are constructed by (a combination of) macroeconomic factors.

Therefore, the main research question of this report is defined: *What is the influence of macroeconomic factors on the risk parameters for the mortgage portfolios?*

Subquestions describe the steps before the main question can be answered:

1) Which macroeconomic factors have influence on the mortgage portfolio?
2) What is the default rate under certain scenarios till December 2015, only based on the macroeconomic factors?
3) What is the loss rate under certain scenarios till December 2015?
4) What are the credit losses under certain scenarios till December 2015?

From a literature study, focusing on the Dutch housing market, a list of macroeconomic factors probably influencing the mortgage portfolio will be derived and forms the basis of the macroeconomic input factors. Later on, the impact of the factors on the DR or LR is determined [Subquestion 1].

The default rate is calculated without concentrating on the size of the mortgage, i.e. a default is defined as arrears of three months and is constructed by the best combination of macroeconomic factors [Subquestion 2].

The loss rate is the fraction of the total loan value that the financial institution loses on a default. This fraction is approached by macroeconomic input [Subquestion 3].

In the end, it is possible to calculate the expected total credit loss by multiplying with the total exposure in a financial period [Subquestion 4]. To answer this subquestion the multiplication of DR, LR and EV is made and is corrected for interdependencies is investigated.

1.3 Scope

In scope are three portfolios: (1) the Intermediary Channel, (2) the White Label and (3) the Consolidated Portfolio. The latter one includes labels (1) and (2) expanded with two small passive portfolios. Analyses are on aggregated level – one representative loan describes a portfolio (the average) - within the available time span. For (1) and (3) the reliable historical data is available from January 2003, for the White Label the data from November 2003 till today is in scope. The reason to choose these portfolios is based on the available information and the size of the portfolios. The Intermediary Channel is the largest portfolio and still active, the White Label is quite large too, but has a different audience.

The forecast period is from April 2011 up until December 2015. For extreme scenarios (stress test scenarios), the forecasts end on December 2012, because of the highly improbability of the scenarios over a longer horizon.

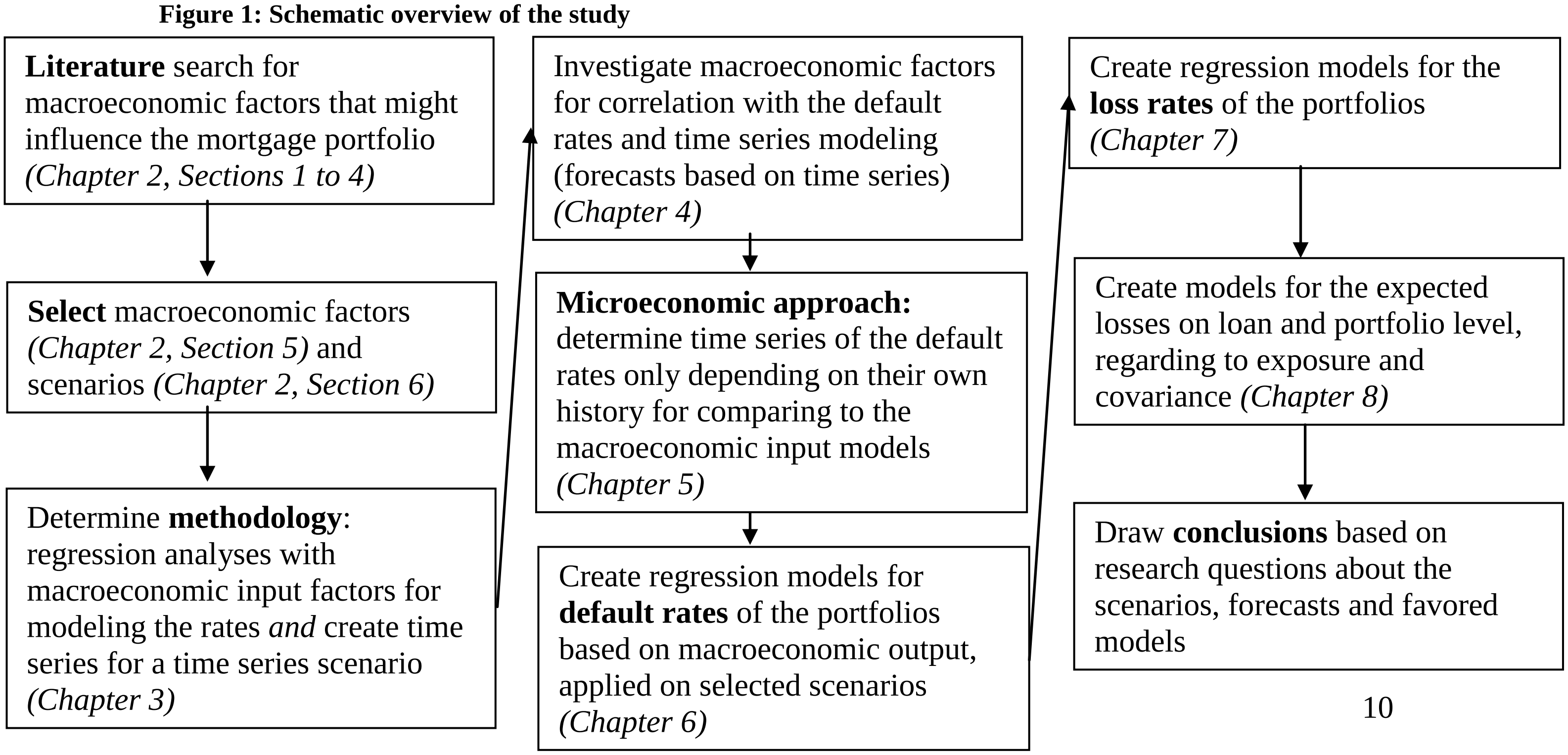
1.4 Outline of the report

The study starts with an introduction of the Dutch housing market, an international comparison, a preview of future developments and the risks related to mortgages *(Chapter 2)*. This literature study ends up with a list of factors that might influence the default portfolio of a financial institution. Several existing scenarios are selected for the analyses in this study.

In *Chapter 3* the methodology is described for all rates and conversions, including time series techniques and regression analyses. This chapter is a guide through the thesis. *Chapter 4* deals with the investigation and modeling of the macroeconomic factors (time series approach). In *Chapter 5* the same steps are applied for the default rates and related rates. The default rate part of the study ends up in predictions based on scenarios selected *(Chapter 6)*.

Chapter 7 constructs the loss rate by using portfolio variables Loan-to-Value and Loan-to-Income. Both are linked to a macroeconomic factor. In *Chapter 8* the study is completed by extrapolating the exposure of defaults and multiply by the default and loss rate.

Figure 1: Schematic overview of the study



2. A brief introduction of the Dutch housing market

To introduce the main underlying subject of the mortgage market, a short review of the actual Dutch housing market is presented. The review is based on recent sources and is restricted to focus on the Dutch housing market and mainly on resale houses *(Section 2.1)* and a comparison with other countries *(Section 2.2)*. A glimpse of the future is added to the short review, also based on recent literature *(Section 2.3)*. The risks related to mortgages are for customers as well as financial institutions described in *Section 2.4.*

The literature part, expanded with an unspecified question and answer session performed within the department, will end up in a list of important factors that could influence risk parameters of the bank *(Section 2.5)*. In *Section 2.6* scenarios are collected and selected. These scenarios are used throughout the whole study.

In *Sections 2.1 till 2.4* all potential causes of changes in the mortgage (default) portfolio are <u>underlined</u> to justify the list in *Section 2.5*.

2.1 The Dutch housing market: the odd one out

Although the government interventions are without doubt well-meant, these provide The Netherlands an exceptional position on the housing market in international perspective. A large number of relevant statistics concerning the housing market can be compared with other, especially western, countries, but in many cases a side note should be made and a corresponding link to government interferences is not rare. This part starts with the actual and most important problems on the housing market that come up over and over again in the recent years, which is an appropriate blueprint of the current situation, and will proceed in statistics and comparisons in an international context.

An expression often heard about the housing market, mostly plaintive, is the extreme tension on the <u>market</u>, mainly caused by the <u>government regulations</u> that can be characterized by stimulating the housing demand and restricting the supply. Stimulation on the demand side is designed by mortgage interest deduction for resale property – this was introduced in the late nineteenth century already - and individual grants and protection for rental houses. On the supply side are restrictions as the planning policy executed. The restrictive planning policy deteriorates the <u>affordability of houses</u>. A second restriction on the supply is the long construction procedure. The throughput time increased from 33 months of preparations to start a housing project in 1970 to 90 months nowadays. Another restriction is the micromanagement conducted by the municipality. Municipals determine the qualitative supply of houses, mostly on ideological motives with emphasis on social rental houses. Problem: there is no shortage of social rental houses, only a maldistribution (NVB 2008).

Other problems of the Dutch housing market are (1) a <u>lack of flow</u>, mainly caused by high transaction costs, that is cutting off the entrance of the bottom of the market, resulting in a forced demand of low quality houses, and (2) the current situation on the <u>housing market</u> that affects the broader <u>economy</u> negatively. Main points are the indebtedness of families, flexibility on the <u>labor market</u> and the waste of tax money (NVB 2008). The lack of flow and high transaction costs resulted in a decline of housing supply registered by Kadaster. In 2005, more than 560 thousand mortgages were registered, but in 2009 this number was below 260 thousand. The average mortgage amount increased in the same period from just above 203k to almost 250k. Because of the economic downturn and increasing unemployment, people postponed the purchase of a house (Van de Pas, L. 2010).

The precarious housing market with a low number of transactions, obviously initiates a high number of days houses are for sale before sold or taken off from the market. The average number of days a house is for sale increased over the last years dramatically, from far below the 200 days (2005) to well above 300 days (2009-2010). Nowadays detached houses are labeled as for sale on the internet for more than 400 days, townhouses and apartments around 250 days. In the previous 18 months about 30 percent of the houses with a value below 750k Euros was not sold on the market, for houses above 750k Euros the percentage was even higher, around 40 percent. Salient detail: the time to sale is shorter in cities with more than a hundred thousand residents (Dankers & Frank 2011).

Worth mentioning is the remarkable statistic of the high ratio of total outstanding mortgage debt and the Gross Domestic Product that is slightly below 100 percent, partly caused by mortgage interest deduction (NVB 2008).

2.2 International comparisons: an image outline

2.2.1 Development of house prices and affordability

Unlike what is often assumed, house prices in The Netherlands develops not significantly higher nor they are growing faster than prices in surrounding countries or other western states. The development of nominal housing prices of different countries is shown in Figure 2 (NVB 2008). House prices are strongly affected by government interferences, lower real interest rates and quality improvements, resulting in a decline of the affordability of houses in the previous year (ABN AMRO 2010). Woningmarktcijfers.nl suggests that general housing expenses are not unacceptable high (Van de Pas, L. 2010).

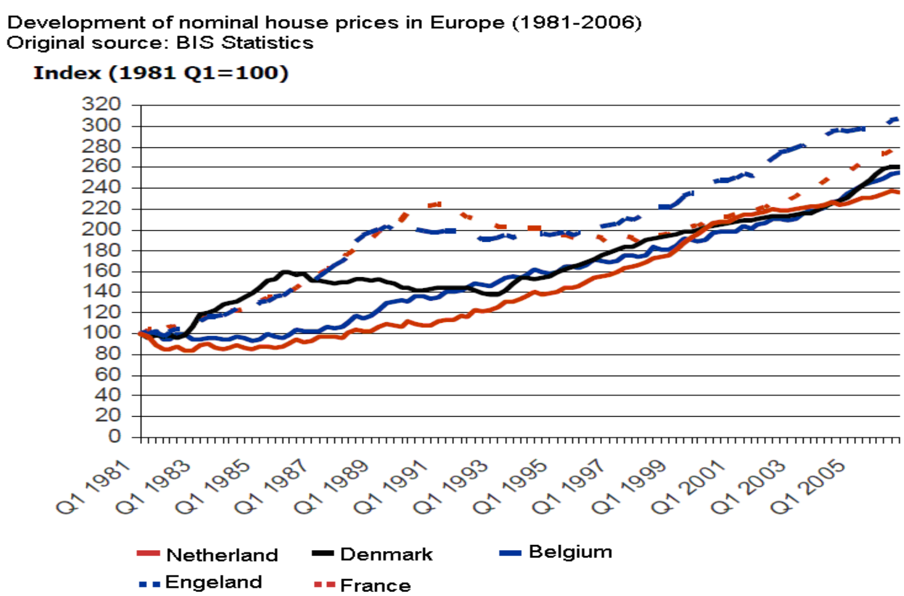
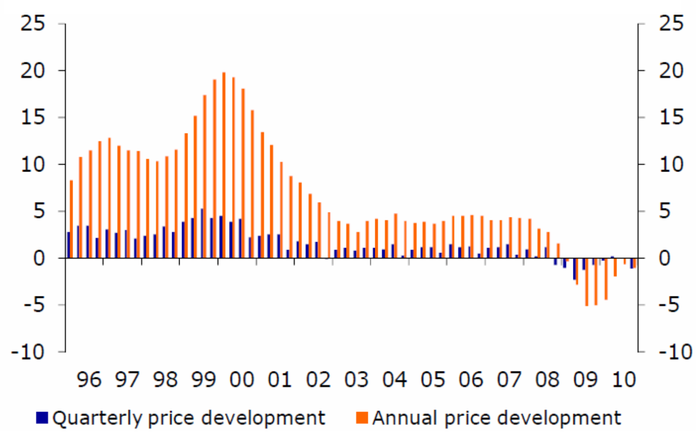


Figure 2: Development of nominal house prices in Europe (Source: NVB 2008)

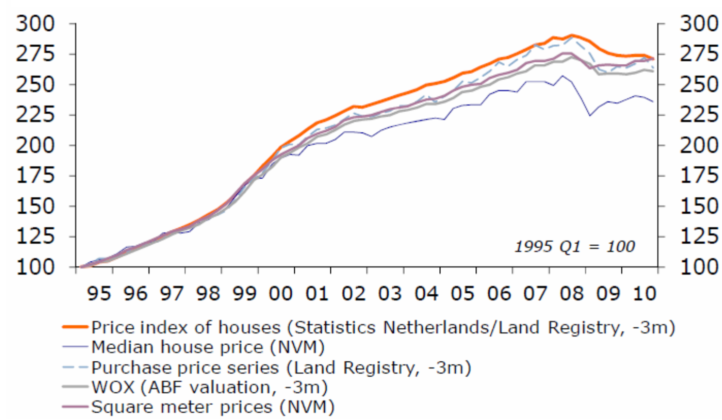
Real house prices are doubled over the last forty years in The Netherlands, which is not an unusual development in international context. Exceptions are Switzerland, Korea, Germany and Japan where prices raised significantly less (Van de Pas, L. 2010).

The latest figures show declining house prices, as can be seen in Figures 3 and 4. An obvious observation is the rarity of the recent decline in the last decades (Rabobank 2011). The WOX® house price index is an alternative Dutch index that describes the price development of the total inventory of houses, designed by ABF Valuation, a subsidiary of Calcasa.

House prices decline in fourth quarter
Original source: Statistics Netherlands



Price development of different indicators
Original sources: Statistics Netherlands, NVM, ABF valuation, Land Registry



Figures 3 and 4: House price developments (Source: Rabobank 2011)

The difference in house prices among the Dutch provinces is very large, although several lagging places are catching up in the recent years (Van de Pas, L. 2010).

In the most recent statistics is concluded that the two most important problems on the housing market – the lack of flow and affordability - still exist and it is even getting worse. House prices reduced slightly, forced by the low transaction rate. The number of deals is still reducing and the number of houses supplied keeps on increasing and the time to sale idem, mainly caused by the high house prices. The house prices in the last months are high in a historical as well as an international perspective. The price rose faster than the building expenses and the ratio of house prices and rent expenses and disposable income is high in comparison with surrounding countries. Structural factors explain the high prices. That fact indicates that there is no question of a bubble (ABN AMRO 2010, Rabobank 2011).

Not only the average demand price decreased over the last quartiles of 2010, the inflation-adjusted demand price growth turns negative for more than two years, starting with a price reduction of about 0,75% till over 4% in the last months of 2010. The nominal demand price continues with a reduction with the lowest value on the most recent data up to -3% (Dankers & Frank 2011).

Despite the negative trends is the number of foreclosure auctions not alarming. Most relevant causes of foreclosure auctions are divorces, financial mismanagement, frequent loan upgrading, buying on credit and unemployment (Van de Pas, L. 2010).

2.2.2 Supply, stock and house diversification

The total house supply increased enormously over the last 2 years, from 80.000 up till almost 160.000. In the intervening time, the diversification never changed significantly: apartments constitute the largest segment in supply, followed by detached houses and townhouses, respectively (Dankers & Frank 2011). A resulting problem is the extremely low elasticity of supply and the increase in housing shortage (Phanos Capital Group 2011).

The number of existing houses per 1000 inhabitants is low, compared to other western countries. Also the Dutch housing stock is tight and there is a great regional diversity, depending on the degree of urbanization and population density. The tightness of the stock could be a risk, because it appears that there is no buffer to prevent for upwards cyclical fluctuations and/or increasing demand (Phanos Capital Group 2011).

The Dutch private sector in the housing market is small. The regulated rental sector, including 2.4 million social rental houses, is by far the largest in the western world and the private almost the smallest. More than half of the social housing is situated in the provinces Noord-Holland and Zuid-Holland. In the Netherlands are about the same number of rental and resale houses, a slight advantage for resale houses. Phanos Capital Group concludes that, compared to other western countries, the Dutch housing stock is of high quality with a relatively low price (Phanos Capital Group 2011).

2.3 Looking forward: a glimpse of the future

Even though the housing market is hard to predict, a lot of documents are written about (near) future trends and expectations. It is worth mentioning some aspects about the Dutch housing market to give a short future look. Some caution is in order here, because it is a selection of sources and there are no guarantees that the future will be as predicted.

The downturn in the Dutch housing market will continue in the near future, according to Wegwijs.nl. The most important causes for continuing the downturn are mainly due to regulatory changes. The Authority for Financial Markets (AFM) will oblige customers of mortgages to make an additional repay in the first years and starters with strongly increasing wages are stronger restricted to scale to higher mortgages. Nibud initiated the reduction of mortgage payments that may be provided. Partly due to these stricter rules, the flow of houses will reduce and the inequality between the rental and second-hand housing market increases (Wegwijs.nl 2010).

The growth rate of the Dutch economy will slow in 2011 to barely 1.5 percent, according to the ABN AMRO Snapshot of the economy, as a result of the slowdown in the second half of 2010 and the cuts in government spending. The export will increase sluggish, private consumptions will grow and a positive investment activity is expected. Employment and consumer spending increase slightly (Kiene, N. 2010).

The volume of the Gross Domestic Product will increase, caused by increasing labor productivity, and the estimated grow is between 1.7 and 2.0 percent in the next years, according to the Dutch Statistics Office. The growth of labor supply reduces, but the participation is still increasing from about 75 percent just before the millennium change to over 80 percent. The labor productivity will increase by about 1.5 percent a year and the unemployment equilibrium is expected to be 6 percent in the coming years (CPB 2010).

The demand of housing will increase in the coming years, right ascending the growth of the number of households (Phanos Capital Group 2011).

In the nearest future the increasing unemployment rate is the most important risk factor on the housing market. A possible worry on the middle long term is the potential increase in real interest rates and adjustments of government policies on the demand side. On the long run, policy changes on the supply side can cause a potential risk (ABN AMRO 2010).

Financial institutions have to deal with the characteristics of the population, in the broadest sense. Forecasts are of great importance, especially about the composition of the population represented.

Because of the aging of the population, elderly will represent a greater proportion of the population and the proportion of young and non-retired residents population will reduce over the coming years (De Jong, A. & Van Duin, C. 2010). The life expectancy is still increasing (CPB 2010).

The number of births will remain below 200 thousand a year, but the number of deaths will increase sharply to over 200 thousand in 2040. Because the migration will stabilize on a slightly positive level, this will result in a future decrease of the population (De Jong, A. & Van Duin, C. 2010).

2.4 The risks of a mortgage

Despite the huge variety in mortgages, it is possible to indicate the most common (and obvious) risks, for the customer as well as the lender, in most cases the financial institution. In fact, all customers' risks are risks for the financial institution too, because financial problems of customers (could) lead to a default and losses for the financial institution.

The risks for consumers can be summarized by payment risk and risk of residual debt (equity risk) (DNB 2009). The payment risk can be described as the risk that the consumer is, on a certain moment, unable to pay the monthly mortgage payments, for example because of an increase in the interest rate or a fall in disposable income. Generally spoken, three situations might happen:

(1) An increase in the costs of living (increase of interest rates);
(2) Other expenses increase (inflation, government interventions, diseases or, for example, family extensions);
(3) A drop in income, caused by a reduction in working hours or job change, or, for example, in an unexpected situation such as a serious disease, divorce or unemployment (DNB 2009).

The above situations do not necessarily lead to payment problems, but the probability is higher when living expenses form a large part of the budget. This indicator is called the housing ratio, the ratio between housing costs and income. A higher quote means that the mortgage payments are a greater part of the income (DNB 2009).

Equity risk related to mortgages is defined as a general decline in the house prices, caused by a downfall in macroeconomic development – lower growth in national income, higher unemployment rate -, or an increase in pressure on house prices due to higher interest rates or stricter credit conditions (DNB 2009).

In case house prices fall, the real value of the house could become less than the mortgage debt, especially in non-overvalue mortgages. The difference is called the residual debt, which could be a risk. The indicator for the risk of residual debt is the Loan-to-Value (LTV) ratio on closing time. This ratio has increased over the recent years, as can be seen in Figure 5 (DNB 2009).

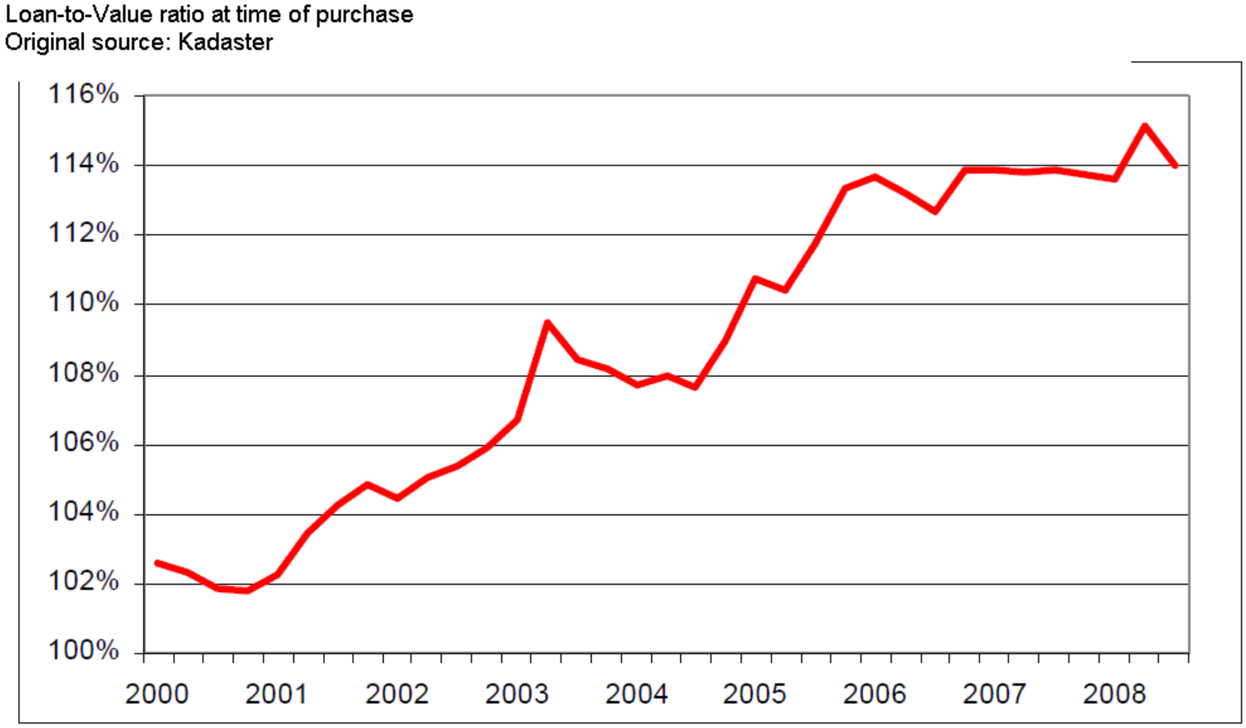


Figure 5: LTV-ratio on closing time (Source: DNB 2009)

Risks for consumers (payment and equity risk) are risks for the financial institution too, extended with generally smaller risk issues as prepayment risk, which is the uncertainty of available money, and quotation risk, that arises in the time between the offer and the acceptance in which the interest rate could change.

The risk for lenders is called credit risk and is split in customer related risks (payment and equity risk) and risks taken by the lender itself (prepayment and quotation risk). In the first situation, the risk can be summarized as the consumer is unable to repay (parts of) the loan to the lender and the lender is confronted with a loss on the loan. This risk can be translated in a Probability of Default (PD) and Loss Given Default (LGD). In the worst case a house ends up in foreclosure auctions. The number of houses auctioned is an indicator of forced sales. The increase in foreclosure auctions is shown in Figure 6 and the cohesion between the risks is drawn in Figure 7 (DNB 2009).

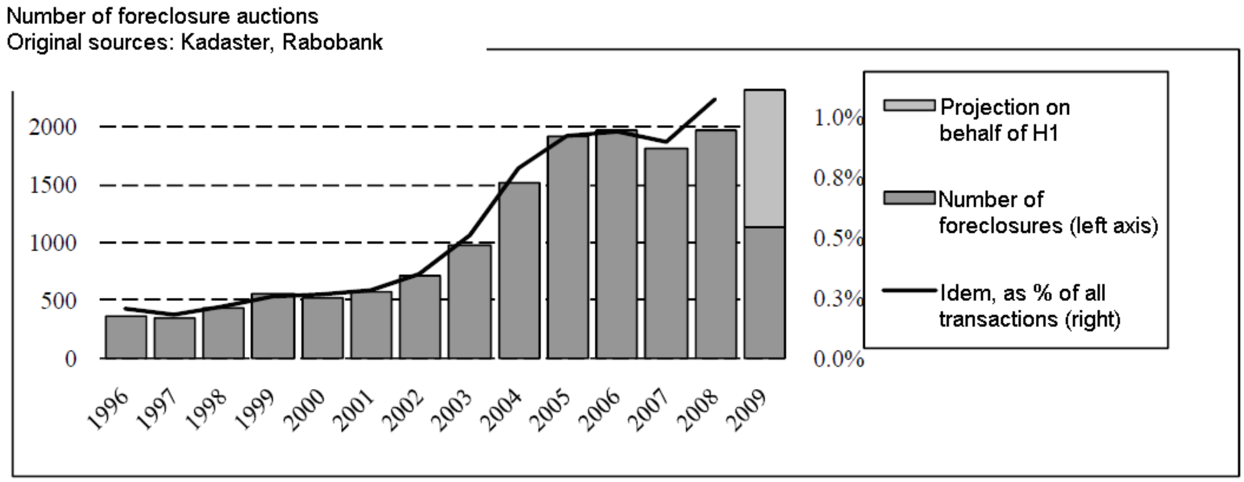


Figure 6: Number of foreclosure auctions (Source: DNB 2009)

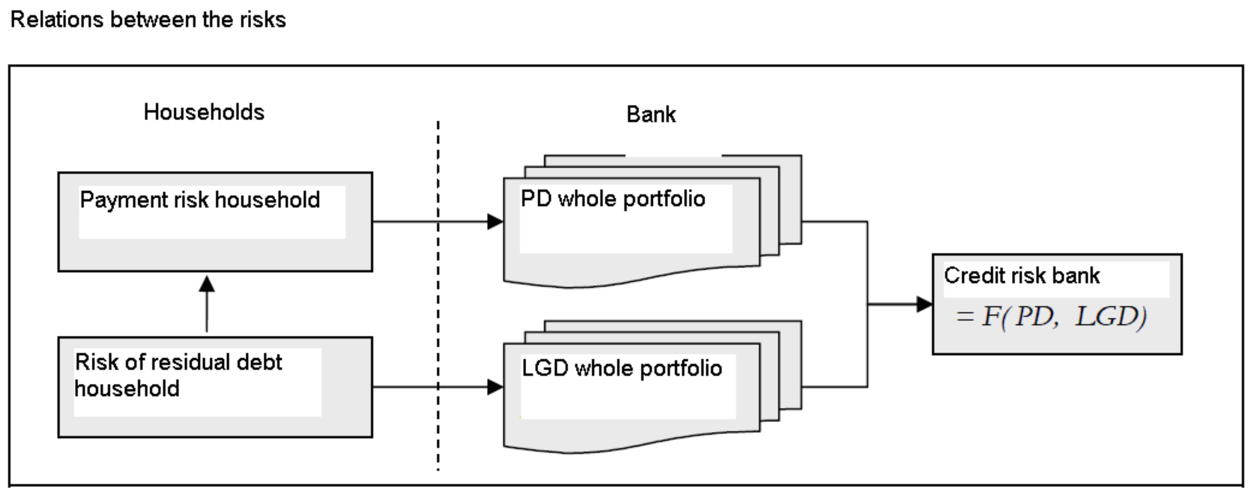


Figure 7: Relations between the risks (Source: DNB 2009)

2.5 Macroeconomic factors affecting credit risk

In the previous paragraphs was implicitly written about causes of defaults. In this section, a list of macroeconomic factors that might influence the credit risk will be drawn, started with the most obvious derived from literature used in the previous paragraphs (those are underlined in the text). At this stadium the fund market is added by reasoning that financial institutions should be able to hedge mortgages. Other obvious factors like internal measurements – think about product development and fraud - , are out of scope.

Below the underlined factors are summarized in categories that will be linked to external data sets.

- ✓ Diseases;
- ✓ Inflation;

- ✓ Interest/yield rates;
- ✓ Investing and trading;
- ✓ Fund market;
- ✓ General economy;
- ✓ Government regulations and interventions;
- ✓ Housing market, transactions and house prices;
- ✓ Population characteristics, including aging and growth/shrinkage;
- ✓ Social changes: divorces, family expansion/reduction;
- ✓ (Un)employment.

In the data collection process an important precondition has to be met: the macroeconomic factors chosen should be quantifiable on monthly basis with an easy-to-find indicator or real number. For each of the factors (categories) that might influences the mortgage portfolio, in databases of the Dutch Central Bank (DNB) and the Central Statistics Office (CBS) explanatory time series are gathered. This means that in each category one or more indicators that quantify the factor are collected and assessed as arguable. When no indicator is mentioned, no suitable time series is found. After an extensive search, the factors below are obtained. These factors are (1) an indicator of the category and (2) reasonable in terms of possible influences the mortgage portfolio. These time series are the starting point of the study and later on it will be tested if there is a relation with the mortgage portfolio.

- ✓ Investing and trading/Fund market/General economy:
  (1) Consumer confidence indicator, or
  (2) Economic environment indicator, or
  (3) Willingness to buy indicator.
- ✓ Diseases: (1) Occupational disability, and/or
  (2) Deaths by new diseases or deaths by heart infarcts.
- ✓ Inflation: (1) Inflation rate, or
  (2) Consumer price index.
- ✓ Government regulations and interventions:
  None, cause of other changes;
- ✓ Housing market: (1) Average house prices, and/or
  (2) The number of houses sold.
- ✓ Population characteristics, including aging and growth/shrinkage:
  Could be used for specifications of Credit Losses.
- ✓ Social changes: divorces, family expansion/reduction:
  None (measured on yearly basis)
- ✓ (Un)employment: (1) Registered unemployment, and/or
  (2) Unemployment rate.
- ✓ Yield rates: (1) 10 years bond yield rate, and/or
  (2) Mortgage yield for over 10 years.
- ✓ Some additions: (1) Number of accounts with a negative balance, and/or
  (2) Total value negative balances.

These factors are taken into account throughout the whole research, until good reasons will lead to exclusion. As will be described in *Chapter 4*, it can be argued for all of these factors to affect the mortgage portfolio, and thus the default portfolio, of a bank. Because of suspected redundancy, a maximum of one factor per category will be included in each model. This restriction is applied in the whole thesis. For example, registered unemployment can be subdivided in ‘total inflow’, ‘inflow from business activity’ and ‘other inflow’, or the same kinds of outflow. In- and outflow could cooperate, but are analyzed separately. Within the groups ‘inflow’ and ‘outflow’ the best fitting factor is chosen. This investigation of macroeconomic factors is the preparation of answering *subquestion 1* of *Section 1.2*. The whole answer will be presented in *Chapter 6*.

The data for these factors was derived from the website www.cbs.nl (Central Statistics Office) and the www.dnb.nl (Dutch Central Bank). Data from January 2001 until now are in scope for the default rate models, because for most of the selected time series this period is available and it is preferred to have more information before January 2003 (starting point of suitable data of the Intermediary Channel and Consolidated Portfolio) to be able to include time lags. Together with the scenario selection *(Section 2.6)* this is the complete set of external data used in this thesis.

2.6 Scenario selection

The overall goal is to analyze the effects of potential changes in the macroeconomics for the mortgage (default) portfolio. Therefore, scenarios are selected. For this thesis, the process of scenario creating is out of scope. Existing scenarios are collected and applied. An informed choice is to be the underlie for the selection of scenarios. For this project, the inclusion of different scenarios is useful to give insight in potential risks. Therefore, scenarios are selected for the short and middle long run.

The most important topic nowadays is the mortgage interest deduction (MID). Immediately total elimination of this deduction would lead to a house prices decline of about 18 percent and the number of houses sold will be reduced by nearly 1/3, according to ECORYS. Several workarounds are discussed in politics and a smooth (and slow) decline would only postpone the house price and transaction reduction. If the MID is partly abolished, the house prices reduction will be less (ECORYS 2005).

The tax rule change - the MID switches from box 1 to box 3 – would trigger an estimated reduction of 5 percent on the house prices and 8 percent on the total number of houses sold. The most important side effect is the unemployment increase. ECORYS expects that in the total abolish of the MID scenario about 60.000 people will lose their job (ECORYS 2005). The current unemployment is about 400.000 people. Derived scenarios are tabulated in Table 1.

Scenario	House prices	Transactions	Unemployment
Total abolish of MID	-18%	-30%	5,8%
50% MID deduction	-9%	-15%	5,4%
Tax Change: MID from box 1 to 3	-5%	-8%	5,2%

Table 1: Scenarios related to mortgage interest deduction (ECORYS 2005)

In current days, stress testing is a hot topic. Generally, for the Dutch housing market a Benchmark and an Adverse stress scenario are used. These scenarios are usually calculated over a two years scope.

Scenario	House prices	Unemployment	Interest rate (10 yrs)
Benchmark	0% (per year)	6% end 2011	4,1% end 2011
Adverse	-10% (per year)	7% end 2011	4,9% end 2011

Table 2: Stress scenarios for a two year scope[1]

Within the financial institution other scenarios are analyzed recently for scenario analyses. Those were verified by the economic office of the financial institution. The interpretation of the AFM scenario is applied over the time horizon up until the end of 2015 with constant unemployment (at 5 percent), the current housing market and a small reduction of the house prices (3 percent).

In the scenario designed by the Dutch Authority for Financial Markets (AFM) there are only small changes involved. In the Expected scenario, small increases are expected and in the interpretation of the D66 scenario, based on MID abolish spread over a long time and designed by a political association called Democraten'66, it seems to be in the middle of the 50% deduction and tax rule change. The economic office and risk management team of the financial institution added three scenarios: the Expected, the Yield Boost and the Unemployment Boost scenarios.

Scenarios 2015	Unemployment rate	Housing market	House prices	Interest rate
AFM	5%	+5%	+2%	
D66	5%	-10%	-8%	
Expected	5%		-3%	
Unemployment boost	6,5%	+2,5%	+1%	
Yield boost	5%		-5%	+1,5%

Table 3: Other scenarios (used in earlier analyses within the financial institution) of the end of 2015

Unfortunately, in different scenarios different macroeconomic factors are involved. Therefore a mathematical time series model is introduced. The undefined macroeconomic factors in scenarios are replaced by time series of the factor. Therefore an extra scenario is added: a scenario based on time series only, the so-called Time Series Scenario.

Scenarios are tightened to make them useful for calculations in scenario analyses. More or less based on rationality, consistency and variety, the scenarios and assumptions as in Table 4 will be used throughout this thesis. The 50% deduction scenario is eliminated by means of redundancy. These scenarios are used in all models.

The so-called Time Series Scenario can be re-engineered and therefore added in the end (determined in Chapter 4). To get a general overview of all scenarios, the Time Series Scenario is already included in Table 4. This scenario is based on time series models. So, the time series of the unemployment rate, housing market, house prices and interest rate are forecasted based on the own time series with time series modeling techniques (ARIMA models). When evaluating the changes realized in the end of 2015 compared to the current value, the scenario can be described in this table. The extensions of the stress test scenarios are added to create a scenario to the end of 2012. The extension is determined by a same increase as in the period before.

[1] http://www.dnb.nl/openboek/extern/id/nl/ki/40-198319.html

Scenarios	Unemployment rate	Housing market	House prices	Interest rate
S1: Adverse (Stress)	7% linear increase (end 2011) *Extension: 9% (end 2012)*	Time series (end 2012)	-10% linear decrease (end 2011) *Extension: Another -10% (end 2012)*	+1,5% linear increase (end 2011) *Extension: Another +1,5% (end 2012)*
S2: AFM	5% constant (current rate)	+5% linear increase (end 2015)	+2% linear increase (end 2015)	Time series (end 2015)
S3: Benchmark (Stress)	6% linear increase (end 2011) *Extension: 7% (end 2012)*	Time series (end 2012)	Current rate	+0,7% linear increase (end 2011) *Extension: Another +0,7% (end 2012)*
S4: D66	5% constant (current rate)	-10% linear decrease (end 2015)	-8% linear decrease (end 2015)	Time series (end 2015)
S5: Expected	5% constant (current rate)	Time series (end 2015)	-3% linear decrease (end 2015)	Time series (end 2015)
S6: MID abolish	5,8% linear increase (suppose end 2015)	-30% linear decrease (suppose end 2015)	-18% linear decrease (suppose end 2015)	Time series (end 2015)
S7: Tax Change (MID: box 1 → 3)	5,2% linear increase (suppose end 2015)	-8% linear decrease (suppose end 2015)	-5% linear decrease (suppose end 2015)	Time series (end 2015)
S8: Unemployment boost	6,5% linear increase (end 2015)	+2,5% linear increase (end 2015)	+1% linear increase (end 2015)	Time series (end 2015)
S9: Yield boost	5% constant (current rate)	Time series (end 2015)	-5% linear decrease (end 2015)	+1,5% linear increase (end 2015)
S10: Time Series Scenario (Calculated in Chapter 4)	Current rate	Current rate	+7,8% time series increase (end 2015)	Current rate

Table 4: Scenario selection for this study

3. Methodology: building blocks of the forecasting model

This chapter illustrates an overview of the model development, starting with notes about the available data in *Section 3.1*. In *Section 3.2* a roadmap is sketched, intended to structure the development of the models in the report. In *Section 3.3* the use of time lags is explained and in *Section 3.4* regression is described. In *Section 3.5* the ARIMA time series model is introduced, including tests, selection and evaluation. In *Section 3.6* the assessment of the macroeconomic models compared to the microeconomic models is elaborated.

3.1 Data collection

For this research internal data from the data warehouses of the financial institution and external data from sources describing the Dutch economy is required. The internal data mining process was primarily done by using SAS® Enterprise Guide to collect, sort, combine and create necessary data taken from the data warehouses within the financial institution's environment. The same software is used to obtain the results. All external data, i.e. time series of macroeconomic factors and scenarios, is selected in *Sections 2.5 and 2.6*.

The available data is on monthly basis and includes a lot of characteristics for each loan, such as the age of the applicant, test income, total value of the mortgage, monthly payments, weighted interest rate, default or recovery date eventually, and a lot of derived factors. For most studies in this thesis, a combination of different data files is made, including selections.

3.2 Roadmap of methodology

This roadmap is intended to give global insight in the followed route to come to the models and functional design of the report for the default rate and loss rate based on macroeconomic factors. This paragraph forms a short overview of the methodology. For several steps there will be referred to another paragraph for the extensive description.

The default rate (DR) is estimated by a combination of macroeconomic factors. The DR is the fraction of the total number of loans with more than three months arrears, calculated for each financial period (i.e. monthly). The fraction line based on the logistic fraction of defaults is constructed by macroeconomic factors, only. This means that no conventional approach is applied; only a combination of macroeconomic factors and an autoregressive term are included in the model for the default rate of the portfolio. See Figure 8.

The loss rate is based on Loan-to-Value- and Loan-to-Income-ratios connected to macroeconomic factors, house prices and unemployment respectively. The LR is defined for defaults as the average loss fraction with respect to the total loan value. A cross-table is drawn to assess the dependency of the loss rate on LTV- and LTI-ratios (divided in classes). This approach is time independent and less useful for forecasting. Therefore, the LTV-ratio can be converted into a loss rate, which makes it possible to create a time dependent loss rate.

The total credit losses are calculated by multiplying DR and LR with the exposure value of the portfolio. The exposure value is determined by extrapolating the trend (so, this risk parameter is calculated without involving macroeconomic factors). Only one issue has to be solved: when multiplying DR and LR, it is assumed that these are uncorrelated, but it would be reasonably to check if there is a substantial correlation involved.

3.2.1 Default rate

The default rate is estimated by macroeconomic factors directly, based on a logistic regression function, which means that all involved factors are multiplied by a separate calculated parameter and a constant and autoregressive term are added. The default rate is first written as logistic factor and afterwards rewritten as rates. This is called logistic linear regression and the general formula used is written in Equation 1, with y indicating the time lag of the factor. The constant and parameters are calculated by minimizing the sum of the absolute errors between model and realized values. The absolute error is favored to the squared error, because of the relatively lower weights on outliers.

$$ln\frac{dr}{1-dr}(t) = constant + parameter1 * macroeconomic\ factor1\ (t-y) + \ parameter2 * macoeconomic\ factor2\ (t-y) + \ \ldots + \ ln\frac{dr}{1-dr}(t-1),$$ **[Equation 1]**

Suppose that the three macroeconomic factors are included, with parameter 0,0001 for macroeconomic factor 1, 0,0004 for macroeconomic factor 2 and -0,0005 for macroeconomic factor 3 with a constant of 0,04. Then the DR can be constructed point for point as is graphically shown in the Figure 8.

The best fitting parameters are determined by minimizing the absolute error between model DR and realized DR. The formula is expressed in Equation 2.

$$\min|DR(realized) - DR(model)\,| = \min \textstyle\sum_{t=0}^{n} |DR(realized)(t) - DR(model)(t)|,$$ **[Equation 2]**

Before regression analyzing, hypotheses about the expected influence of each macroeconomic factor on the mortgage portfolio are formulated. Factors are tested one-on-one for the best time lag with the DR, with regard to the hypothesis for negative/positive correlation. The reason is that, for example, negative developments in the housing market could not have a positive effect on the mortgage portfolio. Positive correlation would be counter-intuitive. A visualization of time lag shifts is expressed in Figure 9. This visualization is based on a negative correlation (=hypothesis): if the macroeconomic factor goes up, the rate goes down. If the hypotheses was to find a positive correlation, both lines would have to move in the same direction. Note that time lags are determined before the regression analyses is performed.

The best time lag is observed by the highest or lowest correlation coefficient for lag zero up until a lag of 24 months, according to the hypothesis. Observing a lag of zero means that the macroeconomic factor series of January 2003 till now corresponds best with the default rate series of January 2003 up until now. A lag of, for example, 12 months means that the macroeconomic factor series of January 2002 up until a year ago (from now), corresponds best with the default rate series from January 2003 till now.

Because scenarios are predicting the house prices, housing market, interest rate and unemployment rate *(Section 2.6)*, it will be tried to include these four factors or suitable substitutes. To fit the model well, sometimes other factors have to be included. When it is impossible to create a proper model with those four factors, one or more have to be eliminated. The best combination is determined by the average percentage error over the last year, when the same model was created a year ago. So, the last 12 months are forecasted and the average of the absolute error divided by the realized rate, decides for the best combination.

With the time lag and best combination of factors, the logistic line of the DR will be constructed. Scenarios are implemented by defining the end value and the end time and drawing the line from the current value to the scenario end value (see Figure 10). Suppose, a time lag of 3 months is observed, the scenario end value is 3 months postponed (because actual data influences the third month from now on).

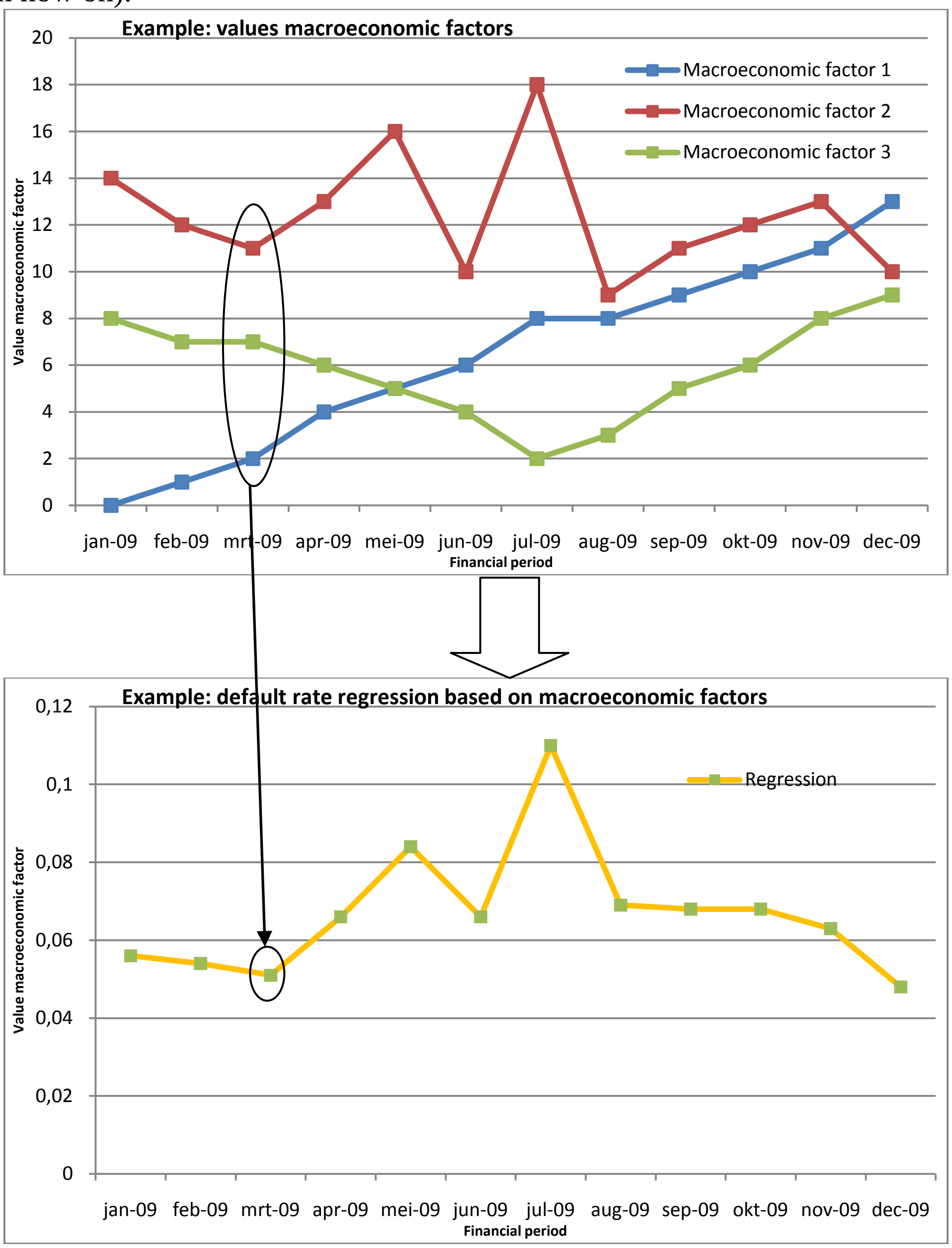


Figure 8: Default rate construction

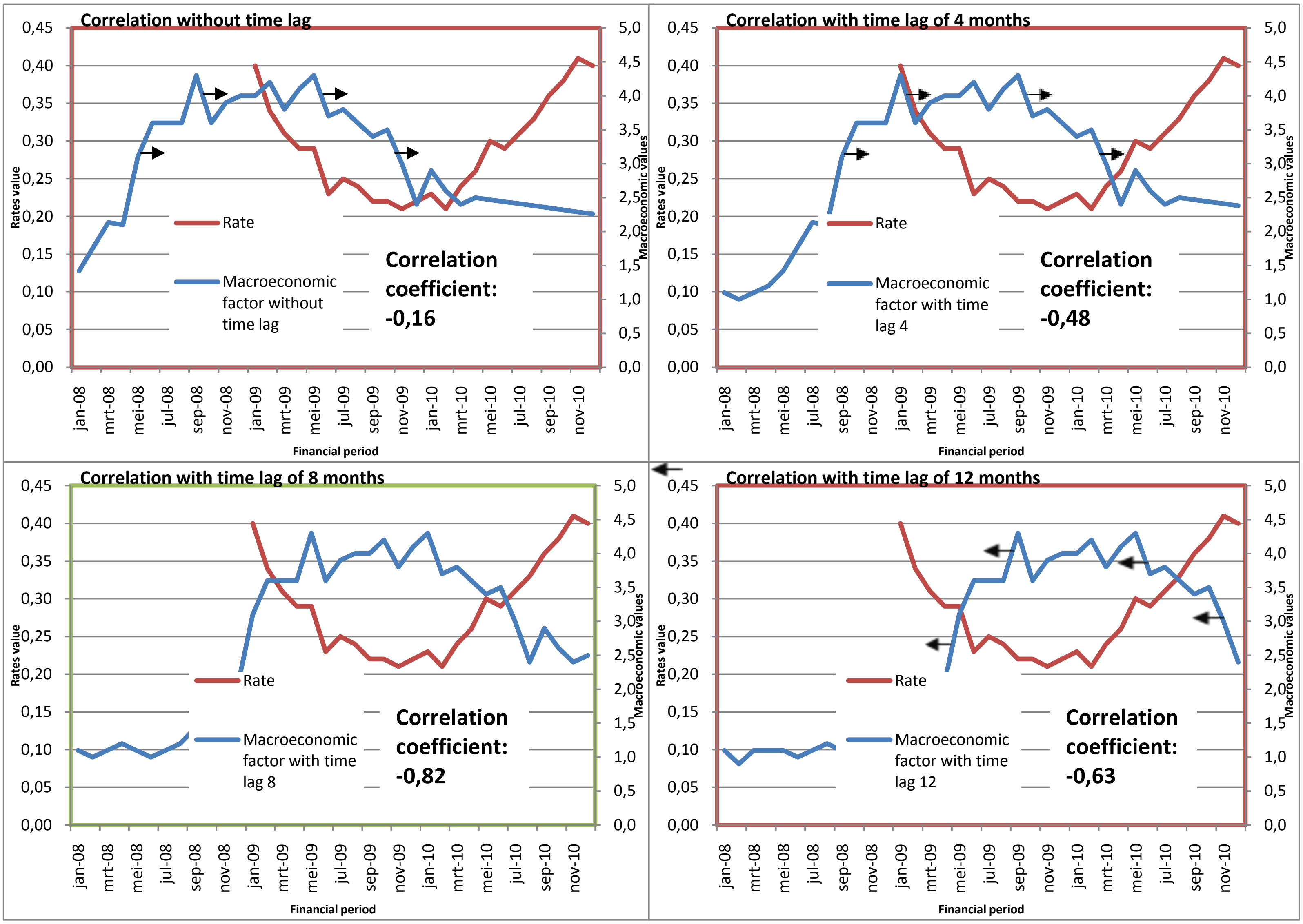


Figure 9: Time lag shift, graphical example, the green lined graph indicates the best fit out of these 4 trials

Besides the scenario forecasts, a Time Series Scenario is created too. All macroeconomic factors are tested for suitability for forecasting. The best-fitting logistic regression is determined. The time series of the underlying factors of the DR are forecasting till the end of 2015. Besides the nine selected scenarios, the created Time Series Scenario is added.

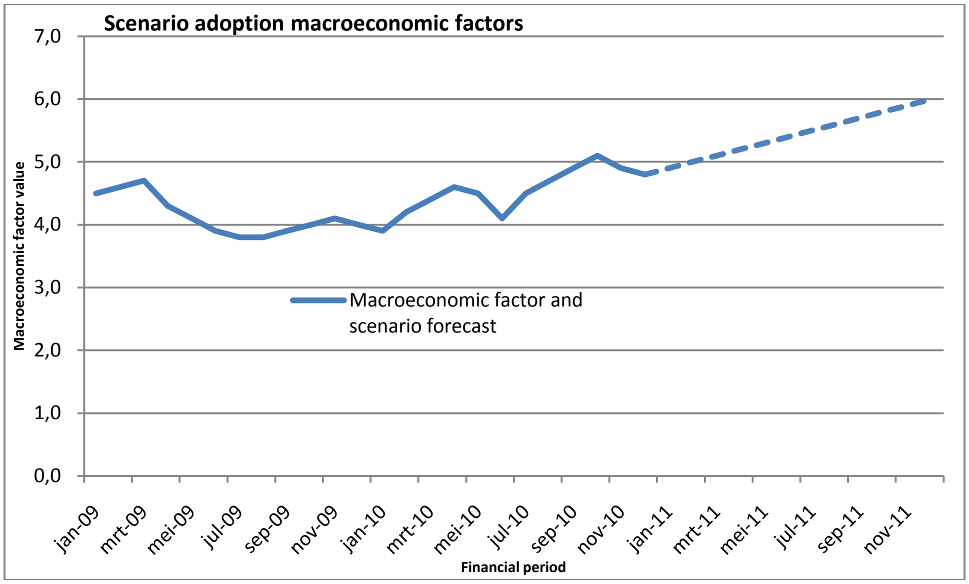


Figure 10: Macroeconomic factor forecasting on scenario interpretation (values 2009-2010 are known and 2011 is forecasted as linear increase to scenario defined value of 6 in December 2011)

Default rate:

1) Set **hypotheses** describing effects of macroeconomic changes on the mortgage portfolio;
2) Determine **time lags** for input factors *(extension in Section 3.3)*;
3) Apply **logistic regression** and look for the best combination of macroeconomic factors *(extension in Section 3.4)*;
4) Implement **scenarios** with respect to the time lag and regression parameters;
5) *For the Time Series Scenario: Determine **suitability** of macroeconomic factors for forecasting and select the best fitting ARIMA-model (extension in Section 3.5):*
    a. *Stationarity: Augmented Dickey-Fuller test;*
    b. *Time series approach: ARIMA time series models;*
    c. *Best fitting ARIMA(p,d,q)-model: Schwarz info criterion;*
    d. *Residuals test for normality: Ljung-Box test.*
6) **Add the Time Series Scenario** to the selected scenarios and publish the results.

Exactly the same methodology is applied on the (1) inflow rate (IR), (2) recovery rate (RR) and (3) foreclosure rate (FR). The inflow rate is the number of new defaults divided by the total number of loans. The recovery rate is the number of loans that were in default, but now meet payment obligations (or at least less than 3 months in arrears), with respect to the total number of loans. The foreclosure rate is the non-recovered fraction of defaults, in most cases the house is sold in foreclosure auctions. All are defined as a fraction of the total loans in the financial period and results can be found in the Appendices (references in text).

3.2.2 Loss rate

The loss rate is a rate that describes the loss fraction on a default with respect to the total loan value. The loss rate is linked to portfolio variables (such as Loan-to-Value) that divide the LR into classes or the variable (LTV) is rewritten as the loss rate. Drawing cross-tables based on

LTI- and LTV-classes is a time independent approach, because the available data does not show a reliable month to month LR, due to a lack of data. The loss rate values directly derived from the LTV is time dependent and therefore preferred.

The portfolio variables are linked to macroeconomic factors to analyze the loss rate in different scenarios *(Section 2.6)*.

Loss rate:

1) **Link** loss rate with a **descriptive variable** of the portfolio (choose variable);
2) **Link** descriptive variable with a **macroeconomic factor;**
3) Determine **time lag** macroeconomic factor on descriptive variable *(extension in Section 3.3)*;
4) **Determine parameters** model (input macroeconomic factor, output descriptive variable) *(extension in Section 3.4)*;
5) **Determine parameters** model translation or draw table (input descriptive variable, output loss rate);
6) **Write** loss rate dependent on macroeconomic factor or publish result table;
7) Combine variables in a cross-table to investigate **loss rate** (time independent) or rewrite variable (LTV) as loss rate (time dependent);
8) Determine or calculate **scenario consequences** including Time Series scenario (see default rate);
9) **Publish parameters** of the LR.

3.2.3 Exposure value and credit losses

The exposure value (EV) is not a fraction, but a value in Euros. This value is simply extrapolated, because no explicit reason to connect the exposure to macroeconomic factors can be imagined.

The credit losses are calculated by multiplying DR, LR and EV. This multiplication can be an under- or overestimation due to correlation between DR and LR. Therefore, the covariance is calculated and the multiplication of DR and LR will be corrected for the covariance value, if necessary.

$$Cov(DR, LR) = E(DR * LR) - E(DR)E(LR) \qquad \textbf{[Equation 3]}$$

Because no individual default rates are calculated, the covariance cannot be calculated on loan level. Therefore a moving average over 6 periods for the DR and LR are used for calculating the covariance. So, the covariance for DR and LR values from January till June 2003 is calculated. This covariance is assumed to be the covariance on June 2003. Then, the covariance for DR and LR values from February till July 2003 is calculated. This covariance value is assumed to be the covariance on July 2003, and so on. The covariance value is added to the multiplication of the expected DR and LR values.

3.3 Time issues and correlation macroeconomic factors

Since the focus is on consequences of macroeconomic changes for the default portfolio, time issues concentrate on delay effects. A default is noticed in case of three months arrears.

Therefore, it could be argued that the time lag is at least three months. On the other hand, causes are in recent history and therefore, a lag of zero will be in scope (and adapted common).

Because of the restricted time horizon available, two years lag is taken as the maximum for default rates models. The correlation coefficient is calculated for period lags from zero up until twenty-four months. The lowest/highest coefficient depending on the hypothesis is the best fitting lag. For example, an increase in unemployment should correlate positive with the increasing defaults (and negative with the recovery rate), because this kind of relation is expected and in scenarios the opposite relationships will be unreliable and not plausible. Graphical expression in Figure 9.

3.4 Regression

Common regression techniques are used for creating models for the default rates and related rates defined as (1) default rate at moment t for each label, (2) inflow of defaults at time t for each label, (3) recovery of defaults at time t for each label and (4) outflow (no recovery) of defaults at time t (called Foreclosure Rate). Also, the loss rate deals with regression by connection to a macroeconomic factor.

The goal is to construct a line similar to the risk parameters and design forecasts based on selected scenarios including the Time Series Scenario. For rates between 0 and 1, a common approach for the treatment of rates is to use logistic regression, which in fact means that the rates are rewritten before linear regression on the available data is performed (and afterwards re-engineered as rates). The conversion is calculated as

$$ln\frac{p}{1-p} = a + bx, \qquad \textbf{[Equation 4]}$$

where p is the rate (in this thesis the default rate), a the trend and b the parameter in the regression of the rate (Pagano 1996). A logistic conversion is used for DR, IR, RR and FR.

All macroeconomic factors and all rates are multiplied by 10^x to converse all time series to the same size order. For example, house prices of 200.000 Euros are written as 2,00 in the parameter estimation phase. In the final formula the conversion will be restored. So, in the end it is possible to fill in 200.000 Euros because a factor 10^{-5} is added.
The linear regression on the raw data:

$$Rate\ on\ time\ t = c + \left(\sum_{i=1}^{n} a_i F_i\right) + \epsilon, \qquad \textbf{[Equation 5]}$$

with constant factor c and parameters a_i for each included macroeconomic factor F_i calculated for each month t with an undefined error ϵ. Notice that the rate is the logistic conversion of the original rate for DR, IR, RR and FR and therefore afterwards re-engineered to rates.

For example, a certain DR is based on house prices (HP) with parameter a_1, the interest rate (INR) with parameter a_2 and constant c. The formula: $ln\frac{DR}{1-DR} = c + a_1HP + \ a_2INR + \epsilon$.

Because time lags are used, the regression function includes a time lag y:

$$Rate\ on\ time\ t = c + \left(\sum_{i=1}^{n} a_i F_i\ (t-y)\right) + \epsilon, \qquad \textbf{[Equation 6]}$$

Parameters found are used to fill in for the scenario analyses, including the Time Series Scenario. In the latter case, predictions can be formulated as in the equation below.

The predictions are based on the ARIMA(p,d,q) models, including time lags:

$$Forecast\ rate\ on\ time\ t = c + \sum_{i=1}^{n} a_i ARIMA(F_i(t-y)) + \epsilon\ on\ time\ t. \qquad \textbf{[Equation 7]}$$

This process for creating the regression models was based on the preferred factors, which are actually preferred only because of these factors define the scenarios selected: (1) house prices, (2) housing market, (3) interest rate and (4) unemployment rate. First those factors were tried, and when the result was not satisfying, more factors were included and if no good solution could be found, one or more of the preferred factors was eliminated. The decision of satisfying the model is based on some criteria:

(1) When using the same factors, the parameters calculated without the last year's data are close to the actual parameters. For example, in the actual situation unemployment gets a parameter 0,40, but in the situation without last year's data the parameter would be only 0,01, then this regression is inappropriate. As a rule of thumb, an difference of over 10 percent is considered as too much.
(2) The best regression combination is based on the smallest error in the validation period. The validation period is last year. So the parameters calculated without last year's data are used to forecast the last year. The solution with the smallest average absolute error over the last twelve financial periods is called the best fitting regression.

To determine the best fitting parameters, the method of minimizing the sum of the absolute errors is used. This means that the parameters are determined by minimizing the error (absolute difference between the real value and model value).

3.5 Time series analyses

A common approach before using time series models is to look at the shape of the research curve, plotted over time. In a lot of economic factors a kind of trend (mostly upwards) can be derived from the figure. Two other often observed patterns in time series are seasonal shapes, for example the selling of ice creams, and cyclical shapes, such as business cycles. A combination of those main shapes is absolutely not rare, for example imagine a cyclic pattern with a trend. All other shapes could be called irregular variations. This does not necessarily mean that it is not possible to analyze and forecast these time series (data).

This research focuses on forecasting the risk parameters in the near future, based on historical data and predictions of (strongly) correlated macroeconomic factors. To predict risk parameters without external observations, a univariate model, such as a simple autoregressive model (AR), could be used. Autoregressive terms refer to older data in the same time series. These predictions will be calculated and drawn only to show whether a benefit of using the macroeconomic factors in the forecasts exists *(Chapter 5).*

In first approach the ordinary multivariate models, in which forecasts also depend on other predictions, are used. In case of linearity an autoregressive integrated moving average model (ARIMA) is used (Lee 2010). This is a time series model based on autoregressive terms and moving average terms on the errors.

A successful implementation of the forecast models benefits from an easy-to-use mathematical expression. Therefore, other possible time series were eliminated. The focus is on the ARIMA model.

3.5.1 ARIMA time series

Autoregressive integrated moving average models (ARIMA) are time series based on three pillars, which are all in the name of the model. The autoregressive part means that, the model is based on the *p* previous values of the time series, called the order of the time series and in fact is the number of parameters in the model. This historical data x_t can be multiplied by an unknown parameter ϕ_j that could be estimated by, for example, a least-squares or least-absolute error approach. In the AR model a residual error ε_t based on the signal value t is added (Figueiredo 2011). Usually a constant factor c is included in the model.

$$x_t = c + \sum_{j=1}^{p} \phi_j x_{t-j} + \varepsilon_t \qquad \textbf{[Equation 8]}$$

The term integrated in ARIMA can be interpreted as the generalization of the autoregressive moving average model (ARMA). The MA part is the moving average over the residual errors, usually white noise, of the q previous values. The (theoretical or empirical) average is added as a constant factor.

$$x_t = \mu + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t \qquad \textbf{[Equation 9]}$$

The ARMA model is the (summed) combination of the AR(p) and MA(q) model. The both constant factors are together called c (Alexander 2001, for time series modeling).

$$x_t = c + \sum_{j=1}^{p} \phi_j x_{t-j} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t \qquad \textbf{[Equation 10]}$$

By integrating the ARMA model, the ARIMA model arises. The ARMA (p,q) model is evolved in an ARIMA (p,d,q) model, where d is also an integer and often referred to order of differencing.

$$\phi(B)\nabla^d(\mathrm{x_t} - \mu) = \theta(\mathrm{B})\varepsilon_t, \qquad \textbf{[Equation 11]}$$

Where $\phi(B) = 1 - \sum_{j=1}^{p} \phi_j B^j$ and $\theta(B) = 1 - \sum_{i=1}^{q} \theta_i B^i$ and $\nabla = 1 - B$ and B is a backward shift operator (Khashei 2011).

At first sight, the ARIMA (p,d,q) model seems to be a lot more complex compared to the ARMA (p,q) model. In fact, the parameter d that was added is only necessary for stationarity reasons, because the value of d indicates the number of times differences between successive observations should be included before a stationary time series is obtained (SWOV 2010).

3.5.2 Assessing and creating models: preparation, selection and regression

The main objective of this study is the ability to forecast risk parameters, more or less summarized in the size of the default portfolio and changes in the IFRS defaults portfolio corresponding to credit losses. The financial institution uses the SAS® (Base and Enterprise Guide) software for data analysis and storage. These both facts will be treated as preconditions, mainly affecting the choices of methods. The results in this research are achieved by using the available SAS® Enterprise Guide software (with time series add-in) and calculations in Microsoft Excel.

Stationary time series: randomizing the errors

The selected time series have to be (or be made) stationary to make suitable predictions based on time series techniques. Stationarity is also a constraint for using most data analyzing time series models. This reasoning is clarified by referring to the described ARIMA(p,d,q) model. When trend and seasonal/cyclical influences are removed, the remainder of the time series has to be stationary, which means independent of time lag, to construct a proper forecast. If the remainder is not stationary, the first differences would be tried. Imagine that the remainder terms are non-stationary; the forecasts would not say anything reliable. Note that these restrictions are not influencing the use of regression analyses. Stationarity is not a hard constraint for regression, because no shape (trend, cyclical,..) has to be reconstructed.

Time series are stationary or non-stationary. If the time series is considered to be non-stationary, the order of differencing will be adjusted. The same process will be completed, but now for the differences of the observed values. Is the result non-stationary again, the difference of the difference will be tested. This is called the second order of differencing. The lowest order for creating a stationary time series is important, because it represents the *integrated* term in the ARIMA(p,d,q) model. More specifically, in the ARIMA(p,d,q) model, the *d* is determined in this phase.

Several appropriate tests are used in econometrics to determine stationarity in time series, the so-called unit root tests. The Augmented Dickey-Fuller (ADF) test will be chosen for this analysis. Although in this study the SAS® Enterprise Guide function will be used to find the results and assess the need for a higher order of differencing, the concept of the ADF test should not be left out. This unit root test is based on the simple autoregressive model with *k* lags, AR(k). The choice of lag order *k* is an important issue, because the test statistic can be sensitive to this choice (Cheung 1998).

The mathematical approach is to set the time series for time t, with the following expression of the regression:

$$\Delta x_i = \mu + yt + \alpha_1 x_{t-1} + \sum_{j=2}^{k} a_j \Delta x_{t-j+1} + u_t, \qquad \textbf{[Equation 12]}$$

with t = 1, 2, …, T, Δ is the difference operator and μ_t is a white noise innovation. The unit root test tests for stationarity by testing for $\alpha_1 = 0$, the alternative hypothesis is $\alpha_1 < 0$ (Cheung 1998).

If the null hypothesis cannot be rejected, a unit root exists. Rejection depends on the comparison of the t-statistic of the test

$$\tau = \frac{\alpha_1}{\sqrt{var(\alpha_1)}}, \quad \textbf{[Equation 13]}$$

and the critical value of the Dickey-Fuller t-distribution, depending on the confidence interval and number of observations. If the result is higher than the critical value, a unit root exists. After removing the observed unit root (constant, trend, seasonal effects), the test will be exercised until no unit roots exists or no stationary solution can be found (then the test will concentrate on the first (or second) order of differencing). The remainder is assumed to be normally distributed with zero mean, known as white noise (Roubos 2008). It is worth to take a look at these errors and the distribution and test or assess for normality in the model building process.

Selection of the most appropriate model

After the process of preparation and interpretation has ended, the integrated factor *d* of the ARIMA(p,d,q) model is known for each macroeconomic factor and default indicator. Then the search procedure to estimate the most proper model starts. The p and q are to be estimated for each individual factor. For each *p* between 0 and 3, at least the combination with three *q* values (0, 1, 2) are analyzed. If there are reasons to think that a higher term improves the solution, those combinations will be added.

Model selection will be based on the Schwarz information criterion (SIC) in favor of the four years earlier presented Akaike information criterion. This choice is based on the objective of the study. Schwarz information criterion will tend to favor lower dimension models when more than eight observations are included. For this study, also quite weak relations are included and higher dimensions can lead to worse results. Therefore, the SIC is used for the selection phase (Koehler 1988).

The SIC model favors the model that minimizes

$$-2\ln\,(M_j\big(Y(1),\ldots,Y(n)\big) + k_j \ln(n), \quad \textbf{[Equation 14]}$$

with Y(i) as the observations, M_j (Y(1),…, Y(n)) defining the maximum value of the likelihood for the j^{th} model and k_j representing the number of free parameters (Koehler 1988). The estimation of parameters will be done by the well-known Yule-Walker equations. The consequence is the probability of dealing with non-converging parameters. A not converging model will be left out and indicated as ‘no solution’.

Evaluation of the models

The models will be evaluated by the Ljung-Box test:

$$Q = n(n+2)\sum_{h=1}^{m}\frac{\rho_R(h)^2}{n-h}, \quad \textbf{[Equation 15]}$$

with $\rho_R(h)$ for the autocorrelations of the residuals of lag h, m for the number of lags tested and n for the number of observations included.

The Ljung-Box test tests for white noise on the corresponding Chi-square value with m degrees of freedom (null hypothesis). If the null hypothesis of appearance of white noise is rejected, the model is inappropriate to forecast the macroeconomic factor or risk parameters (Roubos 2008).

The lag choice is very important, and as small as possible for the confidence of the model, for example LN(n), with n observed values. Therefore, a lag of 6 will be evaluated, corresponding to half a year of data. The null hypothesis is independent observations. This hypothesis will be tested for $p>0,05$.

This evaluation step is not a selection, but an interpretation step. The error term will be white noise or not white noise.

3.6 Quality of predictions

For each of the default and related rates, a regression model with only macroeconomic input factors is estimated (including a constant and a autoregressive term). Besides that, for each rate a time series ARIMA(p,d,q) is constructed, without other influences (so only autoregressive and/or moving average terms). These models are created in *Chapter 5*. In the end, the models with macroeconomic components are compared to the models only based on their own information.

Both, the macroeconomic regression model and the time series model of the rates are treated *as they were created a year ago*. For the regression, the parameters are defined for the data up until March 2010 and on basis of that, the twelve months after (April 2010 up and including March 2011) are forecasted. The time series model forecasted these twelve months based on the data known at March 2010.

The validation period is compared and tested with a paired two-sided T-test with confidence of 95%. If there is a significant difference between both validation period values, the model with the lowest sum of absolute errors is favored.

4. Investigation and orientation macroeconomic data

In the first part of this chapter *(Section 4.1)* the macroeconomic data (as collected in *Section 2.5*) is presented and hypotheses are prepared for further study. The best fitting time lag is determined for the first category of factors to show the selection method used (methodology in *Section 3.3*). The sources used are databases on www.cbs.nl and www.dnb.nl. In *Section 4.2* the ARIMA-models are created and tested. In this paragraph methodology of *Section 3.5* is applied.

There are two reasons for building time series models (ARIMA) for the macroeconomic factors:

(1) Forecasts based on the models are needed for the input for the Time Series Scenario;
(2) For the use in scenarios were one or more factors are not defined and the Time Series choice is made.

4.1 Hypotheses and orders of differencing

This chapter describes the shapes of the curves globally (graphically) and determines the best fitting time lags according to the default and related rates (see also *Section 3.3*). To create the Time Series scenario, for the selected macroeconomic factors the orders of differencing for which the time series is stationary, are determined (see also *Section 3.5.2*). The lowest order for which the time series is significantly stationary, will be used as the integrated term of the ARIMA (p,d,q). In this chapter, the *d* term of the time series of the macroeconomic factors is determined. To test for stationarity the Augmented Dickey-Fuller test (ADF) is used in all evaluations in this chapter. The first 6 lags are considered to be significantly different, which means smaller than 0,05 (for the corresponding t-statistic and the measured Tau).

For each macroeconomic factor a hypothesis is formulated. This hypothesis is treated as restriction for creating the regression models *(Chapter 6)*: if the relationship assumed (hypothesis) does not exist in the time lags zero till twenty-four, the macroeconomic factor is excluded in the specific model. In just a very few situations, the restriction could not be satisfied. The most important relationship that does not satisfy the restriction is the relationship between the house prices and DR of the White Label. The few other eliminated relationships were no scenario components.

For all rates the best time lag can be determined as in Figure 9. The results are included in the formulas in Chapter 6 for the default rate and 7 for the loss rate. Only for the first macroeconomic factor the whole process is described, including the correlation determination. For all other factors, the same process is performed and the final results are included in the thesis, Chapters 6 and 7.

4.1.1 Economic indicators

As can be suggested, and supported by Figure 11, the indicators for the consumer confidence, economic environment and willingness to buy, compiled by Central Statistics Office, are co-operating and could have an influence on the default development. The indicators are indirectly describing the general economy experienced by consumers. Because of the suspected redundancy, only the best fitting indicator will be used to describe the specific rates. The economic indicators can be called a category.

Hypothesis 1: *The indicators 'consumer confidence', 'economic environment' and 'willingness to buy' stimulate or discourage buying a (new or second) house. A decrease of these indicators would have a negative influence on the default portfolio, i.e. a negative correlation is assumed.*

The hypothesis argues a possible increase in defaults when the indicators of the Dutch economy are decreasing (of course corrected for time). This theorem could be accepted – the indicator will be included in the model – or rejected. The opposite would not be reasonable: a decrease in economic indicators cannot cooperate with a decrease of defaults. Therefore, a positive parameter result in the default rate and inflow rate, or a negative parameter in the recovery rate, would be eliminated.

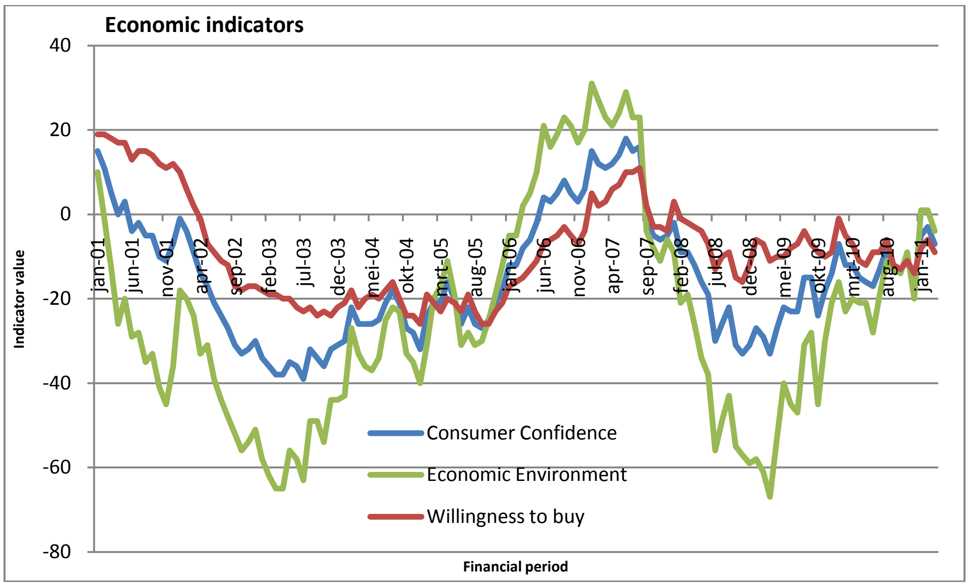


Figure 11: Several indicators of the Dutch economy (Derived values from the CBS website)

The Augmented Dickey-Fuller test obliges to use the first order of differencing for suitable forecasts. The first order of differencing of the indicators is given below. The ARIMA(p,1,q) will be used in further analyses.

For the three observed indicators, the first order of differencing is a stationary time series, according to the ADF-tests. The unit root tests are insignificant: no unit roots found in the first order of differencing. There seems to be no obvious seasonal effect, although January values are significantly higher.

To determine the best fitting time lag of the default rate of the Intermediary Channel, the lowest correlation coefficient between zero and twenty-four lags is searched. This correlation should be a negative value to be in line with the hypothesis. Because the indicators are assumed to be redundant, the lowest correlation result of the three indicators determines the best fitting factor. For the DR of the Intermediary Channel, the indicator of the economic environment with a time lag of eleven months results in the most negative correlation coefficient (-0,74). So, the DR values of January 2003 till now correspond best with the economic environment indicator series from February 2002 till eleven months ago. See Figure 13.

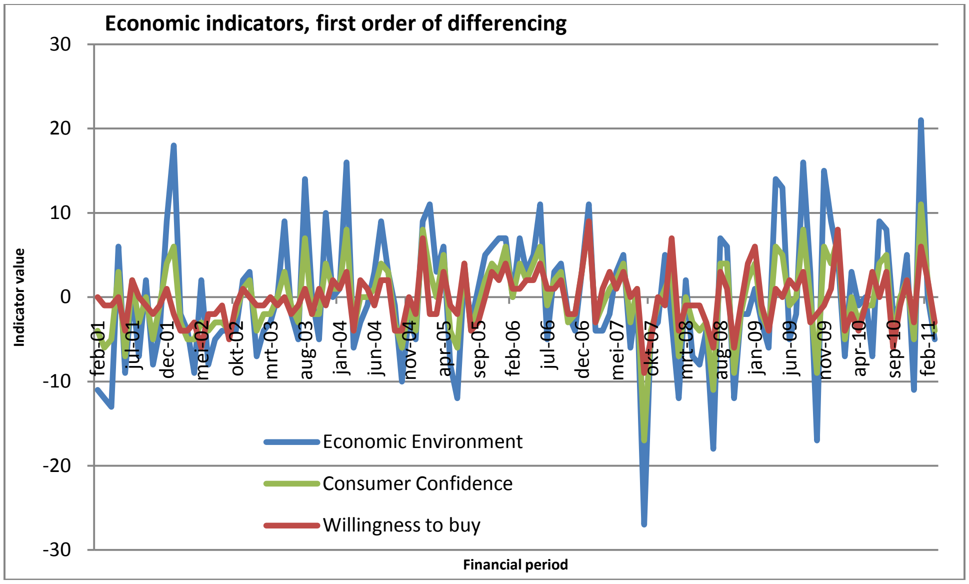


Figure 12: First order plot of economic indicators (Derived values from the CBS website)

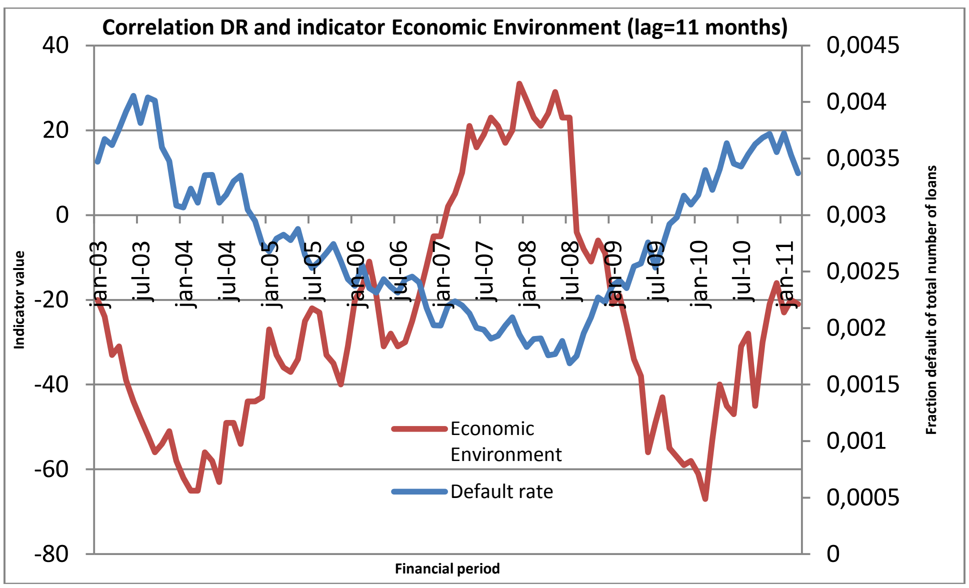


Figure 13: Correlation default rate of the Intermediary Channel and the indicator of the Economic Environment with a lag of 11 months

For all default and related rates the best fitting time lag, based on correlation, with the macroeconomic factors is determined on the same way. Therefore, not all figures and results are presented in this thesis. In the formulas in *Chapter 6* the y component (t-y) represents this time lag.

4.1.2 Occupational disability

Serious diseases can lead to problems in paying the monthly mortgage payments, probably because of sudden high payments or a serious decline in salary. It would be possible to estimate potential defaults by the fraction of employees that leave the active labor force caused by a serious injury. Two different kinds are observed: the total occupational disability rate and the fully occupational disability rate. Both will be divided by the total labor force. The best fitting factor of both will be involved in the model building process. Therefore, occupational disability is considered as a category.

Hypothesis 2: *Occupational disability can be a cause of getting into financial trouble. An increasing rate of occupational disabled (to work force) increases the number of defaults.*

The default rate and inflow rate are expected to increase, when the rate of occupational disabilities increases. The opposite could be true for the recovery rates. People would be earlier in financial trouble when (suddenly) the income is reduced. Odd direction, i.e. an increase leads to a reduction in defaults, could not be argued and are therefore removed from the model. These factors are both stationary on the first order of differencing.

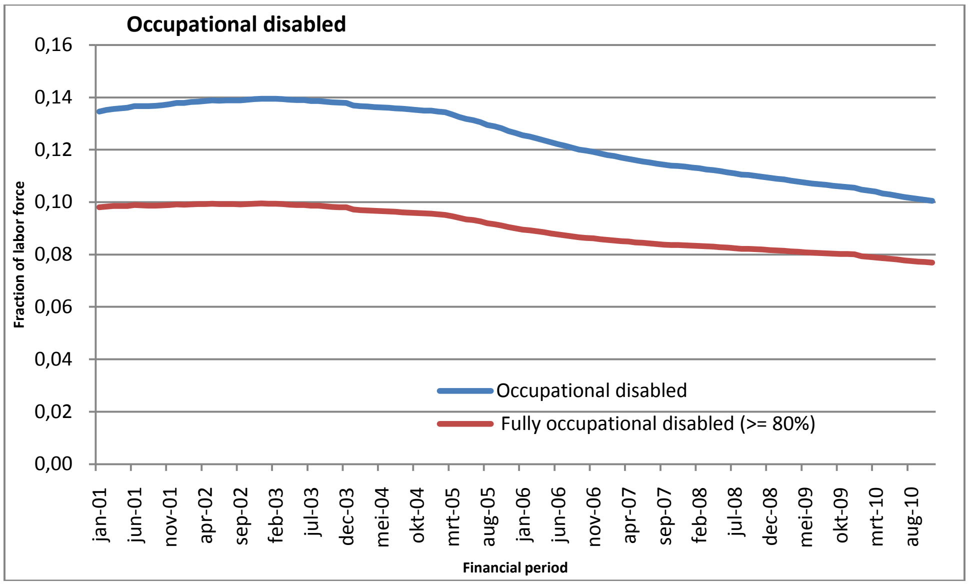


Figure 14: Fraction occupational disabled of work force (Derived values from the CBS website)

4.1.3 Diseases and deaths

In line with the previous paragraph, some diseases on earlier ages or sudden deaths could have an influence on the mortgage payments. Two specific cases are chosen: (1) new diseases and (2) acute heart infarcts. Both are a fraction of the total population and are in the same category (and preferred not to use in combination with occupational disability factors, because of possible redundancy).

Hypothesis 3: *Diseases and (family) deaths could be a reason for financial problems, and therefore for the monthly mortgage payments. An increase in (sudden) deaths could lead to an increased number of defaults.*

Increasing (sudden) deaths, would possible lead to increased default rates and/or reduced recovery rates. The other direction could not be argued and therefore *wrong* parameters will be eliminated.

The trend and single mean evaluation are indicating a significant ADF test on the raw data for the deaths caused by total new diseases. The order of differencing of zero corresponds to the function with certain mean (not equal to zero). For the acute heart infarcts, the mean and trend are necessary according to the ADF test. A first order of differencing is significant for both in all three model types (without mean, with mean or with trend), and by interpreting the results preferred.

4.1.4 Consumer price index and inflation

The consumer price index and the inflation are going hand-in-hand and therefore assigned as category. The consumer price index indicates the price of consumer goods. A buying pattern is not easy changed by people. This could lead to financial problems and initiates the start of a credit problem for the consumer. This reasoning is, again, in one direction.

Hypothesis 4: *The consumer price index (CPI) indicates the prices of products. An increase in prices could lead to an increased default portfolio.*

The raw data of the CPI is not stationary, but the first order of differencing is clearly stationary, as is confirmed by the ADF test.

Inflation could probably influence the default portfolio. An increasing inflation could decrease the defaults for people with a fixed rate mortgage. Most of the mortgages are fixed rated (for a certain period) and therefore a negative correlation is expected. On the other side, it could be argued that the inflation is positively correlated, because a higher inflation could lead to higher costs for living (products are more expensive and probably loans are not increasing with a same rate).

Hypothesis 5: *Inflation could correlate with the default portfolio. The direction is unknown in advance.*

The ADF test does not show a stationary result on the raw data of the inflation rate. The first order of differencing is clearly a stationary time series.

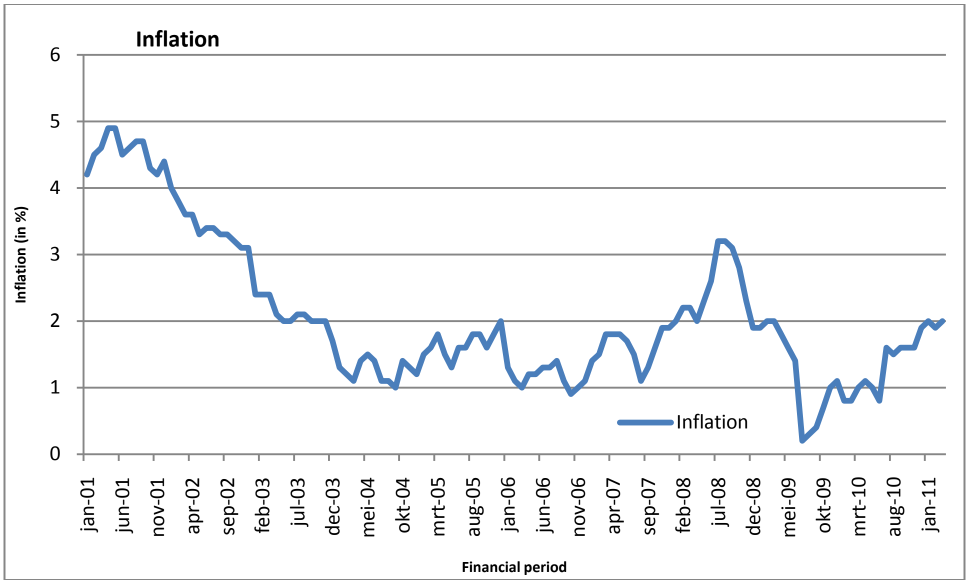


Figure 15: Inflation (Derived values from CBS website)

4.1.5 House prices and transactions

The house prices could have an influence on the defaults. A reduction of house prices could lead to higher losses for a bank in case of default, but the mortgage payments are unaffected. Increasing house prices are an indication of an attractive housing market. Therefore, a negative correlation is expected. A positive correlation result will be eliminated.

Hypothesis 6: *House prices could correlate negative with the default portfolio. An increase in prices could lead to a decreased default portfolio.*

The number of houses sold in a period could be an indication of the throughput of the housing market. If more houses are sold, the market would be in a positive flow, which could be an indication of a better performing default portfolio. Therefore, an increase of transactions is supposed to not cooperate with an increase of defaults. A solution suggesting this relation would be left out of the model.

Hypothesis 7: *The number of houses sold could correlate negative with the default portfolio. An increase in transactions could lead to a decreased default portfolio.*

It can be seen easily in the graph that the house prices is not a stationary time series. According to the ADF test, the number of houses sold is stationary with a trend on the raw data. This stationarity is very weak, so it is better to take the first order of differencing for both. These factors are not in the same category, because the redundancy is not obvious: house prices describe the value of the security and houses sold describe the flow in the housing market.

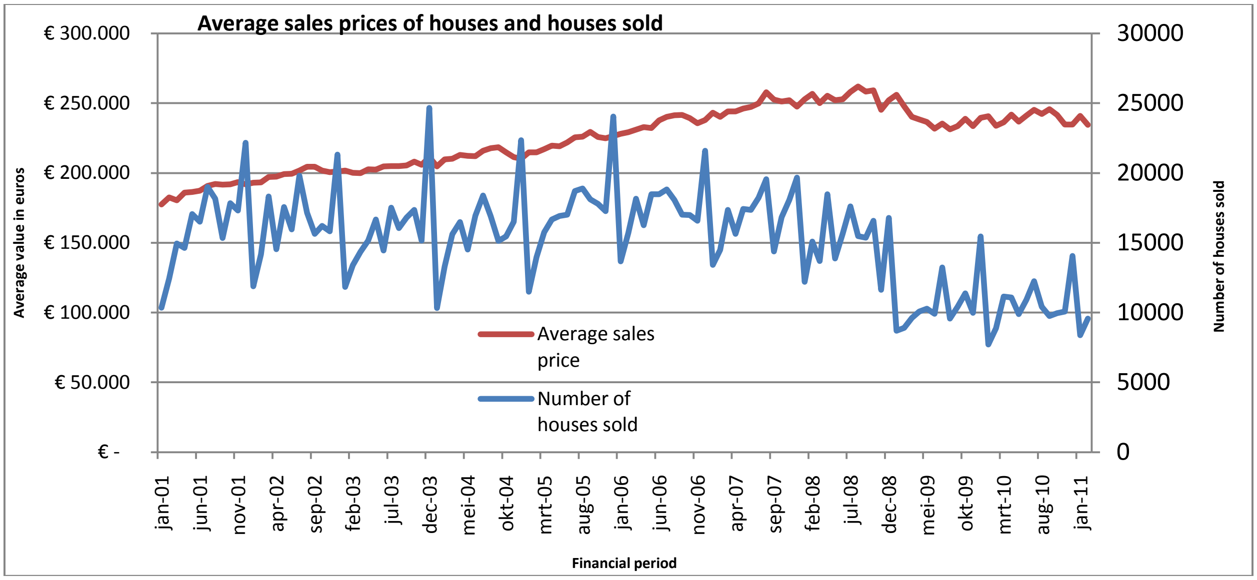


Figure 16: House prices and transactions (Derived values from CBS website)

4.1.6 Negative balance on bank account

In financially difficult times, people would probably have to deal with a negative bank account, faster. This can be split up in two indicators: the total negative balance and the total number of negative bank accounts. These factors are classified in the same category.

Hypothesis 8: *When more people are having a negative bank account, more defaults could be expected. A higher total negative balance would be an indication of a higher number of defaults.*

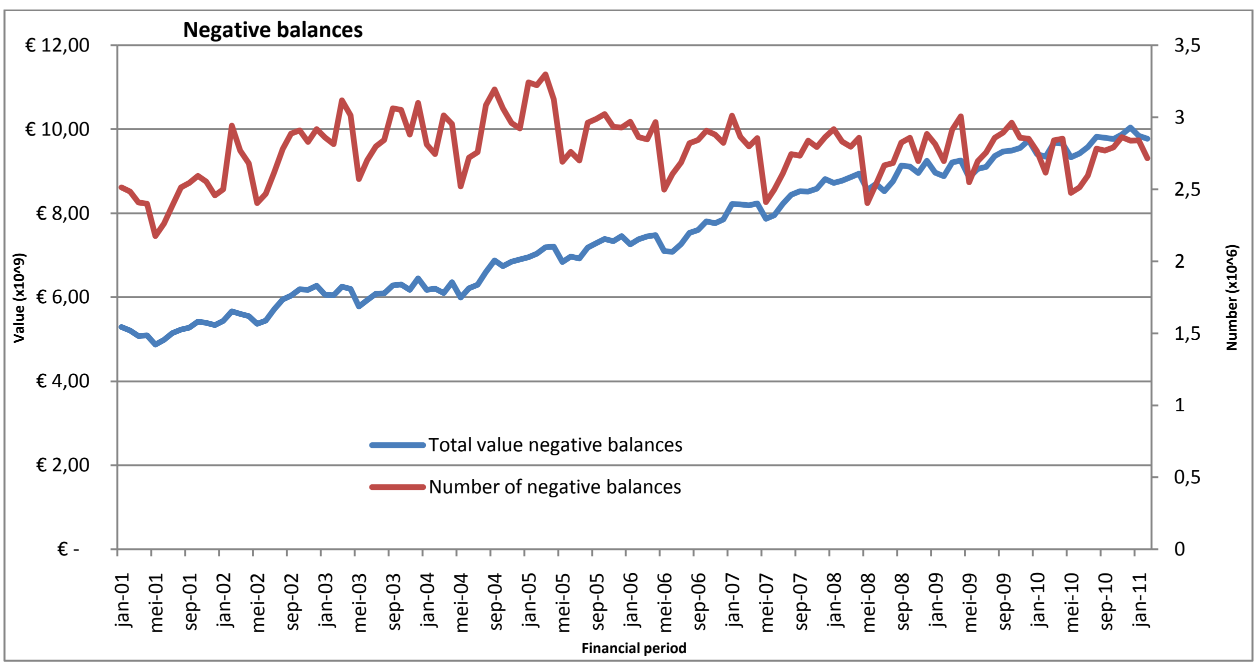


Figure 17: Negative balance and number of negative balances (Derived values from CBS website)

The total balance of negative bank accounts is not significantly tested with an ADF test if a trend is included. The number of negative accounts is of differencing order zero when a certain mean or trend in included. Although, the results are unsatisfying and the first order seems to fit better.

4.1.7 Registered unemployment

Unemployment could initiate financial problems, resulting in defaults, due to a decline in cash inflow. The inflow as well as the outflow of unemployment is observed. An increasing inflow would lead to more defaults; an increasing regular outflow (recovery) could reduce the default portfolio. Other relationships are eliminated in the model.

Hypothesis 9: *More unemployment could lead to more defaults. The indicator registered unemployment can describe the direction of the default portfolio. An increasing unemployment could lead to more defaults, and vice versa.*

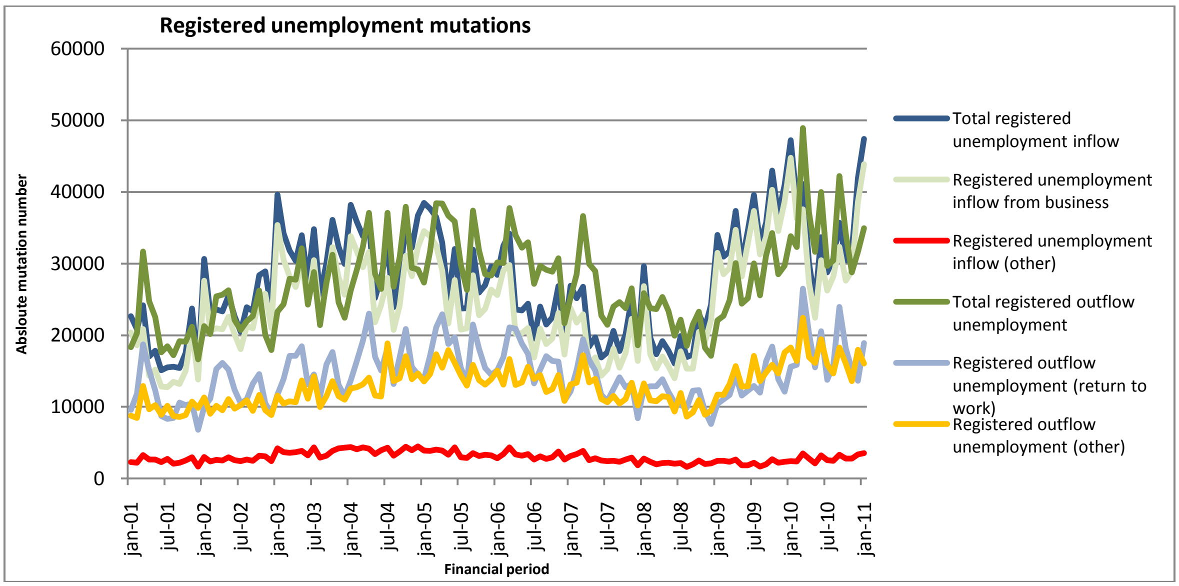


Figure 18: Registered unemployment (Derived values from CBS website)

All kinds of inflows observed are likely to be of the first order of differencing (ADF test). The total outflow (and the specified other outflow) also seems to be a first order differencing function, but the outflow returning to work is, according to the ADF test, stationary on the raw data with mean or trend. Because of this weak stationarity, all factors are treated as first order of differencing factors.

Also the unemployment rate (seasonal corrected) is included in this study. It appears that this factor is also of the first order of differencing.

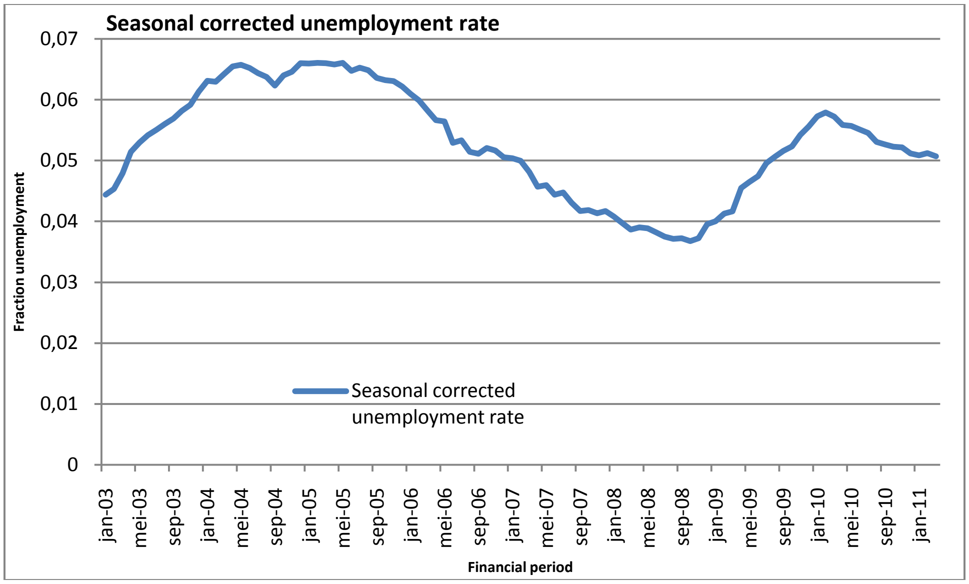


Figure 19: Seasonal corrected unemployment rate (Derived values from CBS website)

4.1.8 Interest rates

A possible well-defined estimate for the yield rate is the ten years government bond yield rate. Upwards trends in yield would probably increase the number of defaults. A decreasing rate would possibly increase the recovery rates.

Hypothesis 10: *Mortgages are directly affected by yield rate changes with a positive correlation.*

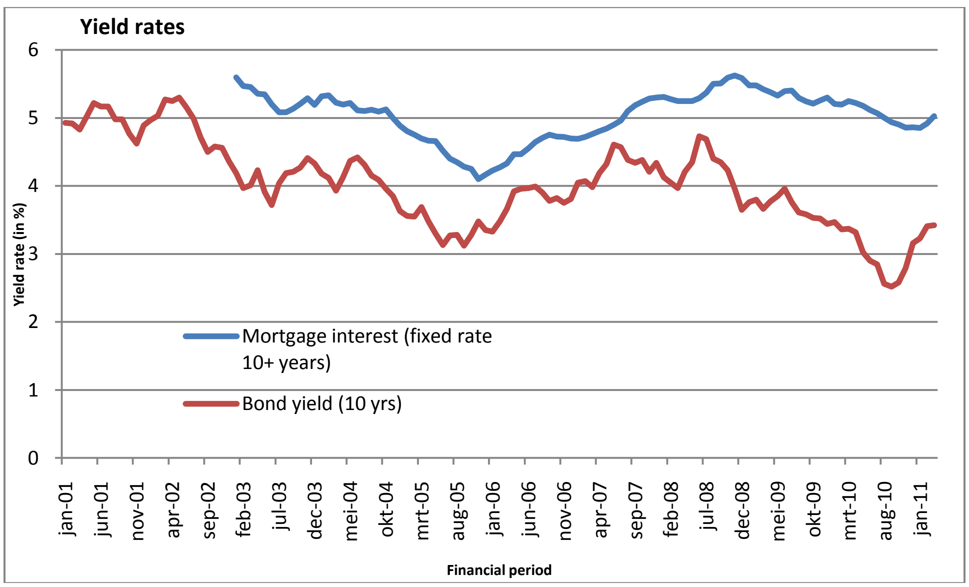


Figure 20: Interest rates 10 years government bonds (Derived values from DNB website)

An increasing mortgage interest could affect the payments problems negatively. As an indicator, the ten years (and more) mortgage interest rate is taken into account. Together with the bond rate, the interest rates are classified as category.

Hypothesis 11: *Mortgages with variable yield rates are directly affected by mortgage yield rate changes. Higher rates could lead to more defaults.*

The Augmented Dickey-Fuller test recognizes a unit root in the time series. The first order of differencing is stationary for both.

4.2 ARIMA models of macroeconomic factors

To determine the underlying time series for the Time Series Scenario and to fill scenario gaps, ARIMA (p,d,q) models are build for the macroeconomic factors (see also the methodology in *Section 3.5*). The *integrated* terms are known from *Section 4.1*. In this chapter the models are estimated, using the Schwarz information criterion (SIC). The p and q are chosen in a range between zero and twelve. At the next stage, before regression is applied, time issues are treated. Now, it is only important to reconstruct the properties of the time series.

The results are tabulated. For each macroeconomic factor, an ARIMA(p,d,q) model is estimated based on the SIC. In the column ARIMA(p,d,q) the models are defined. An ARIMA(0,1,2) means zero autoregressive terms, first order of differencing and two moving average terms. The lags of the autoregressive and moving average terms are written in the same column in vector format. Suppose ([1 2], []), then two autoregressive terms are included in the best solution, AR(1) and AR(2). This means that the formula is based on the values on time t-1 and t-2.

The parameters of the model are tabulated in the fourth column and the 95% standard error in the fifth column. The standard error indicates the stability of the parameter. In the methodology part *(Section 3.5)* the ARIMA-model is described. The parameters of the fourth column are the input parameters of Equation 11, with $\phi(B) = 1 - \sum_{j=1}^{p} \phi_j B^j$ and $\theta(B) = 1 - \sum_{i=1}^{q} \theta_i B^i$. The AR(p)-components belong to the $\phi(B)$ input and MA(q) to the $\theta(B)$ input.

Macroeconomic factor	**ARIMA(p,d,q) ([lags p], [lags q])**	**SIC**	**Parameters**	**Standard error parameters**
Consumer Confidence	(0,1,0) without mean	146,5241	-	-
Economic Environment	(0,1,0) without mean	284,3328	-	-
Willingness to buy	(0,1,1) without mean ([],[12])	48,33982	MA(12): -0,28708	MA(12): 0,09124
Total occupational disability	(2,1,0) with mean ([1 3],[])	-1135,48	MU: 0,0042240 AR(1): 0,56991 AR(3): 0,42118	MU: 0,0016026 AR(1): 0,07518 AR(3): 0,07582
Fully occupational disability	(1,1,1) without mean ([1],[1])	-664,786	AR(1): 0,98768 MA(1): 0,69970	AR(1): 0,01827 MA(1): 0,07320
Deaths by heart infarcts	(2,1,2) without mean ([1 2], [1 2])	598,13	AR(1): -1,74697 AR(2): -0,94840 MA(1): -1,72706 MA(2): -0,95750	AR(1): 0,06898 AR(2): 0,06430 MA(1): 0,06400 MA(2): 0,06023
Deaths by new diseases	(1,1,3) without mean ([1],[1 2 6])	-1551,45	AR(1): 0,99874 MA(1): 0,33029 MA(2): 0,40087 MA(6): 0,01694	AR(1): 0,0083173 MA(1): 0,08598 MA(2): 0,08667 MA(6): 0,07711
Consumer Price Index	(1,1,1) with mean ([3],[1])	-479,747	MU: 0,01589 AR(3): -0,68658 MA(1): -0,43062	MU: 0,00024970 AR(3): 0,07040 MA(1): 0,08409
Inflation	(0,1,0) without mean	13,8223	-	-

Housing market	(1,1,3) without mean ([12],[1 2 3])	-47,125	AR(12): 0,76982 MA(1): 0,75925 MA(2): -0,20789 MA(3): -0,15101	AR(12): 0,06232 MA(1): 0,09163 MA(2): 0,11395 MA(3): 0,09172
House prices	(0,1,1) without mean ([],[1])	-454,966	MA(1): 0,22103	MA(1): 0,08948
Negative bank accounts	(1,1,1) without mean ([1],[1])	-98,6273	AR(1): 0,49984 MA(1): 0,89086	AR(1): 0,11226 MA(1): 0,05848
Value negative accounts	(0,1,0) with mean	-66,8905	MU: 0,03701	MU: 0,01643
Total inflow registered unemployment	(2,1,2) without mean	155,831	AR(1): -1,08524 AR(2): -0,94559 MA(1): -0,80446 MA(2): -0,80757	AR(1): 0,04342 AR(2): 0,04131 MA(1): 0,06983 MA(2): 0,06858
Inflow from business	(3,1,3) without mean ([1 2 3],[1 2 3])	141,6985	AR(1): -2,04026 AR(2): -1,99543 AR(3): -0,90699 MA(1): -1,73202 MA(2): -1,54834 MA(3): -0,70308	AR(1): 0,08803 AR(2): 0,10550 AR(3): 0,08264 MA(1): 0,12596 MA(2): 0,14559 MA(3): 0,11990
Inflow other	(2,1,0) without mean ([1 2],[])	164,2813	AR(1): -0,73955 AR(2): -0,46001	AR(1): 0,08221 AR(2): 0,08256
Total outflow registered unemployment	(3,1,2) without mean ([2 3 4],[1 2])	140,2418	AR(2): -0,66538 AR(3): -0,36251 AR(4): -0,62811 MA(1): 0,86705 MA(2): -0,93940	AR(2): 0,07508 AR(3): 0,07517 AR(4): 0,07527 MA(1): 0,03773 MA(2): 0,03817
Outflow other	(2,1,0) without mean ([1 2],[])	-69,3768	AR(1): -0,79934 AR(2): -0,47632	AR(1): 0,08132 AR(2): 0,08255
Seasonal unemployment	(1,1,1) without mean ([1],[1])	-163,154	AR(1): 0,90066 MA(1): 0,59284	AR(1): 0,06765 MA(1): 0,12493
Mortgage interest	(1,1,0) without mean ([1],[])	-257,726	AR(1): 0,39775	AR(1): 0,09450
Bond yield	(0,1,1) without mean ([],[1])	-111,749	MA(1): -0,33004	MA(1): 0,08610

Table 5: Model characteristics of the macroeconomic factors

In Table 6 in the second column, the significance of each parameter of the models for the macroeconomic factors is determined by a T-test. When a>0,05, it is said to be insignificant in the formula. Those parameters could be eliminated. Sometimes the solution is hurt by eliminating. Therefore, all parameters are observed in the results.

In the third column, the residuals are tested for normality. When a<0,05, the residuals are not normally distributed (indicated by writing the value italic). This is the Ljung-Box test. In the last column the volatility of the model is indicated by the standard error estimate.

Macroeconomic factor (Continued)	**T-value (P > |t|)**	**χ^2 (P > χ^2)**	**Standard error estimate**
Consumer Confidence	-	3,24 (0,7777)	0,441124
Economic Environment	-	2,86 (0,8259)	0,775971
Willingness to buy	MA(12): -3,15 (0,0021)	3,61 (0,6068)	0,290428
Total occupational disability	MU: 2,64 (0,0096) AR(1): 7,58 (<0,0001) AR(3): 5,55 (<0,0001)	7,14 (0,1289)	0,001877
Fully occupational disability	AR(1): 54,07 (<0,0001) MA(1): 9,56 (<0,0001)	1,76 (0,7806)	0,014014
Deaths by heart infarcts	AR(1): -25,33 (<0,0001) AR(2): -14,75 (<0,0001) MA(1): -26,99 (<0,0001)	2,70 (0,2587)	2,691805

	MA(2): -15,90 (<0,0001)		
Deaths by new diseases	AR(1): 120,08 (<0,0001) MA(1): 3,84 (0,0002) MA(2): 4,63 (<0,0001) *MA(6): 0,22 (0,8265)*	*6,31 (0,0427)*	0,000374
Consumer Price Index	MU: 6,36 (<0,0001) AR(3): -5,12 (<0,0001) MA(1): -9,75 (<0,0001)	*11,60 (0,0206)*	0,032331
Inflation	-	3,88 (0,6923)	0,256074
Housing market	AR(12): 12,35 (<0,0001) MA(1): 8,29 (<0,0001) *MA(2): -1,82 (0,0706)* *MA(3): -1,65 (0,1024)*	1,80 (0,4060)	0,187096
House prices	MA(1): 2,47 (0,0149)	*13,23 (0,0213)*	0,036348
Negative bank accounts	AR(1): 4,45 (<0,0001) MA(1): 15,23 (<0,0001)	*10,81 (0,0288)*	0,156017
Value negative accounts	MU: 2,25 (0,0261)	*14,03 (0,0294)*	0,180683
Total inflow registered unemployment	AR(1): -25,00 (<0,0001) AR(2): -22,89 (<0,0001) MA(1): -11,52 (<0,0001) MA(2): -11,78 (<0,0001)	*10,02 (0,0067)*	0,434968
Inflow from business	AR(1): -23,18 (<0,0001) AR(2): -18,91 (<0,0001) AR(3): -10,98 (<0,0001) MA(1): -13,75 (<0,0001) MA(2): -10,63 (<0,0001) MA(3): -5,86 (<0,0001)	*(lag 12):* *25,27 (0,0003)*	0,397497
Inflow other	AR(1): -9,00 (<0,0001) AR(2): -5,57 (<0,0001)	4,25 (0,3730)	0,464904
Total outflow registered unemployment	AR(2): -8,86 (<0,0001) AR(3): -4,82 (<0,0001) AR(4): -8,34 (<0,0001) MA(1): 22,98 (<0,0001) MA(2): -24,61 (<0,0001)	*11,85 (0,0006)*	0,401295
Outflow other	AR(1): -9,83 (<0,0001) AR(2): -5,77 (<0,0001)	2,66 (0,6154)	0,175608
Seasonal unemployment	AR(1): 4,75 (<0,0001) MA(1): 13,31 (<0,0001)	1,72 (0,7875)	0,101486
Mortgage interest	AR(1): 4,21 (<0,0001)	4,74 (0,4483)	0,063792
Bond yield	MA(1): -3,83 (0,0002)	2,35 (0,7988)	0,152437

Table 6: Model characteristics of the macroeconomic factors (continued from Table 5)

5. Investigation and orientation default portfolios

In this chapter the time series for DR, IR, RR and FR are identified for each label. The same split up as in *Chapter 4* is made, first the data is presented and the order of differencing is determined *(Section 5.1)* and thereafter the models are designed *(Section 5.2)*. The goal of this chapter is to create time series models based on the history of the rate (referred to as the microeconomic approach). These models will be compared with the regression models with macroeconomic input in *Chapter 6*. Only when macroeconomic inputs improve the quality of the models, it is worth to use these analyses.

5.1 Determining the orders of differencing

The consolidated portfolio consists of the labels (1) Intermediary Channel, (2) White Label and two small passive portfolios. The constructed consolidated portfolio consists of Intermediary Channel, White Label, and two small labels.

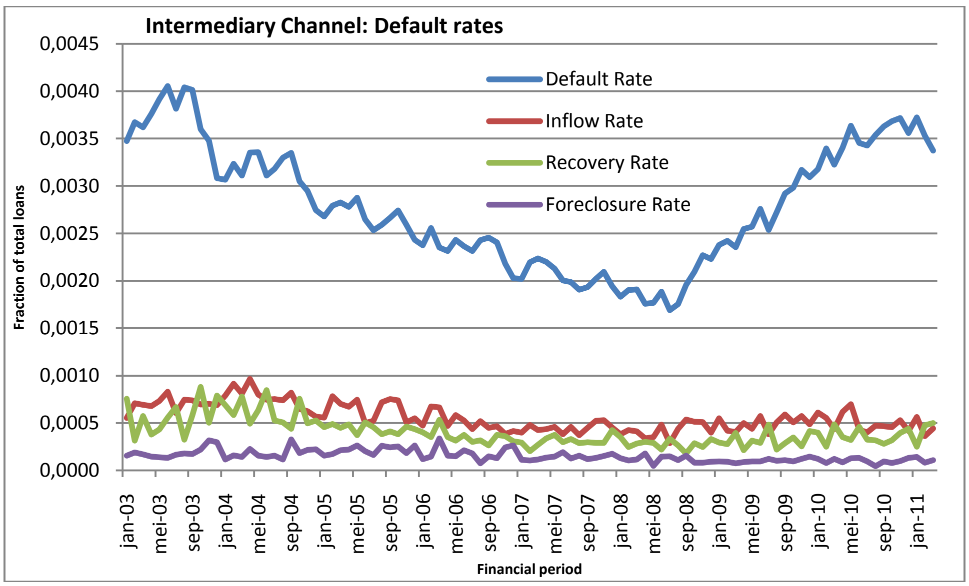


Figure 21: Intermediary Channel default time series

Except the foreclosure rate, which should include a mean or trend to satisfy the ADF-test for order zero, the rates for the Intermediary Channel are of the first order of differencing. Therefore, all rates are treated as first order of differencing rates.

The White Label portfolio default time series are a bit different. The inflow rate, recovery rate and foreclosure rate series should contain a mean or trend to fit an order of differencing of zero. The DR does not satisfy the ADF-test and the integrated term would be one. In the end, it would be better to use the first order of differencing for all factors of the White Label.

The Consolidated Portfolio is largely based on the Intermediary Channel and the orders are similar to the Intermediary Channel's analyses.

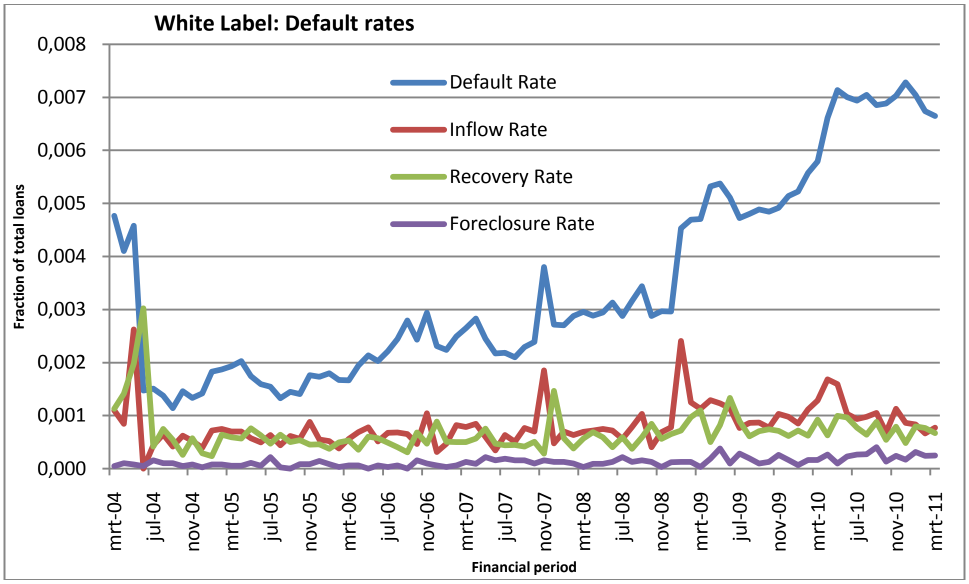


Figure 22: White Label default time series

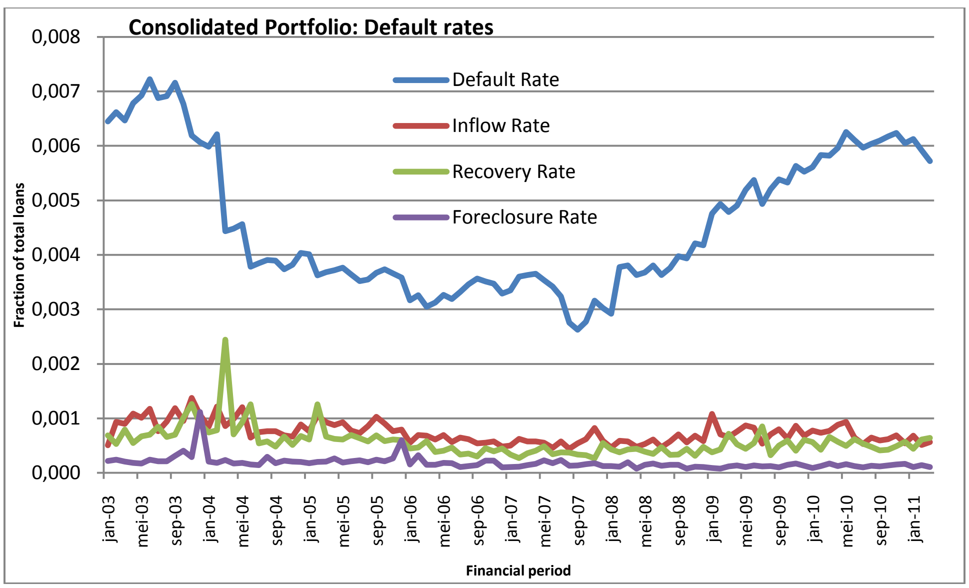


Figure 23: Consolidated default time series

5.2 ARIMA models of risk parameters

The risk parameters can be described by ARIMA(p,d,q) models as well. This is not a separate goal, but in the end, the models can be compared to discuss whether the macroeconomic factor model performs better forecasts than the default time series forecasts based on the own time series, the microeconomic approach. The results are tabulated exactly the same as in *Section 4.2* for the macroeconomic factors. For each rate, the ARIMA(p,d,q) is determined with regard to the result of the Schwarz information criterion. The number of parameters included is written in the second column and in the same column the corresponding lags are mentioned. In the third column the Schwarz information criterion value is written and the values and standard errors (95%) corresponding are in the fourth and fifth column. Table 8 contains the t-values of the parameters and a Chi-square test value for the residuals (test for normality). The standard error of the whole model is added too.

Rate	**ARIMA(p,d,q) ([lags p], [lags q])**	**SIC**	**Parameters**	**Standard error parameters**
Intermediary Channel: default rate	(0,1,0) without mean	-1450,39	-	-
Intermediary Channel: inflow rate	(0,1,1) without mean ([],[1])	-1540,66	MA(1): 0,66537	MA(1): 0,07605
Intermediary Channel: recovery rate	(1,1,1) without mean ([3],[1])	-1504,35	AR(3): 0,26839 MA(1): 0,79237	AR(3): 0,11293 MA(1): 0,07097
Intermediary Channel: foreclosure rate	(0,1,1) without mean ([],[1])	-1659,51	MA(1): 0,84628	MA(1): 0,05389
White Label: default rate	(0,1,0) without mean	-1038,96	-	-
White Label: inflow rate	(0,1,2) without mean ([],[1 2])	-1064,56	MA(1): 0,59737 MA(2): 0,24976	MA(1): 0,10695 MA(2): 0,10726
White Label: recovery rate	(0,1,1) without mean ([],[1])	-1091,09	MA(1): 0,70284	MA(1): 0,07810
White Label: foreclosure rate	(0,1,1) without mean ([],[1])	-1371,04	MA(1): 0,82871	MA(1): 0,06318
Consolidated Portfolio: default rate	(0,1,2) without mean ([],[1 3])	-1323,47	MA(1): 0,18021 MA(3): -0,43828	MA(1): 0,08989 MA(3): 0,09124
Consolidated Portfolio: inflow rate	(0,1,1) without mean ([],[1])	-1440,32	MA(1): 0,58163	MA(1): 0,08268
Consolidated Portfolio: recovery rate	(1,1,1) without mean ([3],[1])	-1357,64	AR(3): 0,22122 MA(1): 0,84355	AR(3): 0,11087 MA(1): 0,06083
Consolidated Portfolio: foreclosure rate	(0,1,1) without mean ([],[1])	-1500,87	MA(1): 0,90169	MA(1): 0,04380

Table 7: Model characteristics of the portfolio rates

Rate (Continued)	**T-value (P > |t|)**	**χ^2 (P > χ^2)**	**Standard error estimate**
Intermediary Channel: default rate	-	4,83 (0,5656)	0,000148
Intermediary Channel: inflow rate	MA(1): 8,75 (<0,0001)	3,68 (0,5967)	0,000092
Intermediary Channel: recovery rate	AR(3): 2,38 (0,0195) MA(1): 11,17 (<0,0001)	1,46 (0,8333)	0,000108
Intermediary Channel: foreclosure rate	MA(1): 15,70 (<0,0001)	4,87 (0,4324)	0,00005
White Label: default rate	-	4,12 (0,6601)	0,000499
White Label: inflow rate	MA(1): 5,59 (<0,0001) MA(2): 2,33 (0,0223)	1,10 (0,8936)	0,000411
White Label: recovery rate	MA(1): 9,00 (<0,0001)	3,39 (0,6400)	0,000385
White Label: foreclosure rate	MA(1): 13,12 (<0,0001)	1,13 (0,9519)	0,000068
Consolidated Portfolio: default rate	MA(1): 2,00 (0,0478) MA(3): -4,80 (<0,0001)	3,13 (0,5369)	0,000273
Consolidated Portfolio: inflow rate	MA(1): 7,03 (<0,0001)	5,78 (0,3278)	0,000153
Consolidated Portfolio: recovery rate	AR(3): 2,00 (0,0488) MA(1): 13,87 (<0,0001)	5,34 (0,2541)	0,000229
Consolidated Portfolio: foreclosure rate	MA(1): 20,59 (<0,0001)	4,30 (0,5076)	0,000112

Table 8: Model characteristics of the portfolio rates (continued from Table 7)

5.3 Seasonal influences

It could be argued that defaults are seasonal dependent, for example by reasoning that in certain months the unemployment raises or overall living expenses are higher. To test this hypothesis roughly, the monthly averages over the available years are calculated and the highest will be compared to the lowest with a one-sided T-test. This test is done for the Consolidated Portfolio as well as for the Intermediary Channel and the White Label. In a few cases, there was a significant difference, mainly caused by one abnormal value. This value was an unreliable high rate in the oldest data. Without this value, no significant difference was observed. It can be concluded that seasonal influences in the default portfolio are not significant.

6. Scenario effects for risk parameters

In this chapter the models are created by regression analyses (see also the methodology in *Sections 3.4*). These are the answers of *subquestions 1 and 2* (see *Section 1.2*). In *Chapter 4* the time lags are described and the time series forecasts for the Time Series scenario are determined. In *Chapter 5* the microeconomic approach is applied on the rates. For the described default rates and related rates (inflow rate, recovery rate and foreclosure rate), there is a model created based on the macroeconomic factors argued to have an influence on the portfolio. For default rates these models are presented in this chapter, for the related rates the models can be found in *Appendix A*. The default rates are modeled for the Intermediary Channel portfolio *(Section 6.1, related rates in Appendix A1)*, the White Label portfolio *(Section 6.2, related rates in Appendix A2)*, and the Consolidated Portfolio *(Section 6.3, related rates in Appendix A3)*. The favored model (macroeconomic or microeconomic) is determined by a T-test (see *Section 3.7* for the methodology, *Chapter 5* for the time series models of the rates and *Appendix A4* for an overview of the comparisons).

The default rate is used in further analyses to estimate the credit losses; the related rates give an indication of the development within the default phase (recovery or foreclosure) and the throughput of defaults (inflow rate) *(Section 8.4)*. For all models, when the error of the time series model from *Chapter 5* is significantly smaller, the macroeconomic model is subordinated to the microeconomic approach. In the text, this judgment is made and in the conclusion section it will be used and summarized. The summarizing table can be found in *Appendix A4* for the three portfolios. Graphs reflecting the validation period of the default models can be found in the text and for related rates these are in the *Appendix A* sections.

6.1 Default rate Intermediary Channel

The observed factors (selected with the least-absolute errors method over all data and over all data minus the last twelve months) can be tested for the best combination that minimizes the average error over the last twelve months of data, April 2010 up until March 2011. The best fitting solution, according to logistic regression, for the default rate that satisfies the preconditions is based on the factors (1) house prices, (2) houses sold, (3) total inflow registered unemployment, (4) total outflow registered unemployment, (5) mortgage yield rate (fixed yield period of ten years and more) and a constant and an autoregressive term for the default series itself. The average error over the last twelve months is 6,23 percent. The corresponding values of the parameters and the graph of the forecasts till the end of 2015 are illustrated by Figure 24, including the different scenarios.

Constant	-0,88
House prices	-0,03
Houses sold	-0,06
Total inflow registered unemployment	0,01
Total outflow registered unemployment	-0,05
Mortgage yield	0,02
AR(1)	0,82

Table 9: Parameters for the default rate of the Intermediary Channel

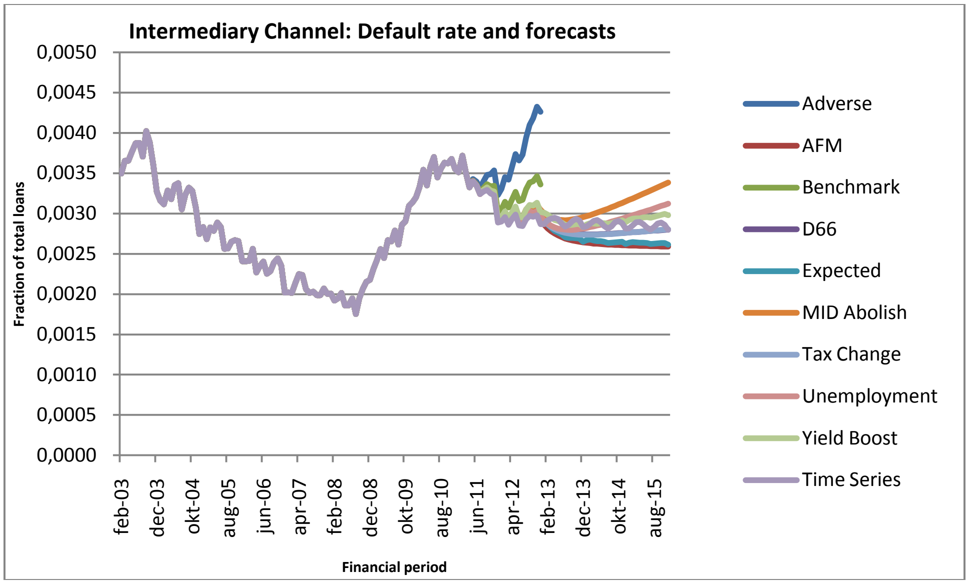


Figure 24: Intermediary Channel default time series and forecasts based on applied scenarios

$$\ln\left(\frac{DR_{Intermediary}(t)}{1-DR_{Intermediary}(t)}\right) = -0,88 - 0,03 * HousePrices(t) * 10^{-5} - 0,06 * HousesSold(t-12) * 10^{-4} + 0,01 * InflowUnemploymentTotal(t-4) * 10^{-4} - 0,05 * OutflowUnemploymentTotal(t-21) * 10^{-4} + 0,02 * MortgageInterestRate(t) + 0,82 * \ln\left(\frac{DR_{Intermediary}(t-1)}{1-DR_{Intermediary}(t-1)}\right)$$

The scenarios selected in *Section 2.6* are applied on the underlying macroeconomic factors and with the regression parameters calculated the effects of the scenarios are expressed in Figure 24. The so called Time Series scenario is a technical and not very reliable scenario, but it gives an indication of the direction when none of the input factors significantly changes of direction. The scenario analysis part is a lot more interesting for the forecasts, but the regular analysis could not lack. The curve in the scenarios is different. In the coming months downsize is expected and a restless stability is supposed in the middle long run. Scenario lines intercept and imaginable slopes move in different directions. In general, the average default rate is supposed between 2,5‰ and 3,5‰. Remember that the stress test scenarios (Benchmark and Adverse) end in 2012.

A validation over the last year is calculated in two situations: (1) macroeconomic approach, considering the current input parameters and calculate (on the same way as the actual model) the values for the period between April 2010 and March 2011, and (2) microeconomic approach, calculating the values for the same period only based on the ARIMA model of the times series itself. The percentage error is defined as the absolute error divided by the realized value and is calculated as the average over the twelve errors for the last year.

In fact the comparison is between the Time Series scenario of the macroeconomic factor model and the ARIMA model of the default rate, without macroeconomic factors. This comparison should be made, because it could support the input of macroeconomic factors. Otherwise, the macroeconomic factor addition is superfluous. The evaluation step advances the macroeconomic model with 6,23 against 9,27 percent for the microeconomic approach and a significant difference as paired samples T-test (two-sided) result. Therefore, the macroeconomic model is favored for forecasting purpose.

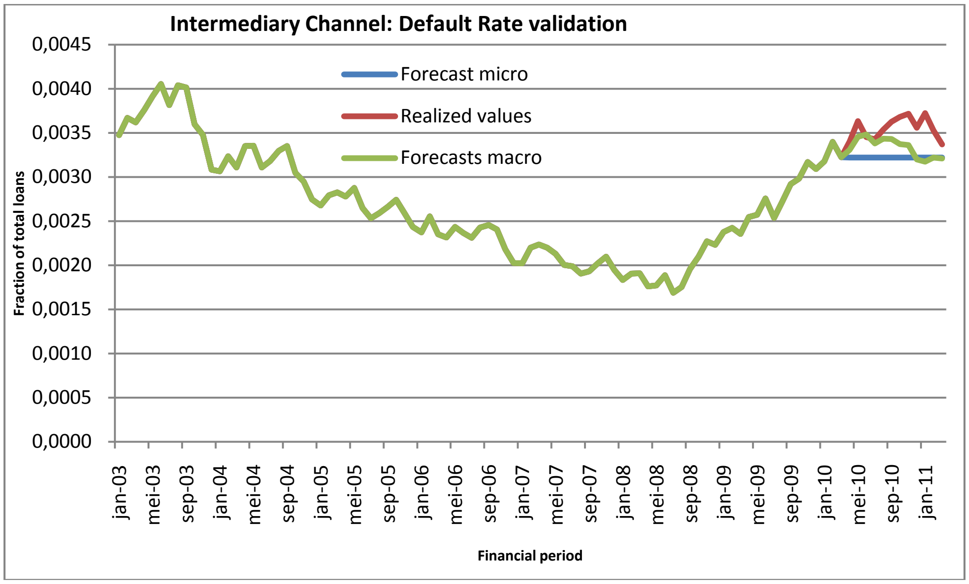


Figure 25: Intermediary Channel default rate time series validation, based on micro- and macroeconomic approaches

The default rate model values are most of the time not exactly equal to the realized values. It is said that the model values include an error compared to the realized values (up until March 2011). These so-called residuals can be characterized by a normal distribution (see Figure 26) with a mean of approximately zero and a little variance. The 95% confidence interval of the errors is between -0,00023 and 0,00025. The R-square value is 0,96. This value is calculated by subtracting the sum of squared errors (between model and observed values) divided by the total sum of squares (between the observation and average) from one. A higher R-square value (between 0 and 1) means a better (forecasting) fit.

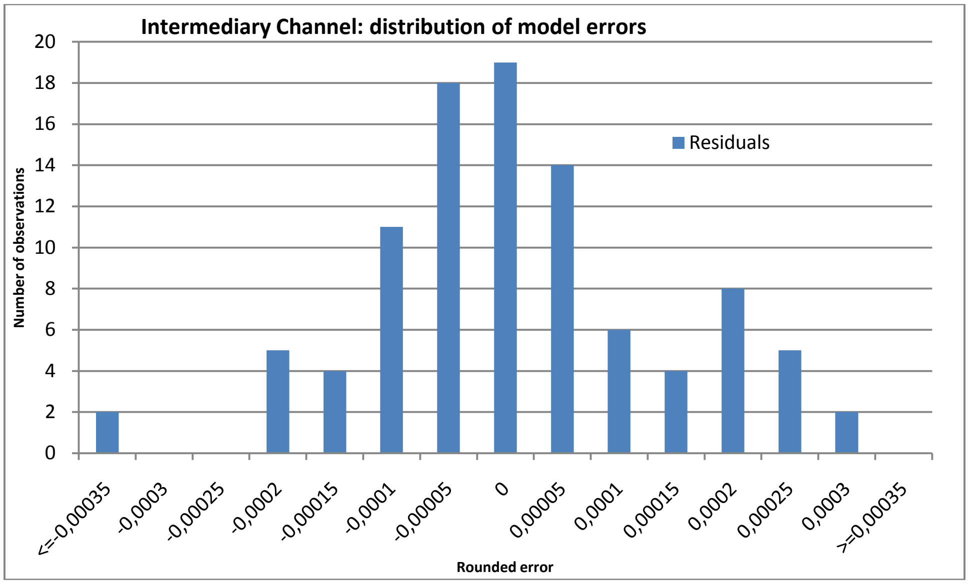


Figure 26: Intermediary Channel distribution of model errors

6.2 Default rate White Label

For all White Label rates, the microeconomic approach is favored by the paired samples T-test. Statistics of the related rates can be found in *Appendix A2*. The White Label portfolio distinguishes due to a positive relation of defaults and the house prices. By means of unreliability, the house prices in the defaults are out of scope for the White Label defaults rate. Probably, defaults in the White Label portfolio are stronger affected by economic indicators and therefore an inflation rate is included, which strengthens the other factors. The error is very huge and the model choice is hard, neither the ARIMA model itself provides good estimates, but is favored by the result of the T-test. In this thesis the scenarios are applied on the macroeconomic approach, but it is good to know the quality of the model (not better than the microeconomic approach).

Looking at the graph (macroeconomic approach) shows again the huge impact of the mortgage interest deduction on the default portfolio.

$$\boldsymbol{ln(\frac{DR_{White}(t)}{1-DR_{White}(t)}) = -0,73 - 0,25 * HousesSold(t) * 10^{-4} + 0,02 * InflowUnemploymentBusiness(t-1) * 10^{-4} - 0,07 * InflationRate(t-20) + 0,04 * MortgageInterestRate(t) + 0,83 * ln(\frac{DR_{White}(t-1)}{1-DR_{White}(t-1)})}$$

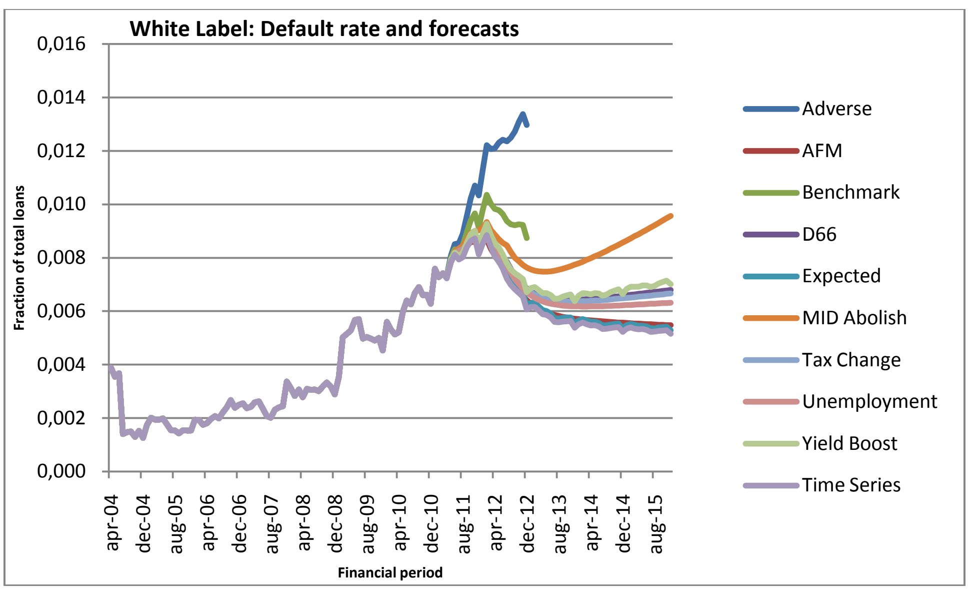


Figure 27: White Label default rate time series and forecasts based on applied scenarios

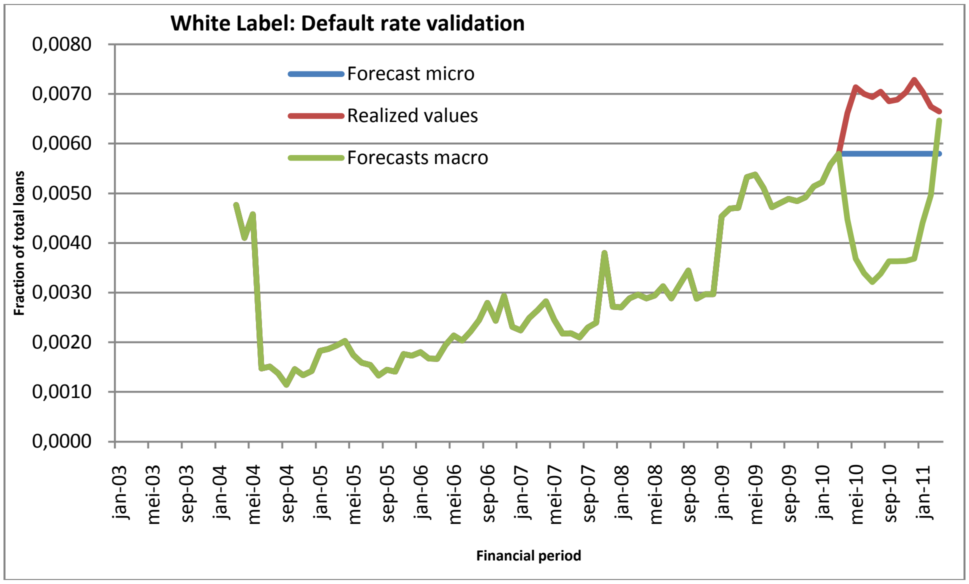


Figure 28: White Label default rate time series validation micro- and macroeconomic approach

The residuals of the White Label model are strange and not comparable to the normal distribution of the Intermediary Channel model residuals. These errors seem to be uniformly distributed between -0,0002 and +0,0002, with relevant outliers. The R-square value is 0,97.

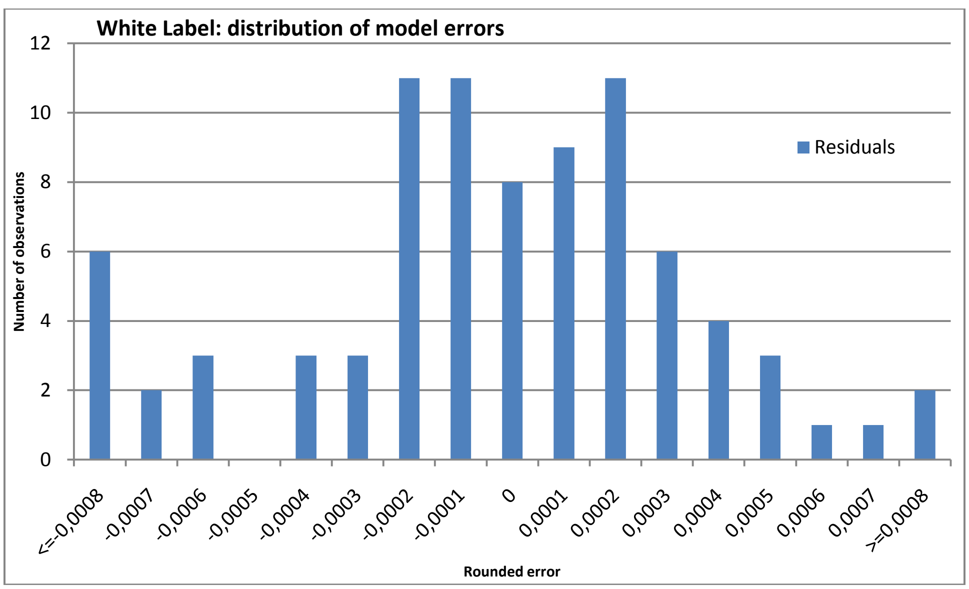


Figure 29: White Label distribution of model errors

6.3 Default rate Consolidated default portfolio

The default rate, as well as the three other rates observed, is quite in line with the Intermediary Channel portfolio, because of the huge impact of Intermediary Channel on the Consolidated Portfolio. The average error in the validation period is very low, in the time series analysis with as well as without external input (both around 2-3% error). The T-test favors the macroeconomic approach significantly.

$$ln\left(\frac{DR_{Cons}(t)}{1-DR_{Cons}(t)}\right) = -0,29 - 0,03 * HousePrices(t) * 10^{-5} - 0,06 * HousesSold(t) * 10^{-4} - 0,02 * OutflowUnemploymentTotal(t-20) * 10^{-3} + 0,01 * MortgageInterestRate(t) + 0,91 * ln\left(\frac{DR_{Cons}(t-1)}{1-DR_{Cons}(t-1)}\right)$$

Most scenarios are between the 0,40% till 0,45% defaults in the upcoming years. After an increase, it is expected that the default rate will decrease a bit. Especially the mortgage interest deduction elimination scenario increases the default rate.

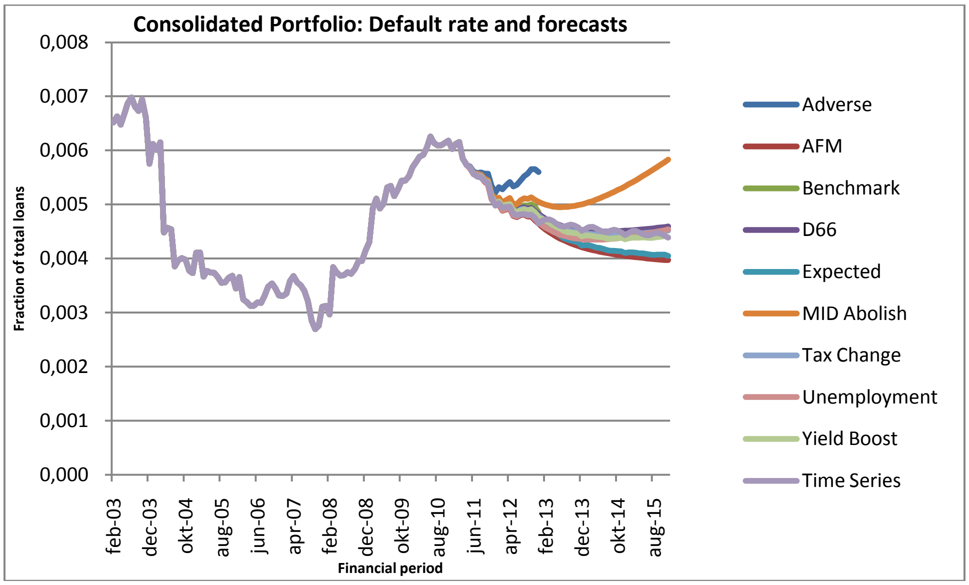


Figure 30: Consolidated Portfolio default rate time series and forecasts based on applied scenarios

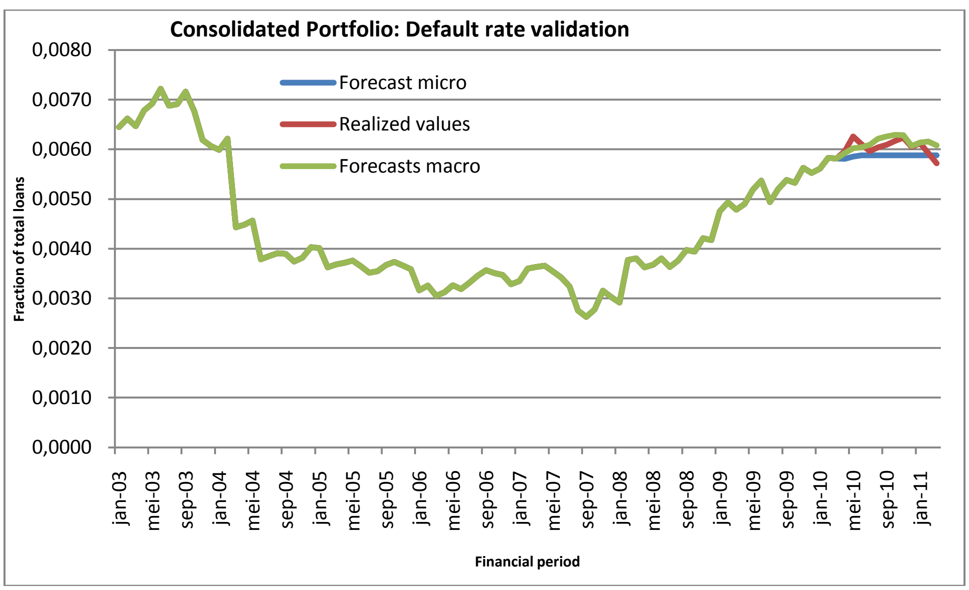


Figure 31: Consolidated Portfolio default rate time series validation micro- and macroeconomic approach

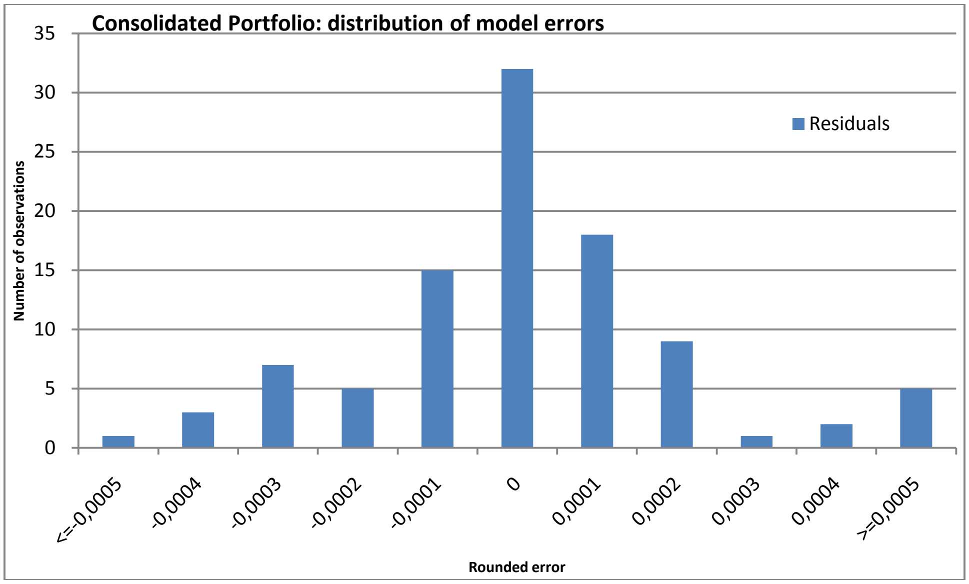


Figure 32: Consolidated Portfolio distribution of model errors

The residuals of the model of the Consolidated Portfolio seem to be normally distributed. The corresponding 95% confidence interval of the errors is between -0,00052 and 0,00056. The R-square value is 0,95.

7. Indicators of the loss rate

An obvious restriction is taken into account before starting the search for indicators within the portfolios that could describe the loss fraction with respect to the total loan value: the intended link to macroeconomic factors has to be arguable, otherwise the analyses of the indicator is meaningless. Sources used are inconsistent and contain fewer data compared to the mutations observed (in the rates), due to the availability of historical data.

The objective is to predict the loss rate, given that the loan is in default. In *Section 7.1* the Loan-to-Value is linked to the losses and to house prices, in *Section 7.2* the Loan-to-Income correlations with unemployment and losses are studied. In *Section 7.3* both are combined to indicate the loss rate in a cross-table and a time dependent approach is proposed by rewriting the LTV-ratio as the loss rate. In *Section 7.4* two possible extensions are discussed, the interest rate and the age of the applicant. This chapter formulates an answer to *subquestion 3* from *Section 1.2*.

7.1 Loan-to-Value

The Loan-to-Value (LTV) is the value of the loan divided by the value of the security, in most cases the house. It is assumed that the Loss Given Default is positively correlated with the Loan-to-Value. The larger the LTV-ratio on a loan is, the larger the loss rate (loss divided by total loan value). In Table 10 the loss fractions classified by LTV-ratio for the labels are shown (time independent).

	Intermediary Channel	**White Label**	**Consolidated Portfolio**
LTV below 60%	0,000	0,002	0,001
LTV between 60% and 70%	0,002	0,000	0,002
LTV between 70% and 80%	0,006	0,031	0,008
LTV between 80% and 90%	0,022	0,047	0,024
LTV between 90% and 100%	0,080	0,074	0,080
LTV between 100% and 110%	0,108	0,109	0,107
LTV between 110% and 120%	0,155	0,122	0,153
LTV above 120%	0,148	0,135	0,147

Table 10: Loss fractions sorted by LTV-ratio

There is a clear relationship between the LTV-ratio and the loss rate. A higher LTV-ratio indicates a higher loss fraction. The only reliable macroeconomic factor that could describe the LTV-value of a portfolio is the factor house prices. Only on loans with a LTV-value above 100%, a loss is obtained. Therefore, time series of defaults with a LTV-ratio equal to or greater than 100% is explained by a constant and the time series of house prices. See Table 11.

Label and rate	**Parameter house prices**	**Time lag (months)**	**Validation error**
Intermediary Channel: LTV defaults LTV>100%	0,116 (constant of 0,784)	30	0,73%
White Label: LTV defaults LTV>100%	0,104 (constant of 0,800)	32	0,52%
Consolidated: LTV defaults LTV>100%	0,106 (constant of 0,801)	30	0,76%

Table 11: LTV (weighted with respect to execution value) and house prices correlation

There exists a very strong relationship between the average house prices, especially with a quite large time lag, and the LTV-ratio. The relationship could even be quantified reliable, resulting in a formula on time t in months (financial period):

$$\text{LTV-ratio}(\text{t}) = c + a(HousePrices(t - y) * 10^{-5}), \qquad \textbf{[Equation 16]}$$

with c representing the constant, a the parameter and y defining the time lag. The start value of t is February 2003, $t = 0$. Table 11 shows the parameters, time lags and corresponding correlations and validation error.

$$\text{LTV-ratio}_{Intermediary}(t) = 0{,}784 + 0{,}116(HousePrices(t - 30) * 10^{-5})$$
$$\text{LTV-ratio}_{White}(t) = 0{,}800 + 0{,}104(HousePrices(t - 32) * 10^{-5})$$
$$\text{LTV-ratio}_{\text{Consolidated}}(t) = 0{,}801 + 0{,}106(HousePrices(t - 30) * 10^{-5})$$

Notice that the errors are small and the forecasts (based only on the house prices) seem to be reliable, but the large time lag restricts the changes on the middle long run, according to the scenarios.

7.2 Loan-to-Income

The Loan-to-Income (LTI) ratio is the total mortgage value divided by the yearly income at the application date. Subdividing the loans of the Intermediary Channel in LTI-classes and calculating the average loss rate for each class, shows a relationship between the losses and the LTI-class. The loss rate is the loss divided by the total value of the loan. For example, in the Consolidated Portfolio the loss of defaults with a loan between 5 and 6 times the yearly income, is estimated to be 10,5 percent. Table 12 shows that a higher LTI-class matches a higher loss, with exceptions in the White Label, because of very few observations.

	Intermediary Channel	**White Label**	**Consolidated portfolio**
LTI below 2	0,006	0,345	0,058
LTI between 2 and 3	0,068	0,051	0,061
LTI between 3 and 4	0,094	0,094	0,094
LTI between 4 and 5	0,098	0,086	0,094
LTI between 5 and 6	0,136	0,081	0,105
LTI above 6	0,121	0,088	0,112

Table 12: LTI-classes and losses

There seems to be a good fit between the economic environment and the Loan-to-Income, which indicates an economic impact. An even better estimator is the seasonal corrected unemployment rate. See Table 13. There, unfortunately, arises a problem in this analysis: there is no documentation about unemployment as a reason of default and the yearly income is only known at the application moment. It is might useful to know that a higher LTI, leads to higher loss fractions, but is seems to be unreliable to forecast losses on the LTI-ratio.

Label and rate	**Parameter unemployment**	**Time lag (months)**	**Validation error**
Intermediary Channel: LTI defaults	-0,09 (constant of 4,81)	22	0,94%
White Label: LTI overall	-0,09 (constant of 4,43)	19	1,01%
Consolidated: LTI defaults	-0,08 (constant of 4,87)	21	0,81%

Table 13: LTI (class with LTI-values above 3) and seasonal corrected unemployment rate correlation

7.3 Loss rate based on LTV (and LTI?)

From *Sections 7.1 and 7.2* the loss rates can be derived, if the results were not completely different. Both indicators, LTI and LTV, describe the loss, but in a different way. For example, in the Consolidated Portfolio the LTV is about 85% and the LTI (class above 3) is about 4,5. According to the LTV-ratio, the loss rate is about 2,4 percent. Otherwise, focusing on the LTI, results in a loss of 9,4 percent. The difference observed is enormous. Therefore, a cross table can be drawn for the portfolios. It is expected, and approved by Table 14, that the combination 'high LTV' and 'high LTI' results in a high loss rate, and that the opposite is also true. The only problem that remains is the time independence.

Intermediary: Loss Rate	*LTI<2*	*2<LTI<3*	*3<LTI<4*	*4<LTI<5*	*5<LTI<6*	*6<LTI*	**Total**
LTV below 60%	0,00	0,00	0,00	0,00	0,00	0,00	*0,00*
LTV between 60% and 70%	0,00	0,00	0,00	0,01	0,00	0,00	*0,00*
LTV between 70% and 80%	0,00	0,00	0,02	0,00	0,00	0,01	*0,01*
LTV between 80% and 90%	0,00	0,05	0,05	0,01	0,00	0,00	*0,03*
LTV between 90% and 100%	0,00	0,05	0,10	0,06	0,11	0,18	*0,09*
LTV between 100% and 110%	0,07	0,16	0,10	0,15	0,08	0,16	*0,13*
LTV between 110% and 120%	0,02	0,14	0,21	0,13	0,27	0,28	*0,19*
LTV above 120%	0,00	0,55	0,12	0,23	0,15	0,21	*0,18*
Total	*0,01*	*0,07*	*0,09*	*0,10*	*0,14*	*0,12*	*0,10*
White Label: Loss Rate	*LTI<2*	*2<LTI<3*	*3<LTI<4*	*4<LTI<5*	*5<LTI<6*	*6<LTI*	**Total**
LTV below 60%	0,00	0,00	0,00	0,00	0,01	0,00	*0,00*
LTV between 60% and 70%	0,00	0,00	0,00	0,00	0,00	0,00	*0,00*
LTV between 70% and 80%	0,00	0,00	0,00	0,03	0,00	0,01	*0,01*
LTV between 80% and 90%	0,00	0,11	0,10	0,03	0,02	0,08	*0,04*
LTV between 90% and 100%	0,00	0,09	0,10	0,07	0,10	0,04	*0,08*
LTV between 100% and 110%	0,00	0,08	0,12	0,10	0,10	0,11	*0,10*
LTV between 110% and 120%	0,00	0,12	0,10	0,10	0,12	0,18	*0,12*
LTV above 120%	0,00	0,12	0,12	0,16	0,05	0,17	*0,12*
Total	*0,00*	*0,05*	*0,09*	*0,09*	*0,08*	*0,09*	*0,08*
Consolidated: Loss Rate	*LTI<2*	*2<LTI<3*	*3<LTI<4*	*4<LTI<5*	*5<LTI<6*	*6<LTI*	**Total**
LTV below 60%	0,00	0,00	0,00	0,00	0,01	0,00	*0,00*
LTV between 60% and 70%	0,00	0,00	0,00	0,01	0,00	0,00	*0,00*
LTV between 70% and 80%	0,00	0,00	0,01	0,01	0,00	0,01	*0,01*
LTV between 80% and 90%	0,00	0,06	0,06	0,02	0,01	0,05	*0,03*
LTV between 90% and 100%	0,00	0,07	0,10	0,07	0,11	0,09	*0,09*
LTV between 100% and 110%	0,07	0,13	0,12	0,13	0,09	0,14	*0,12*
LTV between 110% and 120%	0,02	0,13	0,17	0,12	0,19	0,20	*0,16*
LTV above 120%	0,00	0,23	0,12	0,19	0,09	0,26	*0,15*
Total	*0,01*	*0,06*	*0,09*	*0,10*	*0,11*	*0,11*	*0,09*

Table 14: Cross-table loss rates depend on LTV and LTI

A better option would be to use the LTV-ratio and rewrite the ratio as a loss rate. Several steps are assumed for this reformulation:

$$\text{LTV-ratio} = \frac{Loan\ value}{Security\ value} \qquad \textbf{[Equation 17]}$$

Suppose the loan is in default, then:

$$\frac{Loan\ value}{Security\ value} = \frac{Exposure\ at\ Default}{Market\ Value\ of\ the\ security} = \frac{1}{Fraction\ that\ can\ be\ paid\ back\ to\ the\ bank\ (initially\ called\ cure\ rate)}$$

[Equation 18]

When the cure rate is above 1, there is no loss on the loan, the loss rate is zero. On all loans with a fraction below 1, the loss rate is equal to the remainder:

$$Loss\ rate = \max\,(1 - \text{cure rate}, 0) \qquad \textbf{[Equation 19]}$$

Thus, the LTV-ratio for defaults with a LTV-ratio equal to or greater than 1 (i.e. 100%) can be written as:

$$\text{LTV-ratio}\ on\ default = \frac{1}{cure\ rate} = \frac{1}{1 - loss\ rate} = \frac{1}{1 - \max\,(1 - \text{cure rate}, 0)}$$

[Equation 20]

For defaults the loss rate can be expressed depend on the LTV-ratio for each period:

$$Loss\ rate\ on\ defaults(t) = \max\,(0{,}1 - \frac{1}{LTV - ratio(t)}) \qquad \textbf{[Equation 21]}$$

7.4 Extensions of the loss rate

Some extensions of the loss rate are proposed to improve the correctness over time. In *Section 7.4.1* the interest rate is studied, in *Section 7.4.2* the age development of the population.

7.4.1 Interest rate

The interest rate clearly affects the Probability of Default (PD) as seen earlier, but it is the question if there is a relationship with the default losses too. According to the table created for Intermediary Channel between 2006 and June 2009, the interest rate is out of scope in this context.

This decision is mainly due to the hypothesis that higher interest percentages could generate a higher loss. There appears no increase in losses when the interest rate increases. Therefore, the hypothesis failed and interest input is out of scope for the loss rate. The interest rate is mainly affected by the time of acceptance and therefore the changes in interest over time could initiate problems, but the interest rate on itself is independent of the loss fraction.

Weighted interest	Fraction total loans	Total loss	Number of defaults	Average loss based on number of loans	Average loss based on number of defaults
Below 4%	0,047	€ 528.547	32	€ 43,57	€ 16.517
4 to 4,5%	0,144	€ 1.406.227	91	€ 38,30	€ 15.453
4,5 to 5%	0,249	€ 3.347.222	148	€ 52,53	€ 22.616
5 to 5,5%	0,223	€ 3.601.287	139	€ 63,31	€ 25.909
5,5 to 6%	0,173	€ 569.705	37	€ 12,91	€ 15.397
Above 6%	0,164	€ 624.715	51	€ 15,66	€ 12.249

Table 15: Interest-classes and corresponding losses Intermediary Channel

7.4.2 Population age

Although not very obvious, differences in average losses might be related to age classes, as shown in Table 16. These data is time independent and can be translated into a parameter for the expected average losses in the future, by connecting the classes to population mutations, if there appears a difference in loss related to age.

	Intermediary Channel		White Label		Consolidated Portfolio	
Age classes	**Number of loans**	**Average loss given default**	**Number of loans**	**Average loss given default**	**Number of loans**	**Average loss given default**
<30	60	€ 17.309	117	€ 17.848	366	€ 14.148
31-35	71	€ 27.281	110	€ 14.722	314	€ 15.292
36-40	91	€ 23.830	140	€ 11.055	369	€ 15.498
41-45	77	€ 20.741	150	€ 11.242	357	€ 14.386
46-50	58	€ 16.064	129	€ 11.985	272	€ 14.537
51<	140	€ 17.178	166	€ 7.077	385	€ 10.770

Table 16: Age-classes and corresponding losses

According to Table 16, there is no convinced difference in average loss related to age. In the Intermediary Channel portfolio the highest loss is observed in the class with applicants between 30 and 35 years, in the White Label portfolio the highest loss is observed in the youngest class and in the Consolidated Portfolio the highest loss is observed in the class between 36 and 40 years. Knowing that the Consolidated Portfolio largely depends on the Intermediary Channel and White Label, the differences are strange. Therefore, it is concluded that also the age does not play an active role in the credit loss expectation.

8. Models for credit losses

In this chapter the default rate *(Chapter 6)*, the loss rate *(Chapter 7)* and the exposure *(Section 8.1)* are combined with respect to the covariance *(Section 8.2)*. The expected credit losses are modeled in *Section 8.3*. This the answer to *subquestion 4* from *Section 1.2*.

In *Section 8.4* the fraction of new default in the default portfolio, according to the scenarios, is calculated. Once in default, there are two ways out: recovery or foreclosure. The ratio between those is calculated in the same section. In *Section 8.5* a critical note for using these models is added.

8.1 Exposure value

The exposure value (EV) gives an indication of the outstanding value of the portfolio. For credit losses, it is important to look at the exposure of defaults only. Therefore the exposure value (EV) is used for the total exposure of defaults. No macroeconomic factor is a reliable explanatory variable for EV. Therefore, the EV of the last 24 months is simple extrapolated using MS Excel, which means that a constant is calculated based on the last 24 months. This value is added to financial period t-1 to estimate the value on financial period t. The results are tabulated below.

	Intermediary Channel			
	Average exposure defaults	*Total exposure defaults*	*Average exposure portfolio*	*Total exposure portfolio*
dec-10	€ 208.878,38	€ 135.770.949,02	€ 174.166,54	€ 31.824.233.497,90
dec-11	€ 210.294,37	€ 157.196.400,47	€ 175.013,16	€ 32.421.778.889,17
dec-12	€ 213.346,19	€ 179.076.880,34	€ 175.905,76	€ 33.135.685.718,77
dec-13	€ 216.398,01	€ 200.957.360,22	€ 176.798,36	€ 33.849.592.548,37
dec-14	€ 219.449,83	€ 222.837.840,09	€ 177.690,97	€ 34.563.499.377,97
dec-15	€ 222.501,65	€ 244.718.319,96	€ 178.583,57	€ 35.277.406.207,57
	White Label			
	Average exposure defaults	*Total exposure defaults*	*Average exposure portfolio*	*Total exposure portfolio*
dec-10	€ 214.894,17	€ 45.342.670,54	€ 167.501,21	€ 4.853.347.420,28
dec-11	€ 224.316,01	€ 50.662.588,05	€ 165.989,02	€ 4.516.641.897,37
dec-12	€ 231.596,08	€ 58.186.726,31	€ 164.449,04	€ 4.176.694.381,54
dec-13	€ 238.876,15	€ 65.710.864,58	€ 162.909,06	€ 3.836.746.865,71
dec-14	€ 246.156,22	€ 73.235.002,85	€ 161.369,08	€ 3.496.799.349,88
dec-15	€ 253.436,29	€ 80.759.141,12	€ 159.829,10	€ 3.156.851.834,04
	Consolidated Portfolio			
	Average exposure defaults	*Total exposure defaults*	*Average exposure portfolio*	*Total exposure portfolio*
dec-10	€ 207.025,72	€ 194.811.203,43	€ 172.915,38	€ 39.186.428.674,89
dec-11	€ 210.180,95	€ 222.067.148,55	€ 173.350,27	€ 39.332.902.945,78
dec-12	€ 213.354,83	€ 250.575.342,62	€ 173.799,14	€ 39.584.795.631,77
dec-13	€ 216.528,71	€ 279.083.536,70	€ 174.248,01	€ 39.836.688.317,77
dec-14	€ 219.702,58	€ 307.591.730,77	€ 174.696,88	€ 40.088.581.003,76
dec-15	€ 222.876,46	€ 336.099.924,84	€ 175.145,74	€ 40.340.473.689,76

Table 17: Exposure values extrapolated for the three portfolios

Multiplying the default rate, loss rate and average exposure indicates the expected loss on a loan. When multiplying with the total exposure instead of the average, the total expected credit loss is estimated. It can be argued that a relationship between DR and LR exist. In the next section, a correction for this relationship is proposed. The corrected value of DR times LR is multiplied with the average exposure to estimate the expected loss on a loan (EL) and with the total exposure to estimate the credit losses of a portfolio (CL).

8.2 Covariance correction

Equation 1 from *Section 3.2.3* is applied in this section. It is impossible to determine the covariance on individual loans, because the DR is modeled on the aggregated level. Therefore it is chosen to calculate the covariance with a moving average of the DR and LR of six values (i.e. financial periods).

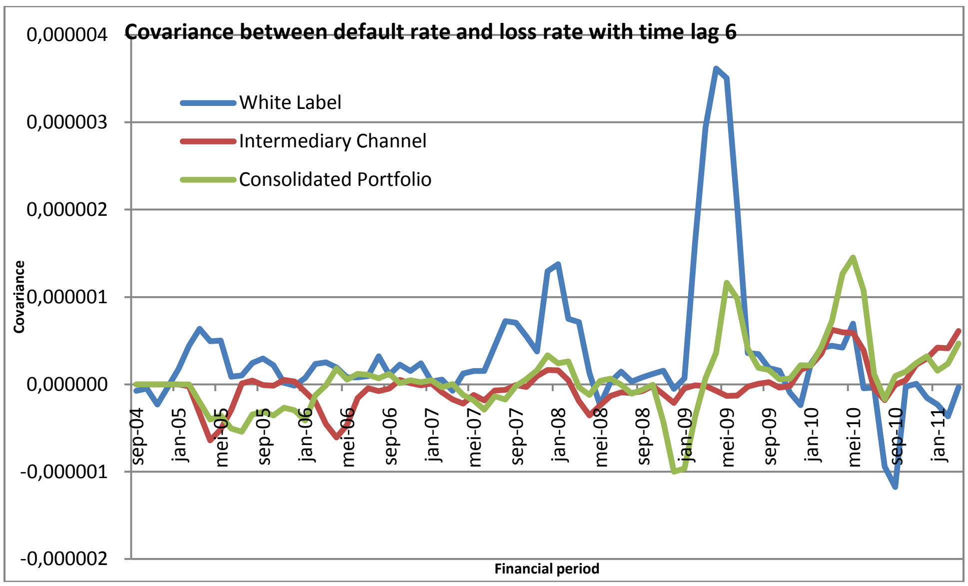


Figure 33: Covariance between default rate and loss rate

Predicting the covariance seems to be tricky, because the covariance should be added to the multiplication of DR and LR to correct for under- or overestimation. When the same (or derived) input factors are used to calculate the DR and LR, a correction could be argued. On the other hand, the default rate and loss rate are constructed with macroeconomic factors based on historical data and therefore implicitly corrected (including a covariance leads to a correction of the history).

The White Label covariance is more volatile than both other portfolio covariances. The scale values (Figure 33) indicate the effect on the total solution: very small influences. Therefore, a decision is not hard to made, because the inclusion of the covariance is hardly not affecting the results. In the remainder of this chapter the covariance correction is applied, but the results of the recent data would only differ sometimes in the second decimal if this correction was not applied.

8.3 Expected loss and total credit losses

The expected loss on a loan (EL) is the multiplication of the default rate (DR), the loss rate (LR) and the average exposure value on a defaulted loan (in this chapter with covariance correction, which does not change the results clearly). The scenario analyses show the average expected loss on a loan in Figure 34 for the Intermediary Channel and for the White Label and Consolidated Portfolio in Figures 35 and 36, respectively. In Appendix B the total credit losses can be found.

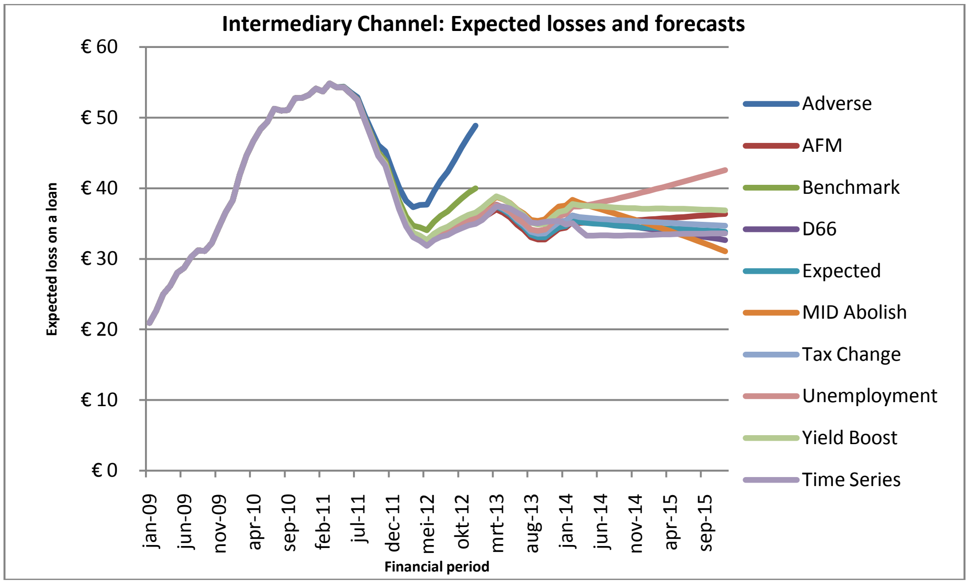


Figure 34: Average expected loss on a single loan for the Intermediary Channel

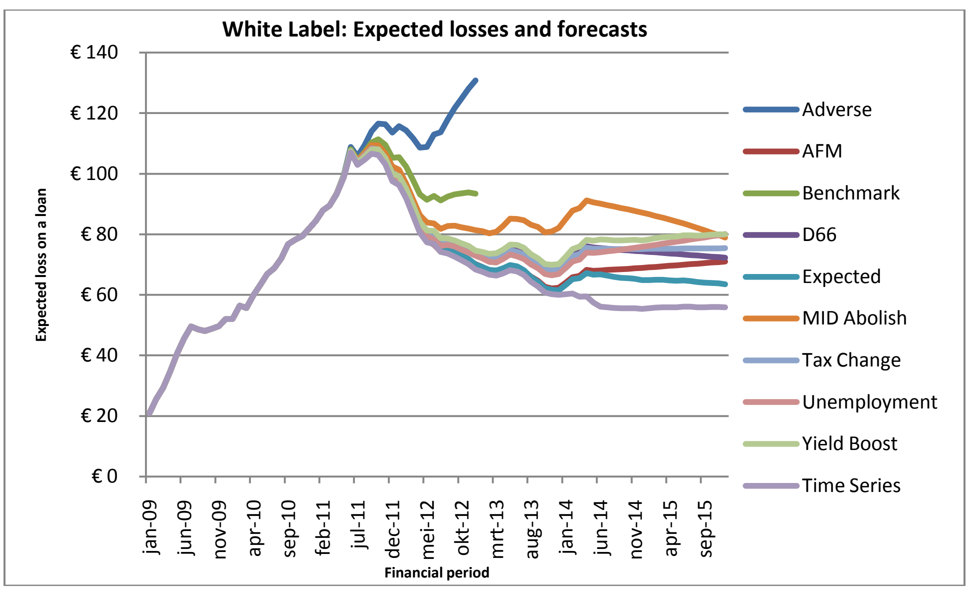


Figure 35: Average expected loss on a single loan for the White Label

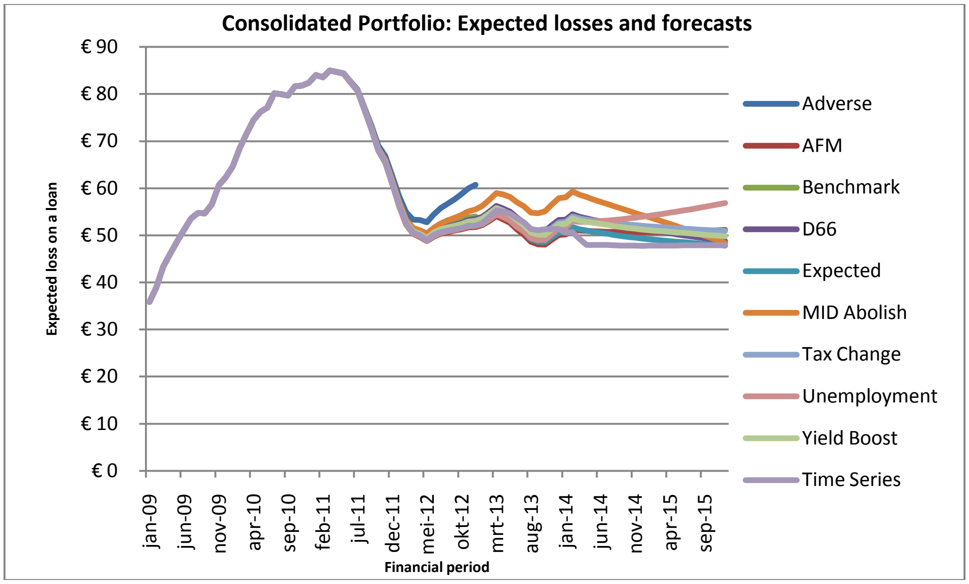


Figure 36: Average expected loss on a single loan for the Consolidated Portfolio

Int Ch	*Adverse*	*AFM*	*Benchmark*	*D66*	*Expected*	*MID Abolish*	*Tax Change*	*Unemployment*	*Yield Boost*	*Time Series*
dec-09	€ 36,58	€ 36,58	€ 36,58	€ 36,58	€ 36,58	€ 36,58	€ 36,58	€ 36,58	€ 36,58	€ 36,58
dec-10	€ 53,21	€ 53,21	€ 53,21	€ 53,21	€ 53,21	€ 53,21	€ 53,21	€ 53,21	€ 53,21	€ 53,21
dec-11	€ 42,53	€ 40,19	€ 40,94	€ 40,28	€ 40,23	€ 40,39	€ 40,26	€ 40,23	€ 40,39	€ 40,02
dec-12	€ 48,88	€ 35,27	€ 39,98	€ 35,76	€ 35,46	€ 36,50	€ 35,67	€ 35,73	€ 36,56	€ 34,94
dec-13		€ 34,22		€ 35,25	€ 34,48	€ 37,41	€ 35,19	€ 36,00	€ 36,63	€ 35,37
dec-14		€ 35,52		€ 34,39	€ 34,42	€ 35,23	€ 35,29	€ 39,38	€ 37,10	€ 33,32
dec-15		€ 36,37		€ 32,67	€ 33,82	€ 31,06	€ 34,68	€ 42,58	€ 36,87	€ 33,59
White	*Adverse*	*AFM*	*Benchmark*	*D66*	*Expected*	*MID Abolish*	*Tax Change*	*Unemployment*	*Yield Boost*	*Time Series*
dec-09	€ 52,03	€ 52,03	€ 52,03	€ 52,03	€ 52,03	€ 52,03	€ 52,03	€ 52,03	€ 52,03	€ 52,03
dec-10	€ 81,74	€ 81,74	€ 81,74	€ 81,74	€ 81,74	€ 81,74	€ 81,74	€ 81,74	€ 81,74	€ 81,74
dec-11	€ 113,58	€ 99,09	€ 105,24	€ 100,36	€ 99,25	€ 102,42	€ 100,25	€ 99,93	€ 100,05	€ 97,54
dec-12	€ 130,80	€ 70,16	€ 93,45	€ 74,28	€ 70,01	€ 81,32	€ 73,90	€ 72,87	€ 74,64	€ 68,37
dec-13		€ 62,30		€ 69,33	€ 61,45	€ 82,14	€ 68,67	€ 66,88	€ 70,13	€ 59,97
dec-14		€ 68,92		€ 74,35	€ 64,91	€ 87,17	€ 75,18	€ 75,87	€ 78,01	€ 55,39
dec-15		€ 71,02		€ 72,25	€ 63,50	€ 78,89	€ 75,40	€ 80,05	€ 79,94	€ 55,89
Cons P	*Adverse*	*AFM*	*Benchmark*	*D66*	*Expected*	*MID Abolish*	*Tax Change*	*Unemployment*	*Yield Boost*	*Time Series*
dec-09	€ 62,28	€ 62,28	€ 62,28	€ 62,28	€ 62,28	€ 62,28	€ 62,28	€ 62,28	€ 62,28	€ 62,28
dec-10	€ 82,40	€ 82,40	€ 82,40	€ 82,40	€ 82,40	€ 82,40	€ 82,40	€ 82,40	€ 82,40	€ 82,40
dec-11	€ 62,54	€ 60,88	€ 61,25	€ 61,22	€ 60,96	€ 61,63	€ 61,15	€ 60,93	€ 61,06	€ 60,91
dec-12	€ 60,75	€ 51,78	€ 53,98	€ 53,44	€ 52,20	€ 55,48	€ 53,10	€ 52,01	€ 53,01	€ 52,02
dec-13		€ 50,08		€ 53,24	€ 50,54	€ 57,89	€ 52,70	€ 51,41	€ 52,28	€ 51,37
dec-14		€ 50,67		€ 51,35	€ 49,28	€ 54,30	€ 52,03	€ 53,94	€ 51,20	€ 47,77
dec-15		€ 51,13		€ 48,78	€ 47,82	€ 48,34	€ 50,94	€ 56,86	€ 49,78	€ 47,88

Table 18: Expected losses on a loan for the three portfolios

8.4 Notes to the related rates

The reason that the inflow rate, recovery rate and foreclosure rates were taken into account was to analyze the possible changes in mutual proportions. The inflow fraction divided by the total defaults fraction indicates the fraction of new defaults in the default portfolio (called throughput of defaults). When this fraction increases, the inflow with respect to the total defaults is increasing. Table 19 shows the fractions according to the selected scenarios.

Intermediary Channel										
Financial period	**Adverse**	**AFM**	**Benchmark**	**D66**	**Expected**	**MID**	**Tax**	**Unemployment**	**Yield**	**Time Series**
dec-11	0,24	0,18	0,19	0,19	0,18	0,19	0,19	0,18	0,19	0,18
dec-12	0,33	0,18	0,21	0,19	0,19	0,21	0,19	0,18	0,20	0,19
dec-13		0,20		0,22	0,21	0,24	0,22	0,20	0,23	0,19
dec-14		0,20		0,23	0,22	0,25	0,22	0,19	0,24	0,19
dec-15		0,20		0,24	0,22	0,26	0,22	0,18	0,26	0,19
White Label										
Financial period	**Adverse**	**AFM**	**Benchmark**	**D66**	**Expected**	**MID**	**Tax**	**Unemployment**	**Yield**	**Time Series**
dec-11	0,08	0,11	0,09	0,11	0,09	0,11	0,11	0,11	0,09	0,09
dec-12	0,08	0,14	0,10	0,14	0,12	0,13	0,14	0,14	0,12	0,13
dec-13		0,16		0,15	0,14	0,14	0,15	0,15	0,13	0,15
dec-14		0,16		0,15	0,15	0,13	0,15	0,15	0,13	0,15
dec-15		0,16		0,15	0,15	0,12	0,15	0,15	0,13	0,16
Consolidated Portfolio										
Financial period	**Adverse**	**AFM**	**Benchmark**	**D66**	**Expected**	**MID**	**Tax**	**Unemployment**	**Yield**	**Time Series**
dec-11	0,25	0,20	0,21	0,21	0,21	0,21	0,20	0,20	0,21	0,20
dec-12	0,32	0,21	0,23	0,22	0,22	0,23	0,22	0,22	0,22	0,21
dec-13		0,23		0,24	0,24	0,25	0,24	0,24	0,24	0,22
dec-14		0,24		0,25	0,26	0,25	0,24	0,24	0,25	0,22
dec-15		0,24		0,25	0,26	0,25	0,25	0,23	0,25	0,23

Table 19: Throughput of defaults (inflow default fraction/default fraction) for the three portfolios

Table 19 shows that only in abnormal (stress) scenarios the fraction changes clearly. Most of the scenarios are dealing with steady throughput fractions.

An increasing foreclosure given default fraction could affect the credit losses negatively. The opposite might be true too: when more defaults recover the losses will probably decline. So, the relation between the RR and FR is interesting. In Table 20 the foreclosure fraction divided by the sum of the foreclosure and recovery fraction is presented. Only in the steady scenarios (AFM, Expected and Time Series) there is no clear change expected, in other scenarios there is –broadly spoken – an increase of foreclosure expected for defaulted loans.

Intermediary Channel										
Financial period	**Adverse**	**AFM**	**Benchmark**	**D66**	**Expected**	**MID**	**Tax**	**Unemployment**	**Yield**	**Time Series**
dec-11	0,29	0,24	0,27	0,24	0,24	0,25	0,24	0,25	0,24	0,28
dec-12	0,46	0,24	0,36	0,25	0,24	0,27	0,25	0,27	0,25	0,28
dec-13		0,24		0,25	0,25	0,29	0,25	0,29	0,25	0,28
dec-14		0,24		0,26	0,25	0,32	0,26	0,32	0,25	0,28
dec-15		0,24		0,26	0,25	0,36	0,27	0,34	0,25	0,28
White Label										
Financial period	**Adverse**	**AFM**	**Benchmark**	**D66**	**Expected**	**MID**	**Tax**	**Unemployment**	**Yield**	**Time Series**
dec-11	0,40	0,08	0,17	0,09	0,08	0,09	0,08	0,08	0,10	0,08
dec-12	0,95	0,11	0,51	0,12	0,11	0,14	0,12	0,12	0,21	0,15
dec-13		0,10		0,11	0,10	0,14	0,11	0,12	0,25	0,14
dec-14		0,10		0,12	0,10	0,17	0,11	0,13	0,32	0,14
dec-15		0,09		0,12	0,10	0,20	0,12	0,14	0,40	0,14
Consolidated Portfolio										
Financial period	**Adverse**	**AFM**	**Benchmark**	**D66**	**Expected**	**MID**	**Tax**	**Unemployment**	**Yield**	**Time Series**
dec-11	0,40	0,35	0,38	0,35	0,35	0,35	0,35	0,35	0,35	0,35
dec-12	0,63	0,35	0,50	0,35	0,35	0,38	0,36	0,39	0,35	0,34
dec-13		0,35		0,35	0,35	0,40	0,36	0,42	0,35	0,34
dec-14		0,35		0,36	0,35	0,42	0,36	0,45	0,36	0,35
dec-15		0,35		0,36	0,34	0,44	0,37	0,48	0,36	0,35

Table 20: Foreclosure fraction with respect to the sum of foreclosure and recovery fraction

8.5 “The future will be better tomorrow.”

(Quote by Dan Quayle)

Models are never perfect and the future is hard to forecast well. This makes the estimates not *true*. In the Consolidated Portfolio another problem arises: this portfolio is a combination of several portfolios with different targets and products. Macroeconomic factors are continuing influencing the risk parameters of the bank, mainly through indirect channels. The causes of changes in risk parameters over time will be found in macroeconomic themes, partly. The remainder of fluctuations is due to very little external influences, inexplicable results, and oversimplification of the model. For example, think about the sample of mortgage applicants.

Another note might be necessary for the understanding of the model fairness. Over the years, several changes in the organization caused structural differences in the models that could not be corrected at this time. Another change is the outsourcing of collections activities, about one and a half year ago. An undefined increase of defaults and some data problems arose in the process.

9. Conclusions

In this experimental study the default rates, loss rates and expected losses are predicted under different circumstances, i.e. scenarios. The goal was to get more insight in the future developments of the mortgage portfolio by linking the characteristics of the portfolios - Intermediary Channel, White Label and a Consolidated Portfolio are in scope - to macroeconomic factors to analyze effects of macroeconomic changes on the mortgage portfolio. This experimental approach to model scenarios that not necessarily assume a normal market development, was an improvement for the DR of the Intermediary Channel and Consolidated Portfolio compared to forecasts on the history of the rate. The White Label behaves different and the macroeconomic inputs did not perform better estimates. The default rates observed were mainly affected by unemployment and interest rates. House prices and the number of houses sold also affect the development of the probability of default.

The loss rates are highly influenced by the Loan-to-Value- and Loan-to-Income-ratios. Because the LTI-ratio is highly uncertain due to the lack of information about the employment of consumers, it is hard to use the LTI-ratio as an explanatory portfolio variable. A cross-table with LTV- and LTI-classes reflects the dependence of the loss rate on both, but this is a time independent approach.

Another approach, that is favored in this thesis, is to rewrite the LTV-ratio, that can be linked reliable to house prices, to a LR. After multiplication with the DR, a covariance factor is added to correct for redundancy. The Intermediary Channel is hardly affected by the correction, but the White Label is clearly influenced. The covariance of the Consolidated Portfolio is, as expected, between those. See Figure 37.

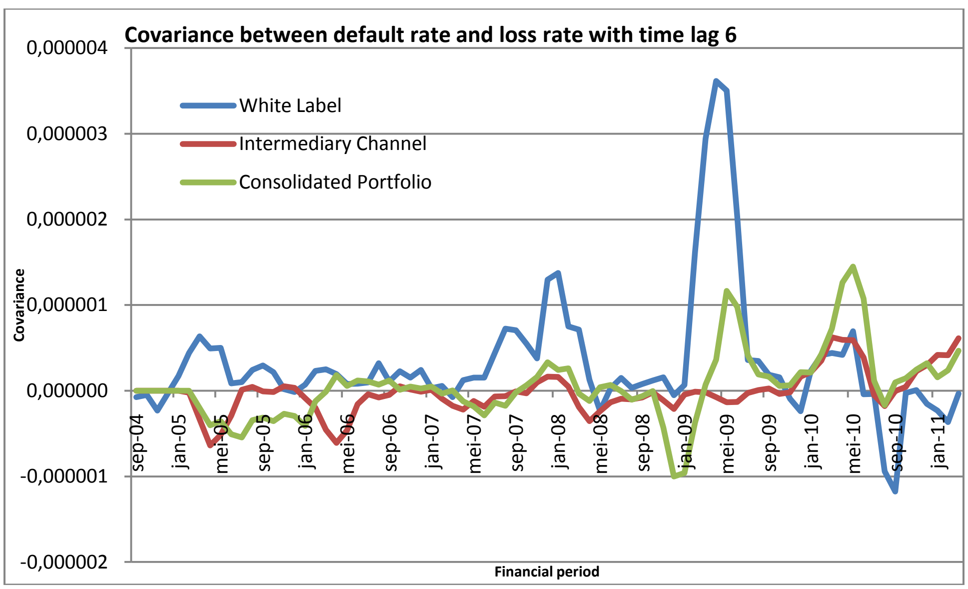


Figure 37: Covariance estimates for the portfolios

The extrapolated average exposure value in the portfolio for defaulted loans is multiplied with the corrected multiplication of DR and LR to estimate the expected loss on a loan. The multiplication consists of the probability of a loan to be in default in a certain financial period, the fraction of the loan that is lost given that the loan is in default, a correction of the DR and LR based on the covariance and the extrapolated average exposure at default. To estimate the total expected credit loss the average exposure of defaults is replaced by the total exposure of defaults. The expected loss on a single loan of the Intermediary Channel is presented in Figure 38 and the estimated credit loss for the White Label in Figure 39.

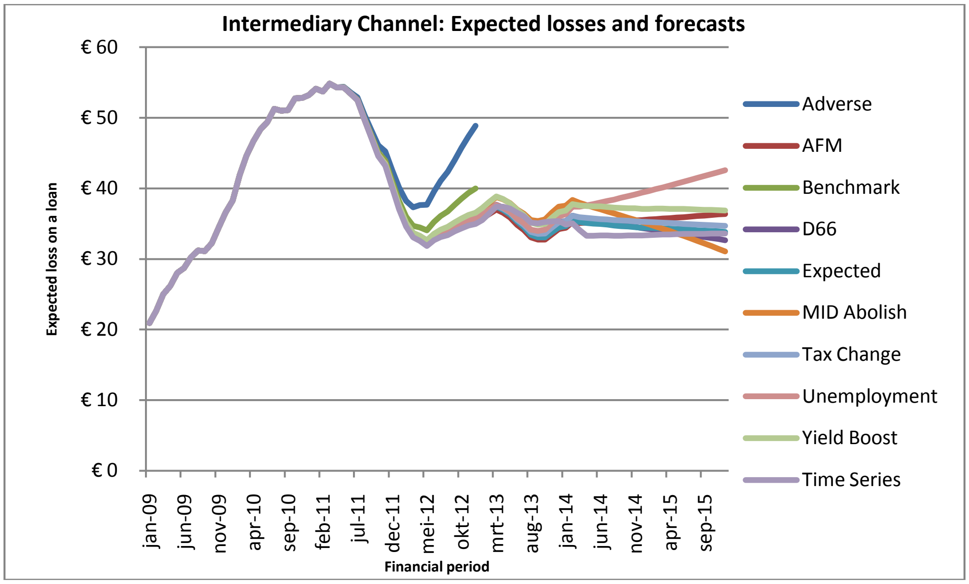


Figure 38: Average expected loss on a single loan for the Intermediary Channel

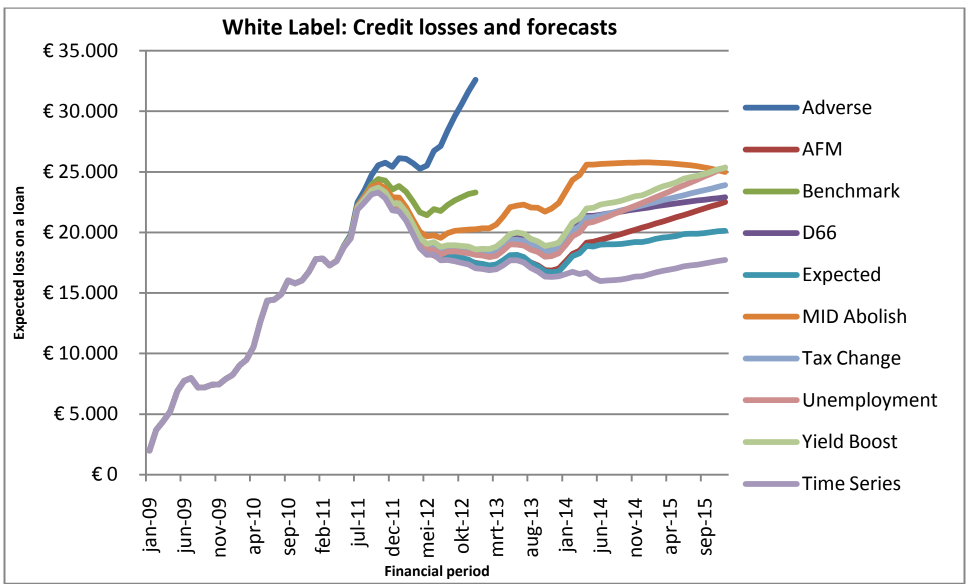


Figure 39: White Label credit losses and forecasts

The main research question was defined as: *What is the influence of macroeconomic factors on the risk parameters for the mortgage portfolios?*

The DR is mainly affected by the scenario variables (house prices, housing market, interest rate and unemployment), the LR is based on the house prices and the exposure is extrapolated. The

DR reacts highly on a change in interest (a higher rate leads to a higher probability of default) and unemployment (a higher unemployment rate leads to a higher probability of default). House prices and the housing market do affect, but not as strong as the interest rate and unemployment. The LR is based on the house prices with a LTV-ratio as intermediate step (higher house prices lead to higher loss rates).

Looking at the credit losses, the elimination of the mortgage interest deduction (MID) is affecting the default portfolios clearly negative. Of course, both included stress scenarios are worse. The effects of increasing unemployment or a higher interest rate appear later, but are also very strong. House price changes and a change in the transactions of houses have smaller effects.

References

Alexander, C. (2001), Market Models

Centraal Bureau voor de Statistiek (2010), Seizoen- en werkdagencorrectie op kwartaalcijfers over de economische groei

Centraal Plan Bureau (2010), Economische Verkenning 2011-2015, No 203

Cheung, Yin-Wong. & Lai, Kon S.(1998), Power of the augmented dickey-fuller test with information-based lag selection, Journal of Statistical Computation and Simulation, 60: 1, 57 — 65

Dankers & Frank (2011), D&F Marktanalyse van Koopwoningen, Vierde Kwartaal 2010, Supplement: In de greep van de mondiale crisis

De Jong, A. & Van Duin, C. (2010), Regionale prognose 2009-2040: Vergrijzing en omslag van groei naar krimp

De Nederlandsche Bank & Autoriteit Financiële Markten (2009), Risico's op de hypotheekmarkt voor huishoudens en hypotheekverstrekkers

Economisch Bureau ABN AMRO (2010), Update woningmarkt, Huizenmarkt: hoog gewaardeerd, maar geen zeepbel

ECORYS (2005), Effecten beperking fiscale aftrek hypotheekrente

Figueiredo, E., et al (2011), Influence of the Autoregressive Model Order on Damage Detection, Computer-Aided Civil and Infrastructure Engineering 26 (2011), 225-238

Khashei, M & Bijari, M. (2011), A novel hybridization of artificial neural networks and ARIMA models for time series forecasting, Applied Soft Computing 11 (2011) 2664-2675

Kiene, N. (2010), Snapshot of Dutch Economy, ABN AMRO

Koehler, A.B. & Murphree, E.S. (1988), A comparison of the Akaike and Schwarz criteria for selecting model order, Applied Statistics, 37: 2, 187-195

Lee, Y-S & Tong L-I (2010), Forecasting time series using a methodology based on autoregressive integrated moving average and genetic programming

NVB (2008), De staat van de Nederlandse woningmarkt, Opinar november 2008 nr. 2

Pagano, M. (1996), Logistic Regression, Nutricion vol. 12 no. 2 1996

Phanos Capital Group (2011), Analyse van de Nederlandse woningmarkt

Rabobank (2011), Kwartaalbericht Woningmarkt Februari 2011

Roubos, A. (2008), Value New Business van Separate Accounts: Tijdreeksanalyse van vermogensbeheer, stageverslag Vrije Universiteit Amsterdam

SWOV (2010), De analyse van tijdreeksen, SWOV- Factsheet Wetenschappelijk Onderzoek Verkeersveiligheid

Van de Pas, L. (2010), Praktische fiscale oplossingen voor de Woningmarkt in Nederland, Woningmarktcijfers.nl

Wegwijs.nl (2010), Woningmarkt stort in januari in, http://www.wegwijs.nl/3882/

Appendix A: Models of inflow, recovery and foreclosure rates

A1: Intermediary Channel

A1.1 Inflow Rate

It was investigated that the best model to describe the inflow rate for Intermediary Channel, according to the validation period and parameter signs, depends on (1) house prices, (2) houses sold, (3) inflow registered inflow (other), (4) mortgage yield rate and a constant and an autoregressive term of the inflow rate itself. Because the T-test indicates no significant difference in results, both models can be used, although the microeconomic approach provides results with a lower error.

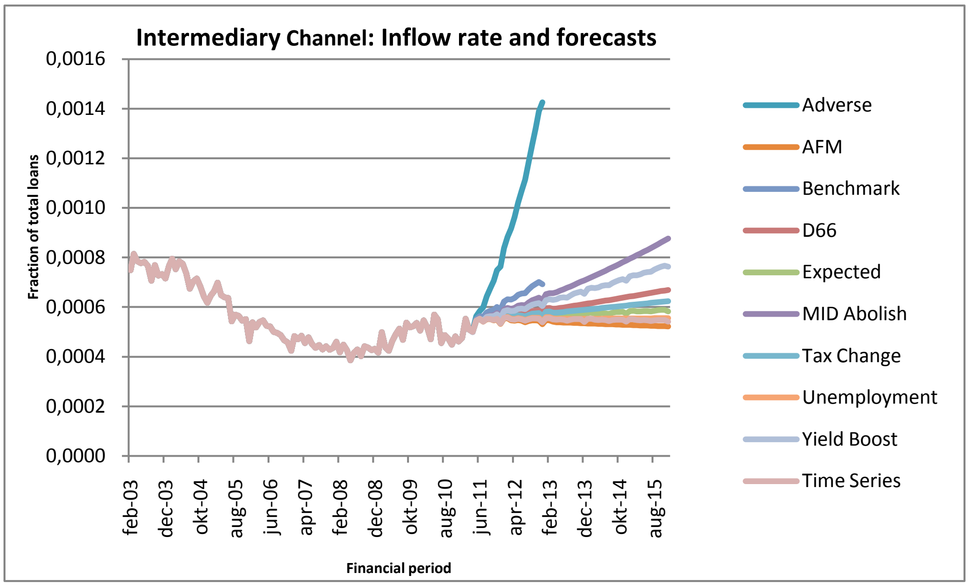


Figure A1: Intermediary Channel inflow rate time series and forecasts based on applied scenarios

Because of the large dependence on the house prices and yield rate, the ECB Adverse Stress-scenario is exploding. Another finding is the importance of the yield rate for the Inflow Rate, which explains the high results of the yield boost scenario. The difference with the ARIMA model of the time series itself is very small and the difference is mainly due to the volatility of the rate. The advantage to analyze scenario affections probably favors the macroeconomic input model.

$$\boldsymbol{ln(\frac{IR_{Intermediary}(t)}{1-IR_{Intermediary}(t)})} \\ = -4,55 - 0,85 * \boldsymbol{HousePrices}(t-3) * 10^{-5} - 0,06 * \boldsymbol{HousesSold}(t-24) * 10^{-4} + 0,03 \\ * \boldsymbol{InflowUnemploymentOther}(t-21) * 10^{-3} + 0,12 * \boldsymbol{MortgageInterestRate}(t) + 0,21 \\ * \boldsymbol{ln(\frac{IR_{Intermediary}(t-1)}{1-IR_{Intermediary}(t-1)})}$$

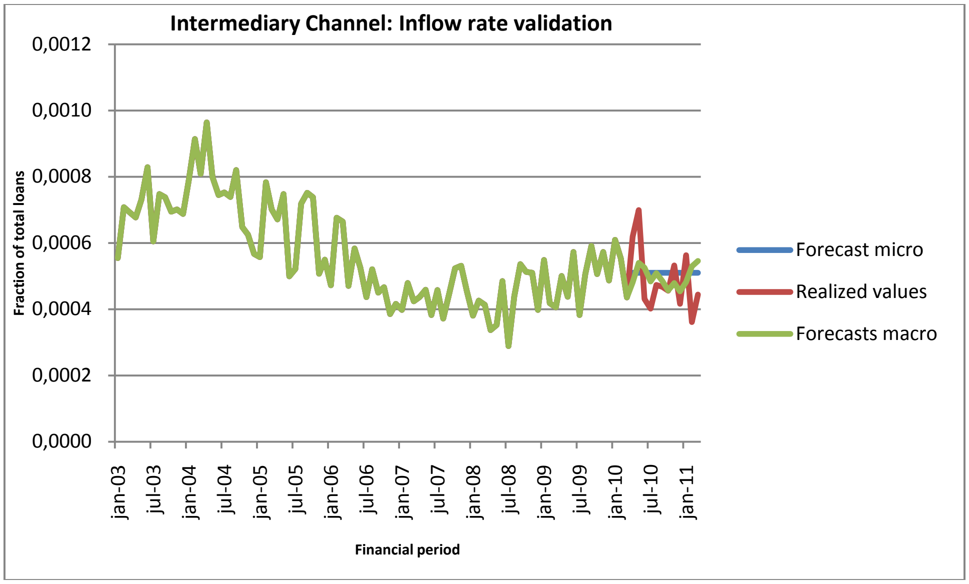


Figure A2: Intermediary Channel inflow rate comparison micro- and macroeconomic approach

A1.2 Recovery rate

The microeconomic approach is favored for the recovery rate, according to the T-test and emphasized by the errors. When building the model only on recovery rate history and therefore, the macroeconomic addition, based on houses sold and outflow of registered unemployment, is not a proper improvement in forecasting.

According to the model with macroeconomic factors involved, the unemployment factor relies heavily on the recovery rate. It could be assumed that increasing unemployment decreases the recoveries.

$$\boldsymbol{ln}\left(\frac{\boldsymbol{RR_{Intermediary}(t)}}{\boldsymbol{1-RR_{Intermediary}(t)}}\right) = -\boldsymbol{5,52} + \boldsymbol{0,13} * \boldsymbol{HousesSold(t-3)} * \boldsymbol{10^{-4}} + \boldsymbol{0,22} * \boldsymbol{OutflowUnemploymentTotal(t)} * \boldsymbol{10^{-4}} + \boldsymbol{0,41} * \boldsymbol{ln}\left(\frac{\boldsymbol{RR_{Intermediary}(t-1)}}{\boldsymbol{1-RR_{Intermediary}(t-1)}}\right)$$

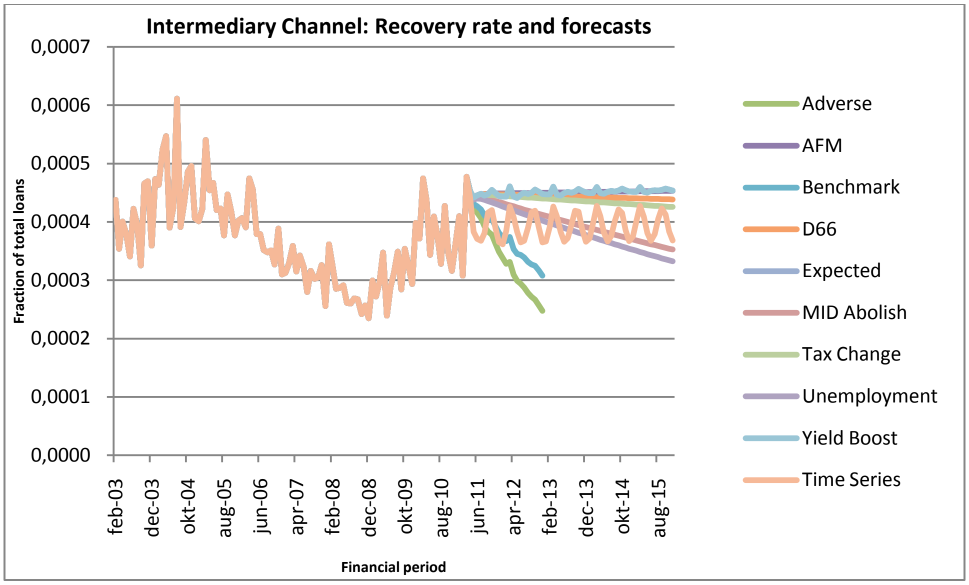


Figure A3: Intermediary Channel recovery rate time series and forecasts based on applied scenarios

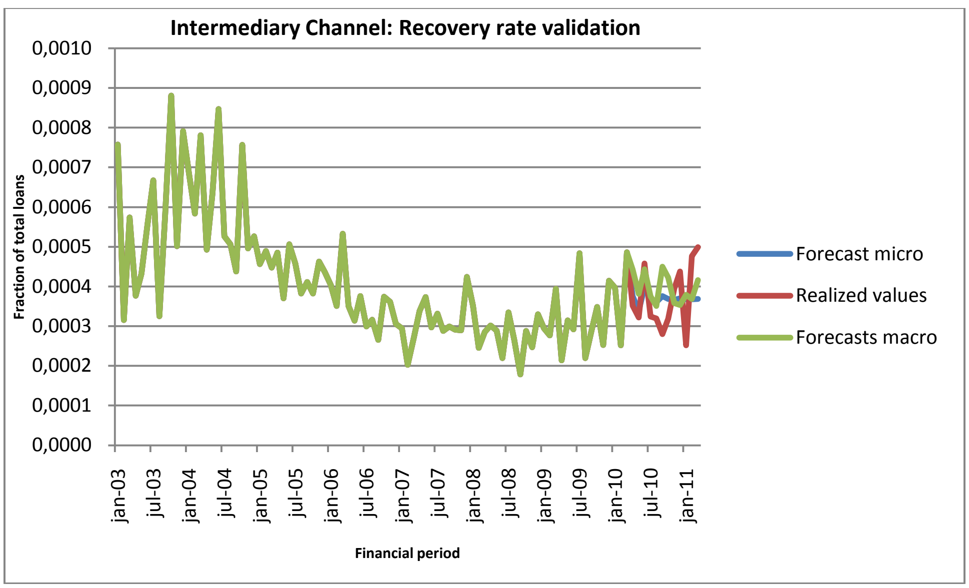


Figure A4: Intermediary Channel recovery rate comparison micro- and macroeconomic approach

A1.3 Foreclosure rate

Although, this rate is very hard to predict, a model is designed. The hardness is explicitly due to the strategy of collection and the actions exercised. The average error of the model depending on house prices, houses sold and the inflow of registered unemployment is about 40 percent, compared 31 percent of the ARIMA model of the Foreclosure Rate. None of the models is significantly better. In most of the scenarios, the recoveries and foreclosures develop in the opposite direction, which indicates a lower recovery percentage of the defaults.

Especially the scenario is with the mortgage interest deduction is totally eliminated, the foreclosure fraction rises dramatically. Nearly the same happens with increasing unemployment.

$$ln(\frac{FR_{Intermediary}(t)}{1-FR_{Intermediary}(t)}) = -7,00 - 0,63 * HousePrices(t-19) * 10^{-5} - 0,05 * HousesSold(t-22) * 10^{-4} + 0,21 * InflowUnemploymentOther(t-9) * 10^{-3} + 0,11 * ln(\frac{FR_{Intermediary}(t-1)}{1-FR_{Intermediary}(t-1)})$$

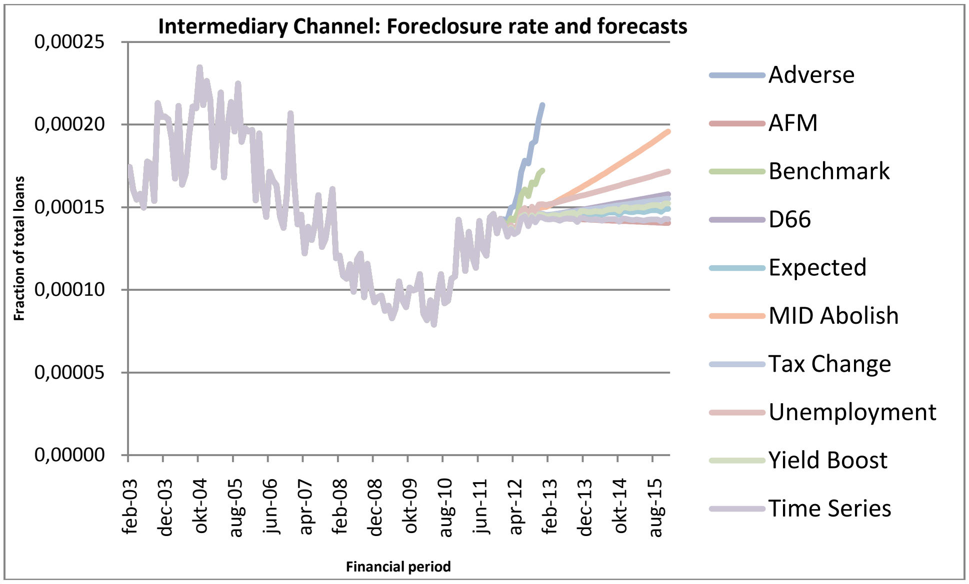


Figure A5: Intermediary Channel Foreclosure Rate time series and forecasts based on applied scenarios

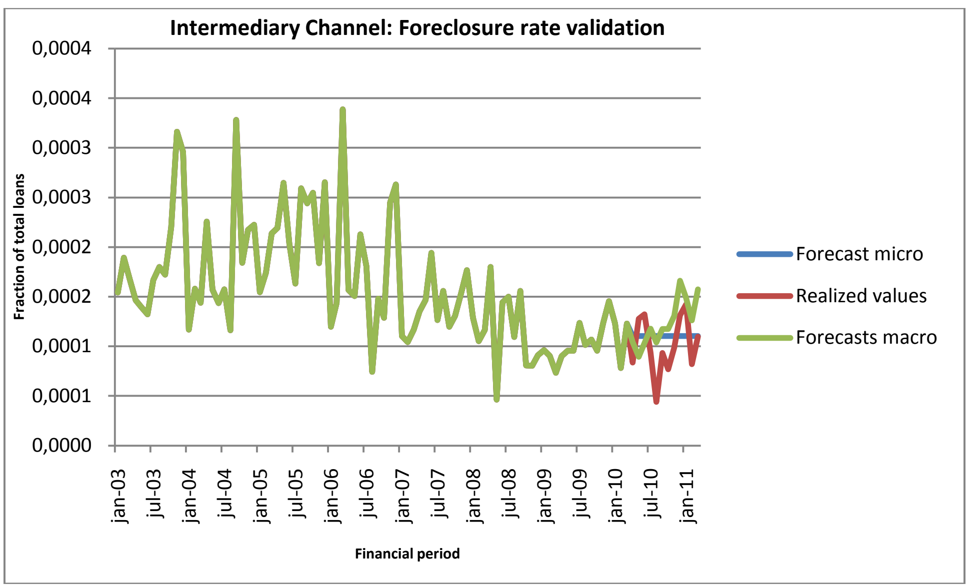


Figure A6: Intermediary Channel foreclosure rate comparison micro- and macroeconomic approach

A2: White Label

A2.1 Inflow rate

The inflow rate direction is clear, according to the scenarios: a stabilization is expected and extreme scenarios are increasing the rate. Also the history data of the inflow rate predicts a stable future. Both errors are around 24 percent, but the best model (microeconomic) is decided by the

T-test. According to the macroeconomic approach, one of the scenarios is explicitly out of the comfort range; the mortgage interest deduction abolish scenario shows a huge inflow of defaults.

$$\boldsymbol{ln}\left(\frac{\boldsymbol{IR_{White}(t)}}{\boldsymbol{1-IR_{White}(t)}}\right) = \boldsymbol{-7,27-0,77*HousesSold(t)*10^{-4}+\ 0,03 * InflowUnemploymentBusiness(t-1)*10^{-4}+\ 0,06*MortgageInterestRate(t)-0,10 * ln\left(\frac{IR_{White}(t-1)}{1-IR_{White}(t-1)}\right)}$$

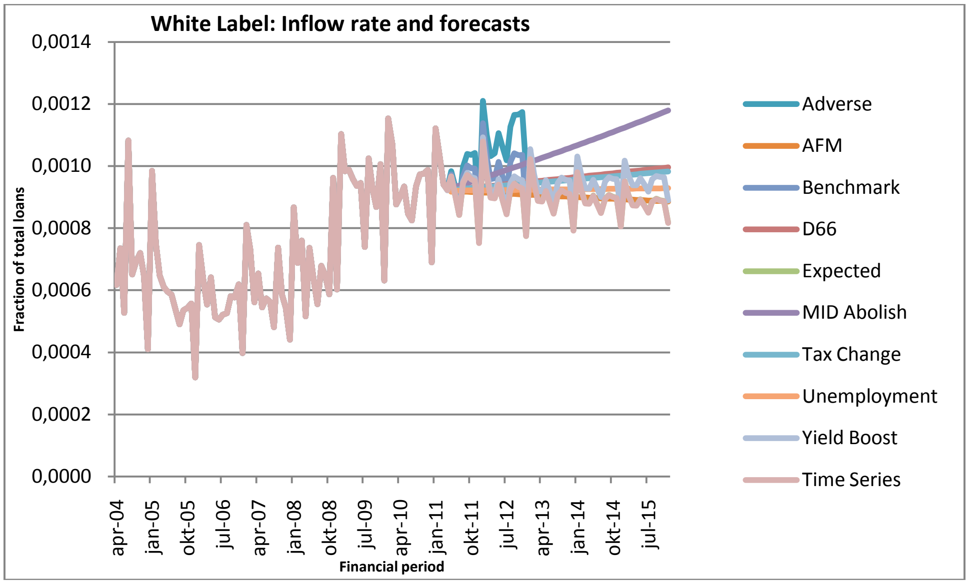


Figure A7: White Label inflow rate time series and forecasts based on applied scenarios

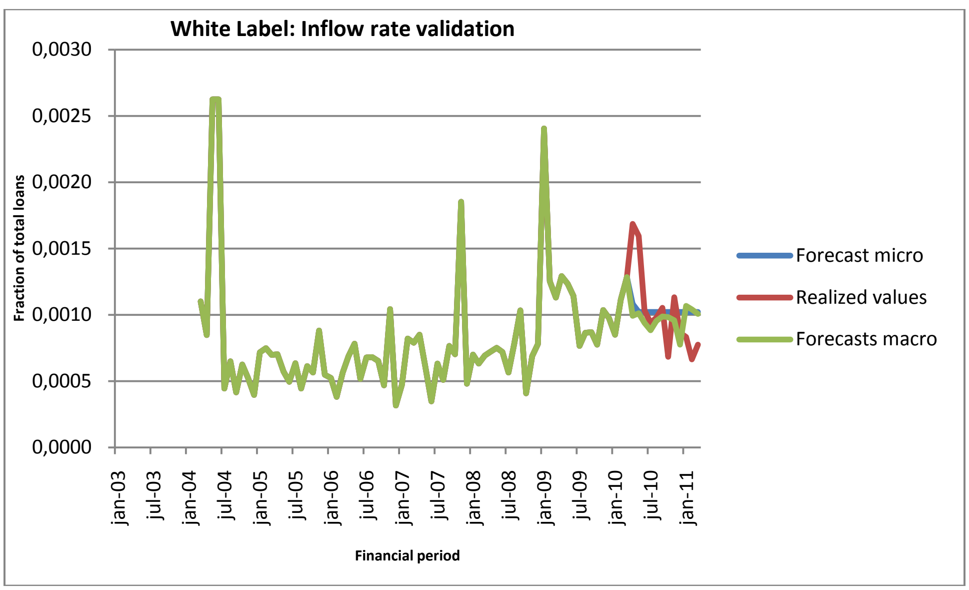


Figure A8: White Label inflow rate comparison micro- and macroeconomic approach

A2.2 Recovery rate

The link with macroeconomic factors is not important for predicting, according to the T-test and therefore the microeconomic approach is favored.

$$\ln\left(\frac{RR_{White}(t)}{1-RR_{White}(t)}\right) = -8,73 + 0,62 * HousePrices(t) * 10^{-4} + 0,08 * OutflowUnemploymentTotal(t-2) * 10^{-3} + 0,31 * BondYieldRate(t-24) + 0,22 * \ln\left(\frac{RR_{White}(t-1)}{1-RR_{White}(t-1)}\right)$$

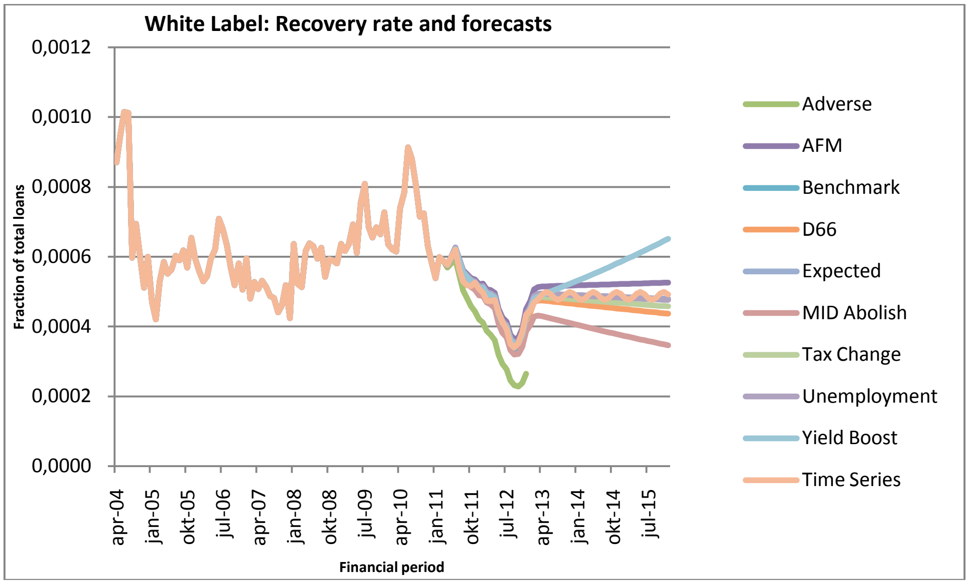


Figure A9: White Label recovery rate time series and forecasts based on applied scenarios

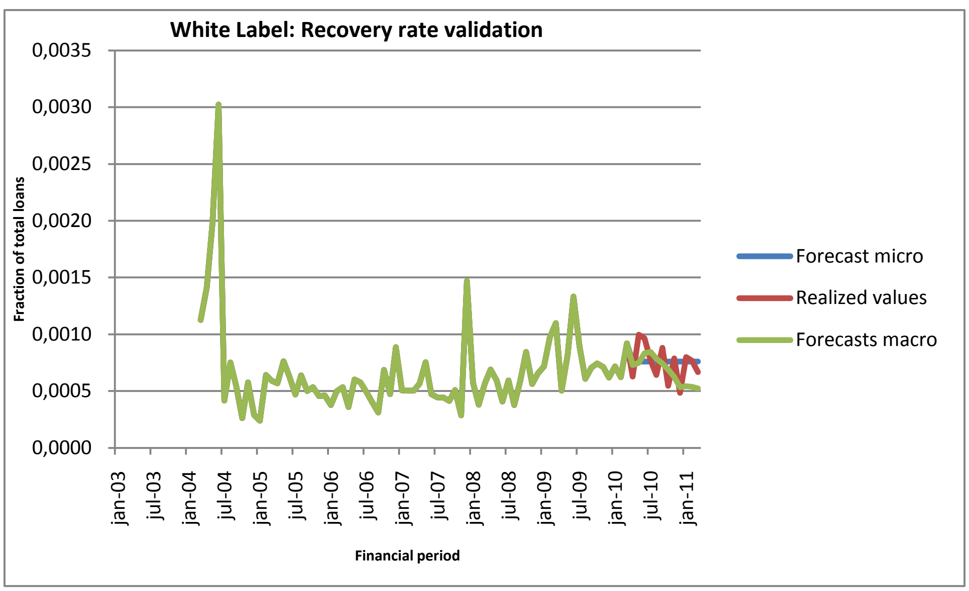


Figure A10: White Label recovery rate comparison micro- and macroeconomic approach

A2.3 Foreclosure rate

The foreclosure rate is, as noticed earlier, very hard to predict. The White Label prediction shows a very strong affection with the yield rate, but is absolutely not better than a prediction without external input. The microeconomic approach is even favored significantly.

The yield rate influences are of such a magnitude that the foreclosures are exploding when yield rates increase.

$$ln\left(\frac{FR_{White}(t)}{1-FR_{White}(t)}\right) = -15,19 - 0,75 * HousesSold(t-6) * 10^{-4} - 0,71 * OutflowUnemploymentTotal(t-19) * 10^{-3} + 1,59 * MortgageInterestRate(t) - 0,09 * ln\left(\frac{FR_{White}(t-1)}{1-FR_{White}(t-1)}\right)$$

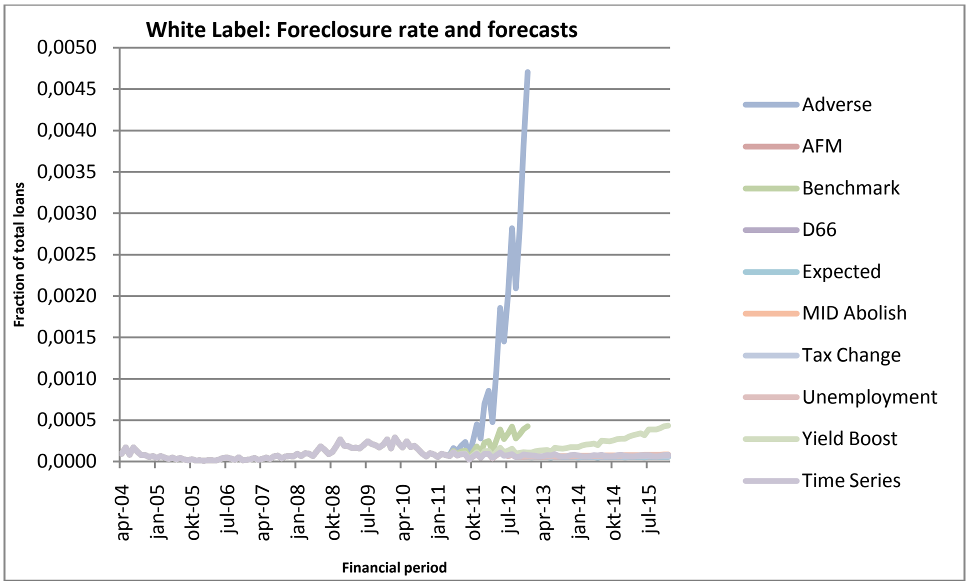


Figure A11: White Label foreclosure rate time series and forecasts based on applied scenarios

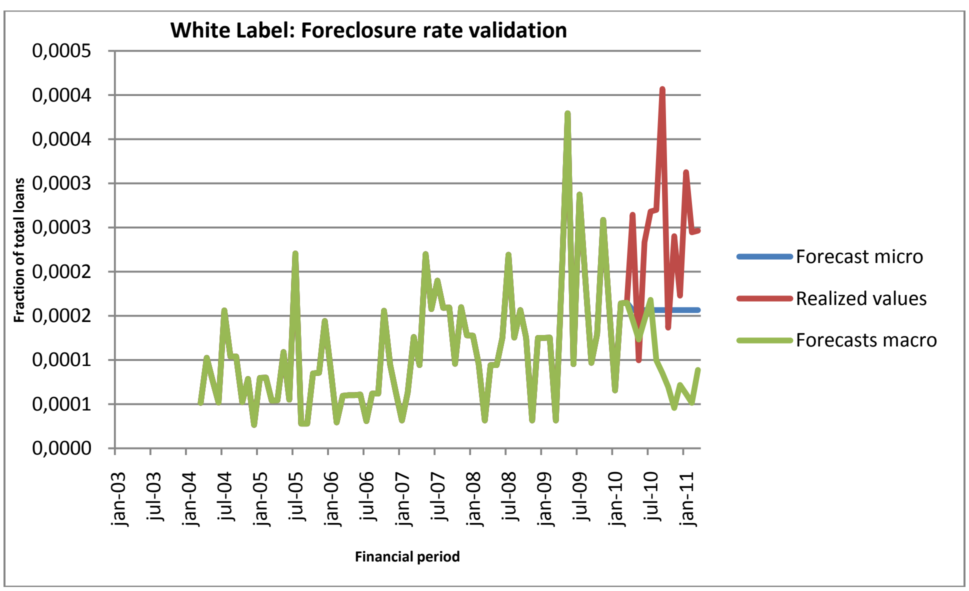


Figure A12: White Label foreclosure rate comparison micro- and macroeconomic approach

A3: Consolidated Portfolio

A3.1 Inflow rate

The inflow rate of the Consolidated Portfolio is based on the house prices, the housing market, interest and unemployment. There is no clear advantage for using the macroeconomic or microeconomic approach.

$$\ln\left(\frac{IR_{Cons}(t)}{1-IR_{Cons}(t)}\right) = -4,58 - 0,66 * HousePrices(t) * 10^{-5} - 0,01 * HousesSold(t-8) * 10^{-4} + 0,04 * InflowUnemploymentTotal(t-1) * 10^{-4} - 0,02 * OutflowUnemploymentOther(t-18) * 10^{-3} + 0,01 * MortgageInterestRate(t) + 0,14 * \ln\left(\frac{IR_{Cons}(t-1)}{1-IR_{Cons}(t-1)}\right)$$

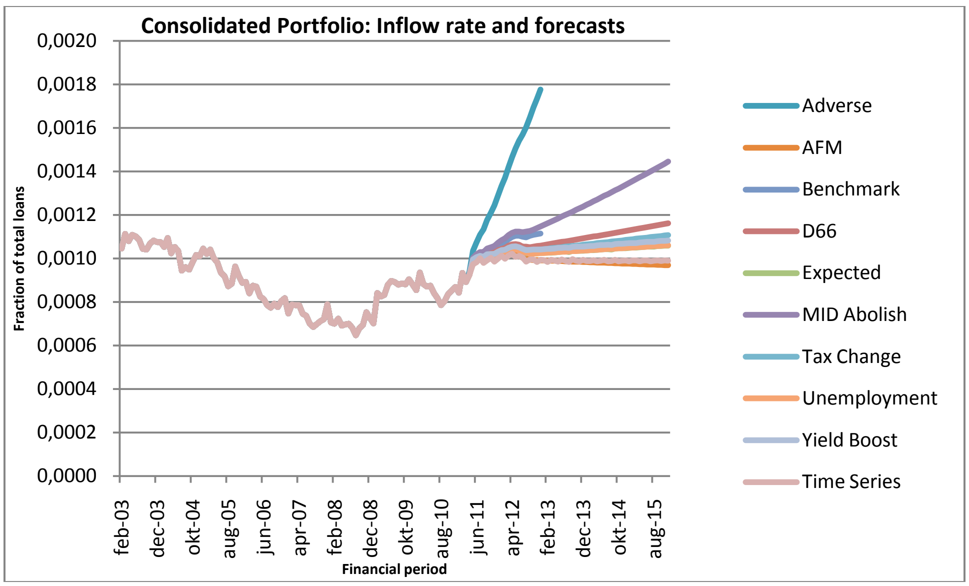


Figure A13: Consolidated inflow rate time series and forecasts based on applied scenarios

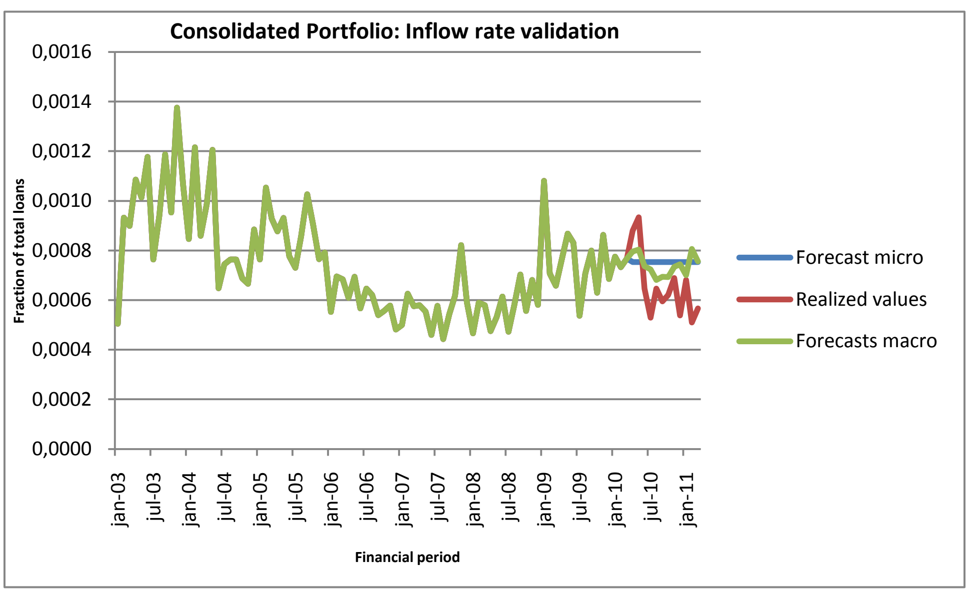


Figure A14: Consolidated Portfolio inflow rate comparison micro- and macroeconomic approach

A3.2 Recovery rate

The recovery rate of the Consolidated Portfolio can be predicted quite suitable with the microeconomic approach, which is favored by the T-test.

$$\boldsymbol{ln(\frac{RR_{Cons}(t)}{1-RR_{Cons}(t)}) = -3,90 + 0,15 * HousesSold(t-3) * 10^{-4} + 0,09 * OutflowUnemploymentTotal(t) * 10^{-4} + 0,55 * ln(\frac{RR_{Cons}(t-1)}{1-RR_{Cons}(t-1)})}$$

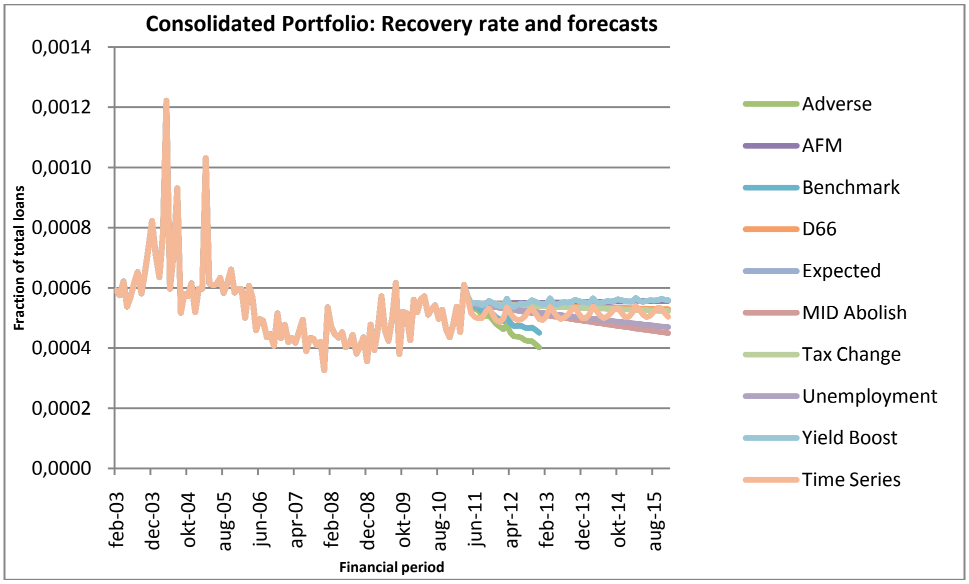


Figure A15: Consolidated recovery rate time series and forecasts based on applied scenarios

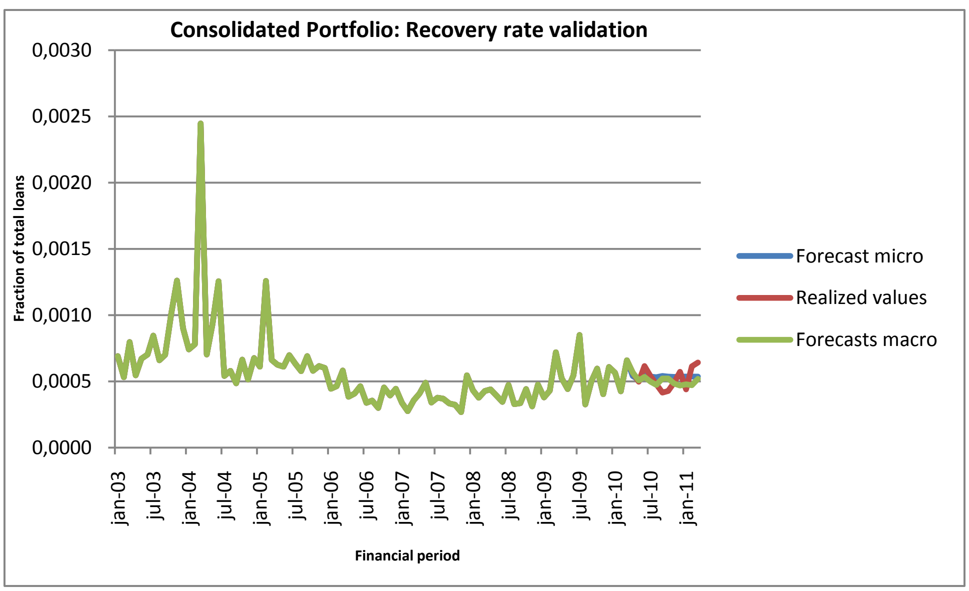


Figure A16: Consolidated Portfolio recovery rate comparison micro- and macroeconomic approach

A3.3 Foreclosure rate

The foreclosure rate can be predicted by the micro- as well as the macroeconomic approach. Although, the error of the microeconomic approach is smaller.

$$ln(\frac{FR_{Cons}(t)}{1 - FR_{Cons}(t)}) = -7,30 - 0,02 * HousesSold(t-23) * 10^{-4} + 0,25 * InflowUnemploymentOther(t - 8) * 10^{-3} + 0,03 * MortgageInterestRate\ (t) + 0,26 * ln(\frac{FR_{Cons}(t-1)}{1 - FR_{Cons}(t-1)})$$

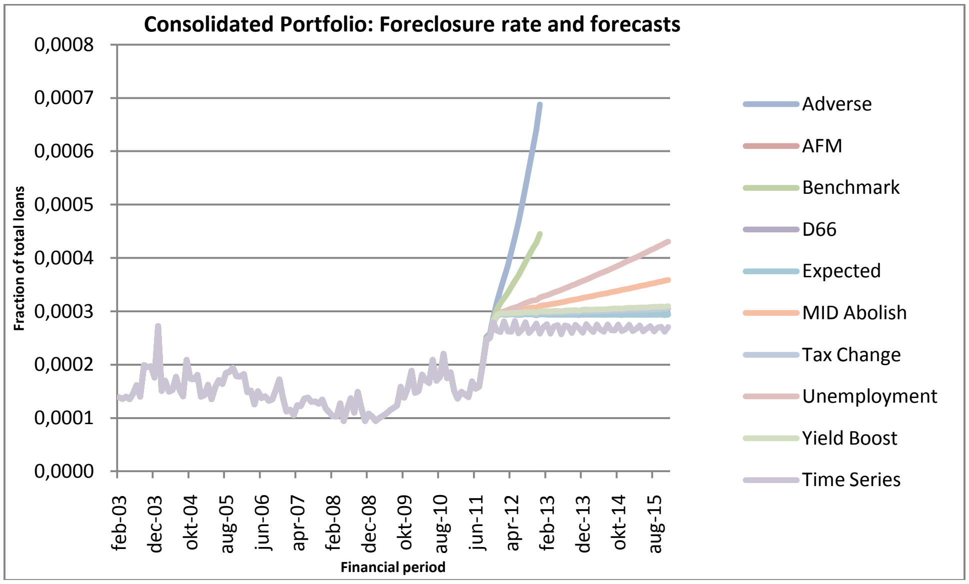


Figure A17: Consolidated foreclosure rate time series and forecasts based on applied scenarios

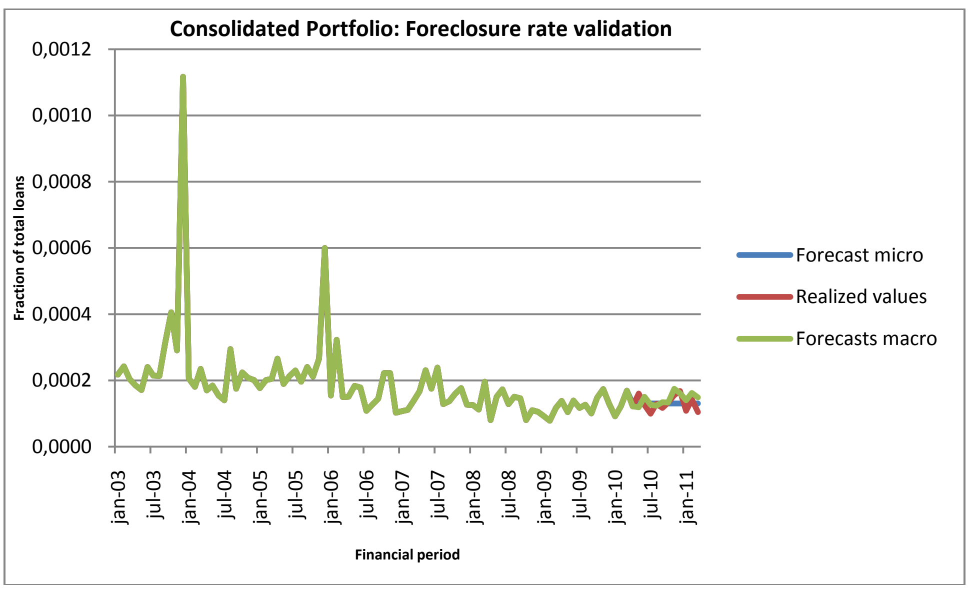


Figure A18: Consolidated Portfolio foreclosure rate comparison micro- and macroeconomic approach

A4. Overview and comparison model estimations

	Intermediary Channel			
	Default Rate	*Inflow Rate*	*Recovery Rate*	*Foreclosure Rate*
Favored model	Macroeconomic	None (micro)	Microeconomic	None (micro)
T-test (95%)	0,004519	0,204594	0,016974	0,077826
Sum of squared errors micro	1,49956E-06	1,10427E-07	6,76014E-08	9,98037E-09
Sum of squared errors macro	8,70302E-07	1,1111E-07	9,83895E-08	1,55942E-08
Average error micro	9%	18%	18%	31%
Average error macro	6%	17%	24%	40%
	White Label			
	Default Rate	*Inflow Rate*	*Recovery Rate*	*Foreclosure Rate*
Favored model	Microeconomic	Microeconomic	Microeconomic	Microeconomic
T-test (95%)	0,000041	0,022689	0,039725	0,000312
Sum of squared errors micro	1,60526E-05	1,0791E-06	2,78411E-07	1,56613E-07
Sum of squared errors macro	0,000112316	1,2154E-06	3,25201E-07	3,42914E-07
Average error micro	16%	23%	18%	38%
Average error macro	41%	24%	20%	58%
	Consolidated Portfolio			
	Default Rate	*Inflow Rate*	*Recovery Rate*	*Foreclosure Rate*
Favored model	Macroeconomic	None (macro)	Microeconomic	None (micro)
T-test (95%)	0,000003	0,297200	0,001869	0,063726
Sum of squared errors micro	6,35149E-07	3,13934E-07	6,41009E-08	5,56519E-09
Sum of squared errors macro	3,38656E-07	2,54688E-07	7,69931E-08	1,8228E-06
Average error micro	3%	25%	12%	14%
Average error macro	2%	21%	12%	73%

Table A1: Comparison microeconomic and macroeconomic approach

Appendix B: Credit losses

B1 Credit losses and forecasts

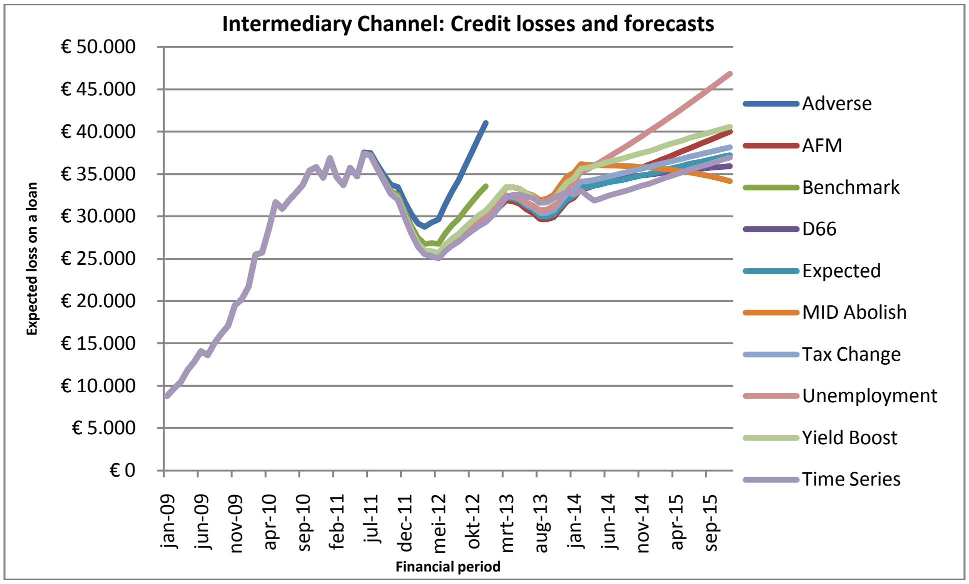


Figure B1: Intermediary Channel credit losses and forecasts

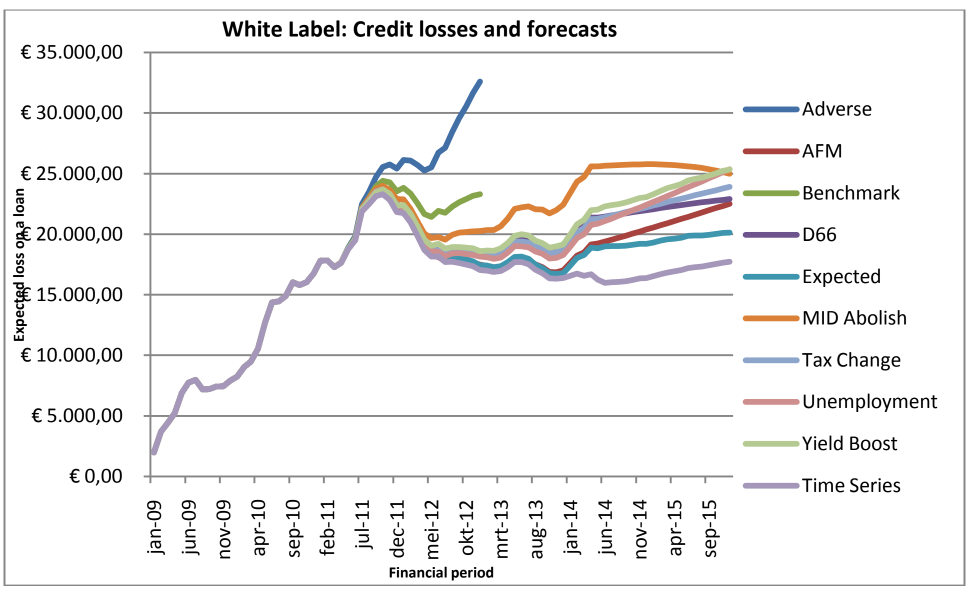


Figure B2: White Label credit losses and forecasts

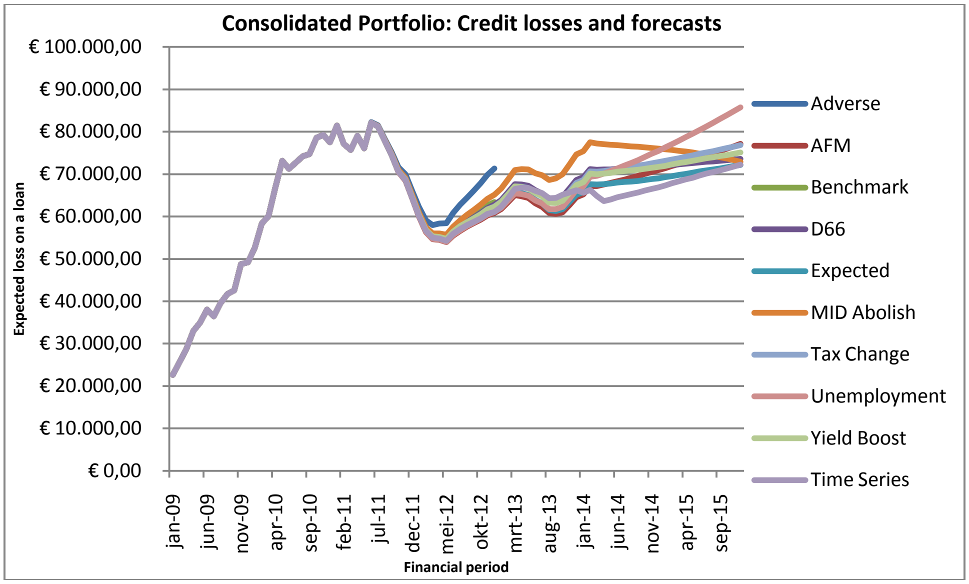


Figure B3: Consolidated Portfolio credit losses and forecasts

B2 Credit losses: summarizing table for monthly data

	Adverse	*AFM*	*Benchmark*	*D66*	*Expected*	*MID Abolish*	*Tax Change*	*Unemployment*	*Yield Boost*	*Time Series*
dec-09	€ 20.195	€ 20.195	€ 20.195	€ 20.195	€ 20.195	€ 20.195	€ 20.195	€ 20.195	€ 20.195	€ 20.195
dec-10	€ 34.588	€ 34.588	€ 34.588	€ 34.588	€ 34.588	€ 34.588	€ 34.588	€ 34.588	€ 34.588	€ 34.588
dec-11	€ 31.791	€ 30.041	€ 30.606	€ 30.110	€ 30.076	€ 30.193	€ 30.092	€ 30.074	€ 30.192	€ 29.914
dec-12	€ 41.029	€ 29.604	€ 33.558	€ 30.015	€ 29.766	€ 30.641	€ 29.938	€ 29.989	€ 30.687	€ 29.330
dec-13		€ 31.783		€ 32.736	€ 32.015	€ 34.738	€ 32.676	€ 33.435	€ 34.012	€ 32.843
dec-14		€ 36.066		€ 34.921	€ 34.955	€ 35.775	€ 35.832	€ 39.986	€ 37.673	€ 33.830
dec-15		€ 40.002		€ 35.933	€ 37.197	€ 34.160	€ 38.146	€ 46.834	€ 40.556	€ 36.943
	Adverse	*AFM*	*Benchmark*	*D66*	*Expected*	*MID Abolish*	*Tax Change*	*Unemployment*	*Yield Boost*	*Time Series*
dec-09	€ 7.909	€ 7.909	€ 7.909	€ 7.909	€ 7.909	€ 7.909	€ 7.909	€ 7.909	€ 7.909	€ 7.909
dec-10	€ 16.756	€ 16.756	€ 16.756	€ 16.756	€ 16.756	€ 16.756	€ 16.756	€ 16.756	€ 16.756	€ 16.756
dec-11	€ 25.404	€ 22.162	€ 23.538	€ 22.447	€ 22.198	€ 22.908	€ 22.422	€ 22.351	€ 22.377	€ 21.816
dec-12	€ 32.594	€ 17.482	€ 23.287	€ 18.509	€ 17.445	€ 20.264	€ 18.416	€ 18.157	€ 18.599	€ 17.037
dec-13		€ 17.016		€ 18.936	€ 16.784	€ 22.437	€ 18.758	€ 18.268	€ 19.157	€ 16.380
dec-14		€ 20.381		€ 21.985	€ 19.193	€ 25.775	€ 22.230	€ 22.435	€ 23.067	€ 16.377
dec-15		€ 22.511		€ 22.899	€ 20.126	€ 25.003	€ 23.897	€ 25.372	€ 25.335	€ 17.713
	Adverse	*AFM*	*Benchmark*	*D66*	*Expected*	*MID Abolish*	*Tax Change*	*Unemployment*	*Yield Boost*	*Time Series*
dec-09	€ 49.199	€ 49.199	€ 49.199	€ 49.199	€ 49.199	€ 49.199	€ 49.199	€ 49.199	€ 49.199	€ 49.199
dec-10	€ 77.534	€ 77.534	€ 77.534	€ 77.534	€ 77.534	€ 77.534	€ 77.534	€ 77.534	€ 77.534	€ 77.534
dec-11	€ 66.077	€ 64.327	€ 64.715	€ 64.684	€ 64.405	€ 65.113	€ 64.611	€ 64.376	€ 64.510	€ 64.360
dec-12	€ 71.350	€ 60.819	€ 63.395	€ 62.760	€ 61.309	€ 65.163	€ 62.361	€ 61.088	€ 62.252	€ 61.097
dec-13		€ 64.543		€ 68.618	€ 65.140	€ 74.615	€ 67.929	€ 66.258	€ 67.385	€ 66.216
dec-14		€ 70.933		€ 71.894	€ 68.997	€ 76.019	€ 72.840	€ 75.515	€ 71.681	€ 66.884
dec-15		€ 77.103		€ 73.560	€ 72.120	€ 72.902	€ 76.813	€ 85.741	€ 75.068	€ 72.200

Table B1: Expected total credit losses on period basis

B3 Credit losses: estimate of yearly loss

	Adverse	*AFM*	*Benchmark*	*D66*	*Expected*	*MID Abolish*	*Tax Change*	*Unemployment*	*Yield Boost*	*Time Series*
dec-09	€ 169.370	€ 169.370	€ 169.370	€ 169.370	€ 169.370	€ 169.370	€ 169.370	€ 169.370	€ 169.370	€ 169.370
dec-10	€ 368.210	€ 368.210	€ 368.210	€ 368.210	€ 368.210	€ 368.210	€ 368.210	€ 368.210	€ 368.210	€ 368.210
dec-11	€ 420.695	€ 414.806	€ 416.587	€ 415.058	€ 414.932	€ 415.339	€ 414.987	€ 414.887	€ 415.223	€ 414.474
dec-12	€ 399.839	€ 326.621	€ 352.072	€ 329.139	€ 327.791	€ 332.834	€ 328.615	€ 328.906	€ 333.679	€ 323.933
dec-13		€ 369.569		€ 378.390	€ 373.497	€ 394.100	€ 377.243	€ 380.649	€ 391.409	€ 383.983
dec-14		€ 412.288		€ 411.883	€ 408.016	€ 430.834	€ 416.855	€ 446.374	€ 437.176	€ 393.902
dec-15		€ 458.209		€ 426.050	€ 434.782	€ 420.629	€ 445.180	€ 523.505	€ 471.631	€ 426.910
	Adverse	*AFM*	*Benchmark*	*D66*	*Expected*	*MID Abolish*	*Tax Change*	*Unemployment*	*Yield Boost*	*Time Series*
dec-09	€ 75.098	€ 75.098	€ 75.098	€ 75.098	€ 75.098	€ 75.098	€ 75.098	€ 75.098	€ 75.098	€ 75.098
dec-10	€ 158.292	€ 158.292	€ 158.292	€ 158.292	€ 158.292	€ 158.292	€ 158.292	€ 158.292	€ 158.292	€ 158.292
dec-11	€ 256.729	€ 245.800	€ 250.487	€ 246.781	€ 246.108	€ 248.358	€ 246.693	€ 246.450	€ 246.524	€ 244.206
dec-12	€ 335.379	€ 224.470	€ 270.748	€ 232.447	€ 227.405	€ 245.797	€ 231.728	€ 229.732	€ 235.519	€ 222.537
dec-13		€ 208.867		€ 227.096	€ 208.899	€ 259.455	€ 225.419	€ 220.798	€ 230.940	€ 203.924
dec-14		€ 231.834		€ 255.489	€ 225.033	€ 303.747	€ 255.340	€ 252.521	€ 264.600	€ 195.926
dec-15		€ 258.337		€ 270.046	€ 237.539	€ 305.963	€ 277.683	€ 288.094	€ 292.863	€ 206.135
	Adverse	*AFM*	*Benchmark*	*D66*	*Expected*	*MID Abolish*	*Tax Change*	*Unemployment*	*Yield Boost*	*Time Series*
dec-09	€ 441.243	€ 441.243	€ 441.243	€ 441.243	€ 441.243	€ 441.243	€ 441.243	€ 441.243	€ 441.243	€ 441.243
dec-10	€ 839.145	€ 839.145	€ 839.145	€ 839.145	€ 839.145	€ 839.145	€ 839.145	€ 839.145	€ 839.145	€ 839.145
dec-11	€ 914.401	€ 908.300	€ 909.621	€ 909.558	€ 908.623	€ 911.068	€ 909.302	€ 908.474	€ 908.900	€ 908.465
dec-12	€ 759.161	€ 687.774	€ 706.831	€ 700.855	€ 693.376	€ 716.864	€ 698.173	€ 689.575	€ 699.229	€ 691.767
dec-13		€ 754.314		€ 791.974	€ 762.911	€ 842.842	€ 784.794	€ 764.555	€ 782.494	€ 783.869
dec-14		€ 822.222		€ 854.452	€ 815.554	€ 919.208	€ 855.504	€ 860.428	€ 846.137	€ 784.268
dec-15		€ 890.425		€ 873.610	€ 848.602	€ 894.951	€ 899.423	€ 970.621	€ 882.617	€ 837.887

Table B2: Expected total credit losses on yearly basis