COMENIUS UNIVERSITY, BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

METHODOLOGY DESIGN FOR STRESS TESTING OF RISK PARAMETER PD

DIMPLOMA THESIS

2016 **Bc. Anna ROSINOVÁ**

COMENIUS UNIVERSITY, BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

METHODOLOGY DESIGN FOR STRESS TESTING OF RISK PARAMETER PD

DIMPLOMA THESIS

Study Programme:	Economic and Financial Mathematics
Field of Study:	9.1.9. Applied Mathematics
Department:	Department of Applied Mathematics and Statistics
Supervisor:	doc. Mgr. Radoslav Harman, PhD.

Bratislava, 2016

Bc. Anna ROSINOVÁ

Acknowledgements:

This way I would like to thank to my supervisor doc. Mgr. Radoslav Harman, PhD. for his endless patience, good questions that helped me to find the answers I needed and help with time management. I would also like to thank to my colleagues Mgr. Eva Karvajová for her trust in my abilities and good advices and Mgr. Gejza Wimmer, PhD. for his optimistic realism and bright ideas when everything seemed to be lost. Lastly, I would like to thank to my family for their patience and support.

96874133

Comenius University in Bratislava
Faculty of Mathematics, Physics and Informatics

THESIS ASSIGNMENT

Name and Surname:	Bc. Anna Rosinová
Study programme:	Mathematical Economics, Finance and Modelling (Single degree study, master II. deg., full time form)
Field of Study:	Applied Mathematics
Type of Thesis:	Diploma Thesis
Language of Thesis:	English
Secondary language:	Slovak

Title: Methodology design for stress testing of risk parameter PD

Aim: Introduce legislative requirements for stress testing of risk parameters and make an overview of stress testing in Eurozone. Design a methodology for stress testing of risk parameter PD based on these requirements and program the methodology in statistical softvare SAS.

Supervisor:	doc. Mgr. Radoslav Harman, PhD.
Department:	FMFI.KAMŠ - Department of Applied Mathematics and Statistics
Head of department:	prof. RNDr. Daniel Ševčovič, CSc.
Assigned:	14.02.2015
Approved:	16.02.2015

prof. RNDr. Daniel Ševčovič, CSc.
Guarantor of Study Programme

Student

Supervisor

96874133

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Anna Rosinová
Študijný program: ekonomicko-finančná matematika a modelovanie (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: aplikovaná matematika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: anglický
Sekundárny jazyk: slovenský

Názov: Methodology design for stress testing of risk parameter PD
Návrh metodiky na stresové testovanie rizikového parametra PD

Cieľ: Introduce legislative requirements for stress testing of risk parameters and make an overview of stress testing in Eurozone. Design a methodology for stress testing of risk parameter PD based on these requirements and program the methodology in statistical softvare SAS.

Vedúci: doc. Mgr. Radoslav Harman, PhD.
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry: prof. RNDr. Daniel Ševčovič, CSc.

Dátum zadania: 14.02.2015

Dátum schválenia: 16.02.2015

prof. RNDr. Daniel Ševčovič, CSc.
garant študijného programu

študent

vedúci práce

Abstrakt v štátnom jazyku

ROSINOVÁ, Anna: Návrh metodiky na stresové testovanie rizikového parametra PD [Diplomová práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky, školiteľ: doc. Mgr. Radoslav Harman, PhD., Bratislava, 2016, 58 s.

V práci uvádzame prehľad základných legislatívnych požiadaviek v oblasti riadenia kreditného rizika a stresové testovanie uvádzame ako integrálnu súčasť tohto procesu. Súčasťou navrhnutia metodiky je aj stručný prehľad prístupov, ktoré už v minulosti boli použité na stresové testovanie. Na odhad vplyvu makroekonomických premenných na portfólio hypoték sme použili tzv. CPH model. Postupne uvádzame matematické východiská použité v modely, metódy použité na odhad modelu na reálnych dátach ako aj viacero typov rezíduí, ktoré boli použité na kontrolu predpokladov modelu. Analýzu sme vykonali na bežne používaných makroekonomických premenných a naviac sme ako vstupný parameter modelu použili aj rating klienta ako ukazovateľ jeho individuálnych charakteristík. Model, ktorý navrhujeme ako vhodný na stresové testovanie rizikového parametra PD je založený na reálnom raste HDP. Nakoniec uvádzame prehľad dopadov na výpočet kapitálovej primeranosti pre rôzny vývoj HDP.

Kľúčové slová: stresové testovanie, CPH model, pravdepodobnosť zlyhania , reálny rast HDP

Abstract

ROSINOVÁ, Anna: Methodology Design for Stress Testing of Risk Parameter PD [Diploma thesis], Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statics, Supervisor: doc. Mgr. Radoslav Harman, PhD., Bratislava, 2016, 58 p.

In this thesis we briefly introduce legislative requirements on credit risk management in bank institutions and introduce stress testing as the integral part of risk assessment. We provide a short summary of approaches already used for stress testing. Cox Proportional Hazard model is used to estimate the impact of macroeconomic variables on a mortgage loan portfolio. We set up mathematical background of the model, estimation techniques as well as residuals used for the model adequacy check. We analyze commonly used economic drivers in order to quantify the impact on the portfolio. Moreover, the Rating of the client is used in order to treat for the individual impact. The final model we propose for the stress testing of PD parameter is based on real growth of GDP. Finally, we provide overview of impacts on capital requirement under the different states of economy.

Keywords: Stress testing, Cox Proportional Hazard Model, Probability of Default (PD), Real Growth of GDP

Contents

List of Abbreviations

BCBS – Basel Committee on Banking Supervision

CCF – Credit Conversion Factor

CPH – Cox Proportional Hazard Model

DR – Default Rate

EAD – Exposure at Default

EBA – European Banking Authority

EL – Expected Loss

HR – Hazard Ratio

IRB – Internal Rating Based

LGD – Loss Given at Default

OLS – Ordinary Least Squares

PD – Probability of Default

RW – Risk Weights

RWA – Risk Weighted Assets

SL – Specialized Lending

UL – Unexpected Loss

Introduction

After the recent crisis (started in 2008) many concerns have arisen about the stability of the banking system. As reaction European authorities introduced stress testing as a way how to disclose shortcomings in bank's capital and preparedness for next macroeconomic turbulences and as a way how to bring back trust into the banking system. Regular stress testing became a part of risk management which requires development of robust methodology. The aim of this thesis is to make short overview of legislative framework for risk management in banks with focus on risk parameter PD, define common techniques for the stress testing and develop a model which could be used to model macroeconomic impact on the bank's portfolio.

Even though this work does not bring a new methodology, the estimation of the model on real data is always unique and very challenging process which requires a complex comprehension of the context as well as mathematical framework. Several aspects have to be considered and there is no unique right solution.

This work covers the development of a stress testing model step by step. In the first chapter we set up legislative framework and define basic terms of risk management. The focus is on basic concepts of credit risk. Next chapter introduces common methodologies used for stress testing. Theory part ends up with third chapter where we choose and describe Cox proportional hazard model which is afterwards used for estimation of the model on the mortgage loans portfolio.

In the second part of the work we describe data preparation, macroeconomic variables which enter the model and we propose the model for stress testing. We provide extensive check of the adequacy of the proposed model by different types of residuals commonly used for CPH model as Cox-Snell residuals, Martingale residuals, Deviance residuals and lastly Schoenfeld residuals. In the last chapter we show how the model is used for calculation of regulatory capital requirement under the different macroeconomic conditions.

1. Legislative Requirements

Banks as any other enterprises face several risks related to the activities they provide. As the banks are crucial institutions in economy it is very important to manage all these risks and anticipate situations which may lead to crisis again. Standards for effective identification, measurement and assessment of risks are anchored in documents issued by EBA and Basel Committee on Banking Supervision. National regulators are required to supervise commercial banks under their jurisdiction and support activities of European authorities in order to preserve stable markets. Definitions in this chapter are based mainly on Basel II and Basel III accord and European regulation no 575/2013.

As we have already mentioned bank is exposed to different risks in its operations. According to the source of the risk we distinguish between three main types:

- Credit risk
  Credit risk is most simply defined as the potential that a bank borrower or counterparty will fail to meet its obligations in accordance with agreed terms [15].

- Market risk
  Market risk is defined as the risk of losses in on and off-balance-sheet positions arising from movements in market prices. The risks subject to this requirement are:
  - The risk pertaining to interest rate related instruments and equities in the trading book
  - Foreign exchange risk and commodities risk throughout the bank.

- Operational risk
  Operational risk is the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events. It includes legal risks such as exposures to fines, penalties, punitive damages resulting from supervisory actions, as well as private settlements. However, strategic and reputational risk is excluded [3].

1.1. Credit Risk

In order to ensure that a bank does not become insolvent due to failures of its clients it holds certain amount of capital as a reserve. The amount of this capital requirement is under the supervision of regulator. According to Basel II accord to determine the capital requirements banks are permitted to choose between two methodologies to calculate their capital requirements the standardised approach and internal rating based approach (IRB). Standardised Approach is predominantly for less sophisticated banks however, they are expected to evolve to IRB approaches.

A bank tries to estimate the probability of default of the clients and the amount of the possible losses. It is a complex process that results in adequacy of bank’s capital and loan loss reserves. It requires a robust system and it underlies to strict regulatory standards. When measuring the riskiness of the portfolio it is important to reflect the heterogeneity of the exposures in terms of risk characteristics they underlie. The main asset classes are as following: corporate, sovereign, bank, retail and equity. We will focus on corporate and retail asset class.

1.1.1. Retail

According to §213 in [3] under the retail asset class are classified exposures where nature of the borrower is an individual. This encounters exposures to individuals with exposure less than € 1 million, residential mortgage loans if the individual is an owner-occupier of the property and loans extended to small business if the exposure of the banking group is less than €1 million. Under the retail asset class banks are required to identify three sub-classes of exposures:

- Exposures secured by residential properties which encounters exposures to individuals, residential mortgage loans and loans extended to small businesses or small business loans if the exposure is less than € 1 million.
- Qualifying revolving retail exposures – exposures to individuals, which are unsecured, uncommitted and where customers are permitted to decide to borrow and repay its outstanding balances up to the settled limit. Maximum limit is €100,000.
- Other exposures

1.1.2. Corporate

As defined in §218 in [3] corporate exposure is in general a debt obligation of a corporation, partnership or proprietorship. Banks are permitted to distinguish separately exposures to small- and medium-sized entities (SME) (§273). These are defined as firms where the reported sales of the whole consolidated group are less than €50 million. For exposures to SME borrowers a firm-size adjustment is made to corporate risk weight formula. (i.e. 0.04 x (1-(S-5)/45), where S is total annual sales in millions of euros and for exposures where the sales are less than €5 million, S is equal to 5).

Within the corporate asset class five subclasses may be identified. Distinction of these five subclasses is important especially when using supervisory risk weights otherwise more detailed segmentation may help to identify homogenous subgroups. Banks using foundation IRB approach must use supervisory parameters for LGD and EAD corporate exposures.

1.2. The Standardised Approach

Banks under the standardized approach use risk weights defined by regulator. Each exposure is assigned a risk weight according to strict rules. The grading scheme is predefined for each asset class. An example for corporate asset class is in Table 1.1. We focus on claims on corporate and retail asset classes. Principle of weighting claims on corporates is as following – each rating falls into prescribed category as illustrated in table. Unrated exposures are weighted at 100% (this may be changed if the supervisory authorities judge that the higher risk weight is warrant).

Tab. 1.1: Risk weights under the Standardized Approach

Credit assessment	AAA to AA-	A+ to A-	BBB+ to BB-	Below BB-	Unrated
Risk weight	20%	50%	100%	150%	100%

As the significant part of retail portfolio constitutes from private individuals and those do not have external ratings a different approach is used to set the risk weight. Claims qualifying as retail portfolio according to definition above are risk-weighted at 75%. Exception are past due loans (unsecured portion of any loan that is past due for more than 90 days).

Weaknesses of this approach were revealed and published in a BCBS's consultative document in December 2014 [17]. The most important are over-reliance on external credit ratings, lack of granularity and risk sensitivity, out-of-date calibrations, lack of comparability with risk measurement under the IRB approach and excessive complexity and lack of clarity which results into international differences. The document was opened for discussion until the end of March 2015. The main purpose of upcoming changes is to create an approach which would reflect the inherent riskiness of exposures and would be a suitable alternative and complement to internal models. The other aim is to set up a floor for IRB approach so as to reduce variability in RWA and reduce the risk the capital requirements would be too low.

The crucial change is in corporate and retail asset classes where the proposal is to alter the external ratings with key risk drivers. This risk drivers were selected on the basis that they should be simple, intuitive, readily available and capable of explaining risk consistently across jurisdictions. For example for corporate exposures instead of using rating the risk-weight is in range from 60% - 300% on the basis of two risk drivers: revenue and leverage (these are calibrations from the consultative document which are just preliminary). Some of the comments on proposed changes criticise the complete removal of rating as a determinant of risk-weight [13], the rationale is that the ratings draw on much wider range of factors than the Basel Committee is proposing.

1.3. Internal Ratings Based Approach

Banks with rich history of data and robust processes for risk control are expected to assess the riskiness of the counterparty by themselves. This risk measurement is based mainly on estimation of three risk parameters – probability of default (PD), Loss given default (LGD) and credit conversion factor (CCF). To use its own estimates a bank must meet certain minimum conditions and disclosure requirements in order to receive supervisory approval for such an activity. There are two levels of IRB approach that may be implemented. The first one is foundation IRB approach when only PD parameter is estimated by bank and supervisory parameters are used for other parameters. When a bank meets more rigorous conditions it is expected to use the advanced IRB approach and the bank is responsible for assessing all risk parameters.

The core of the IRB approach is to measure the expected (EL) and unexpected (UL) losses. Expected and unexpected losses are treated separately. Capital requirements for securitization of UL are calculated by risk-weight functions which are developed by bank. The estimated risk parameters are used as inputs of these risk-weight functions. On the other hand formulas for calculation of EL amount are determined by the supervisor (detailed definitions can be found in Basel II, section III.G).

To sum up three key elements are important when using IRB approach (foundation or advanced) [3]:

- Risk parameters estimated by bank or provided by supervisor
- Risk-weight functions
- These functions transform risk parameters into risk-weight assets and capital requirements
- Minimum requirements
- Standards which must be met in order to use IRB approach

When discussing adoption of IRB approach the key idea is that it should be done systematically and consistently over time. Once a bank adopts an IRB approach for any part of its portfolio it is expected to extend it across entire banking group. Similarly for asset classes once the IRB approach is used for one asset class it should be implemented for other asset classes or business lines as well. The whole implementation process is under the regulatory supervision and the bank must create a detailed implementation plan.

The advantage of using IRB approach is illustrated in the Figure 1.1. We compare the prescribed risk-weights under the standardised approach with the risk-weights (RW) calculated when using IRB approach. As we observe IRB approach asses the riskiness of the counterparty more precisely therefore when conducted properly it may lead to more appropriate allocation of the capital.

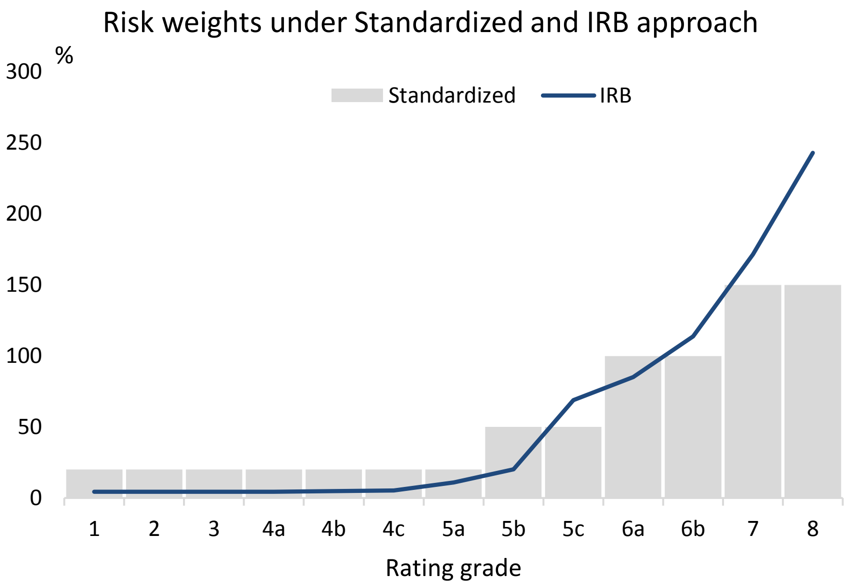


Fig. 1.1.: Comparison of RW under the Standardized and IRB approach.

1.3.1. Requirements for the IRB Approach [16]

So far we have discussed general characteristics of standardised and IRB approach. When characterising IRB approach we have mentioned "more sophisticated" banks and "some" requirements which has to be fulfilled in order to implement IRB approach into practice. These requirements are discussed in Basel II in section II-H and more detailed in Regulation 575/2013 chapter 3 - section 6. The requirements are quite broad including requirements on rating systems, risk quantification, validation of internal estimates, special requirements for equity exposures and internal governance and oversight.

- Rating system

Rating system is a complex process with the outcome of a final rating for each client and assigning a risk-weight to all exposures. The purpose of using IRB approach is to assign each client to a certain rating grade within a predefined rating scale. Rating scale should allow for meaningful categorization of clients according to their riskiness and the asset quality of their exposures.

- Rating structure

In general assigning an obligor to a rating system shall reflect the level of risk and all the criteria and processes shall be documented and periodically reviewed. Rating system should take into account all obligor and transaction risk characteristics. For retail the number of exposures in given rating grade should be sufficient for meaningful analysis. For corporate the rating scale should have at least 7 grades for non-defaulted obligors and one for defaulted. In both portfolios risk differentiation shall be such as to avoid risk concentrations. Relationship between grades and default risk shall be clearly defined. An example of rating scale for retail and corporate portfolio is in Tables 1.2. and 1.3. Rating scale for retail portfolio is the same we use later for data preparation.

Tab. 1.2: Example of rating scale for retail portfolio

Rating grade	Designation
1	Excellent
2	Very good
3	Good
4	Satisfactory
5	Acceptable
6	Bad
7	Very bad
8	Loss risk
R	Loss-making (default)
N	Not rated

Tab. 1.3: Example of Rating scale for corporate portfolio

Rating grade	Designation
1	Extremely strong
2	Very strong
3	Strong
4a	Good
4b	Very satisfactory
4c	Satisfactory
5a	Lower medium risk
5b	Medium risk
5c	Higher medium risk
6a	Vulnerable
6b	Very vulnerable
7	Weak
8	Risk of loss
R	Loss (default)
N	Not rated

- Assignment to grades or pools

Criteria and definitions for assignment to rating grades shall be sufficiently detailed so that the exposures with similar risk are assigned to the same pool. The documentation should be such as the third party is able to replicate the whole process.

- Assignment of exposures

This should be an integral part of the credit approval process. Each counterparty should be treated separately and separate exposures to the same obligor should be assigned to the same risk grade. Integrity of assignment process is ensured by regular (at least annual) reviews approved by risk manager.

- Use of rating models

Models in use should have a good predictive power, input variables shall form a reasonable and effective basis for predictions and the model shall not have material biases. Data used to develop model shall represent the population and should be controlled for accuracy, completeness and appropriateness. Outcomes of the model should be checked by regular validation.

- Documentation of rating systems

The whole decision process and supporting analyses for choosing rating criteria as well as all major changes shall be documented. When using statistical models outline of theory with assumptions and constraints is a part of a documentation and validation on out-of-time and out-of-sample data is required.

1.3.2. Default

In general default means that counterparty is not able of paying-off anymore. However, we need some criteria to distinguish these clients. According to Basel II accord a default is assigned to an obligor when one or both of the following conditions are met: counterparty is unlikely to pay its obligations in full or the obligor is past due more than 90 days. In line with §452 and §453 of Basel II we distinguish 5 events when counterparty is considered as defaulted in our local conditions.

- Unlikeliness to pay – strong expectation of loss based on subjective evaluation of the financial and non-financial standing.
- 90 days overdue – except for days a minimum threshold for amount is set as €50 for retail and €250 for other clients.
- Defaulted Forbearance – when the altering conditions on account including material forgiveness are approved
- Credit loss – if any credit loss actually happened (due to sale of loan with loss, debt forgiveness or write-off)
- Bankruptcy – when a client enters bankruptcy proceedings or similar protection scheme

Once a client is assigned on any of its exposures as defaulted all of its exposures are treated as in default.

1.3.3. Probability of Default

Following definition of Probability of Default is anchored in Article 160 in [16]. One year PD – indicates the probability of a counterparty defaulting within a 12-month period. So far we have discussed segmentation on an exposure level. However, PD is assigned on a client's level based on rating of a client. Once the client has rating, value of PD is then assigned to each exposure based on asset class and PD associated with that particular grade. For corporate there is a prescribed minimum value of PD which is 0.03%. Any defaulted exposure is assigned PD of 100%.

For estimation of PD for corporate exposures banks may use one or more of the three techniques: internal default experience, mapping to external data and statistical default models. Used information and techniques must take appropriate account of the long-run experience. When using any of the three methods a bank must ensure that the estimation is consistent over time and adjust for any differences and changes. Special attention should be given to consistency of definitions when external data are used.

When estimating PD for retail exposures banks must use the internal data as the main source. External data and statistical models may be used if there is a strong link to external source in data pooling and the risk profile of the data. All relevant and material data must be used.

Irrespective of used method and segment the history of data must be at least five years for at least one source and if a longer history is available, relevant and material data this must be used as well. For retail exposures a different importance may be assigned to historical data if the most recent data are a better predictor.

Rating is a crucial for determining risk parameters. Rating of a client represents creditworthiness of a debtor. Under our conditions we will apply different rating scales for retail and corporate clients. For retail clients we use 9 - grade rating scale, for corporate clients 13-grade scale as was illustrated in Table 1.2. and Table 1.3.

2. Stress Testing

Requirements for stress testing are defined in [10]. Apart from credit operational and market risks banks supervisor is obliged to asses risks disclosed by stress testing. Frequency of stress testing is specified by supervisor, at least once a year. According to Directive of European Parliament no. 575/2013 stress testing is an integral part of risk management and is mandatory for all institutions that use IRB approach on regular basis.

Stress testing shall involve identifying possible events or future changes in economic conditions that could have unfavourable effects on an institutions credit exposures and assessment of the institution’s ability to withstand such changes [16]. Stress test is conducted in order to assess the effect of certain specific conditions on its total capital requirements for credit risk. The test should be meaningful and consider the effects of severe, but plausible, recession scenarios. An institution shall assess migration in its ratings under the stress test scenarios.

2.1. Stress testing methodology

To assess the impact of stress scenario several methods may be engaged. We focus on measuring changes in default rates and migrations of clients in PD grades when macroeconomic factors deteriorate. Firstly, we have to determine which macroeconomic factors have significant impact on PDs. Secondly, we try to test our portfolio against these changes and measure the shifts in PDs. Throughout the time several methods have been developed and used for stress testing of PD. We present a short overview of different statistical models used for this purpose based on the book Stress Testing for Financial Institutions [20].

The first and important thing is the segmentation of the portfolio according to different products and customer groups. However, there are no strict rules how to segment the portfolio. The key principle is to create loan cohorts that are “structurally similar”. According to stress test and correlation to macroeconomic factors the dynamics of the loans is far more important than similar PD. Segmentation based just on temporarily similar PD at the certain time may lead to wrong conclusions. Then a suitable model for each segment may be designed. Other important thing is the choice of the time scale.

Once these important decisions are made we can start with modelling default rate subject to macroeconomic factors. We start with the simplest econometric model

2.1.1. Simple linear regression

$$r(t) = \beta^T \mathrm{Y}(t) + \epsilon(t)\,, \qquad (2.1.)$$

where $r(t)$ is the default rate, $Y(t)$ is a set of macroeconomic factors, β are coefficients and $\epsilon(t)$ is the noise term. This model directly relates default rate to macroeconomic factors and was successfully applied in many contexts. However, the volume and quality of new bookings is often dependent on management decisions and therefore these factors show strong autocorrelation [6]. As a result many of the oscillations in the default rate time series are driven by management rather than macroeconomic effects. For that reason this simple model is not effective as a stress-testing method of a portfolio.

2.1.2. Delinquency migration

Known also as roll rate models are based on tracking if the client fail to make a payment, paid only a portion of the amount or paid the full instalment. If for example, the client fails the payment, he or she “rolls” from one stage of delinquency to the lower one. This process becomes quite complicated when we consider all possible scenarios that may happen. The most general approach is to create Markov model for all the possible transitions where the probability of a transition from stage i to j is defined as

$$p_{ij} = \frac{a_j(t)}{a_i(t-1)}, \qquad (2.2.)$$

Where t denotes the time period. If we consider the client to be in default after 90 days past due, then we have three delinquency stages plus one default stage which gives us 4! (=24) equations. This is usually simplified by computing only the ratio of the number of accounts in each delinquency state with the appropriate time lag.

$$r_i = \frac{a_i(t)}{a_{i-1}(t-1)}\,. \qquad (2.3.)$$

However, the main obstacle of this modelling is that the model is the most effective when the transition rates are stable and unconditional [1]. This is in direct contradiction to the idea of stress testing where we are looking for a correlation between macroeconomic factors and changes in default rates. Therefore effective modelling must be built at the vintage or account level, which incorporates the structural components. Moreover for more plausible results further segmentation is required.

2.1.3. Behavior score migration

When building a model at account level, the procedure is to rank the accounts from most likely “bad” (default) to most likely “good”. The ranking is based either on credit bureau score or in-house behaviour score. For the purpose of stress testing, the ranked accounts are grouped into the bins according to the riskiness and then the PD for each bin is computed based on historic data. After a given time period, transition rates within all bins are calculated according to equation (2.2). To create a stress test model, each migration rate is measured as a time series and modelled using macroeconomic factors. This model suffers from two main shortcomings. Firstly, the structure of the model spreads the macroeconomic signal across many variables due to behaviour score. Therefore it may be very difficult to identify the impacts from the economy especially if the recession in data is not severe enough. Secondly, the migrations may depend on non-economic factors such as maturity of the loan. To prevent confusion from lifecycle and macroeconomic effects it is useful to separate these drivers directly during score creation. Such an approach is developed in proportional hazard models as we describe below.

2.1.4. Static pool of default modelling

Static pool modelling is based on following the behaviour of the fixed group of accounts over time. Advantage of this approach is that the effects from lifecycle are directly treated and instead of the transition matrix we use only few variables which help to concentrate macroeconomic impacts. Two main models are currently used for stress testing retail portfolio.

- **Cox proportional hazard models**

For stress testing purposes modification of the original model is used in a following form:

$$\lambda_i(a) = \lambda_0(a) e^{\boldsymbol{\beta}^T X_i} e^{\boldsymbol{\gamma}^T Y(t)}, \quad (2.4.)$$

where $\lambda_i(a)$ is the probability that the i-th account defaults at age a, $\lambda_0(a)$ is the hazard function which is equivalent to the lifecycle curve, X_i are input factors about the account scaled by $\boldsymbol{\beta}$. The last term is added for the stress testing purposes and directly incorporates macroeconomic factors $Y(t)$ scaled by $\boldsymbol{\gamma}$. This model is used in further developments in next chapters.

- **Dual-time dynamics [11]**

$$r(v, a, t) = f_q(v) e^{f_m(a)} e^{f_g(t)}. \quad (2.5.)$$

The aim is to explain the default rate r by three decomposed factors: maturity (a), vintage (v) and exogenous factor (t). The formula is very similar to the proportional hazards model though the model operates on vintage level rather than on single account level. Moreover the exogenous function $f_g(t)$ and the quality curve $f_q(v)$ are non-parametric functions. The maturation curve $f_m(a)$ is analogous to the hazard function.

2.1.5. Merton's model

Merton's model was originally developed in 1974 for forecasting and pricing corporate defaults. The key idea is that we look at the firm's equity as at the call option on its assets. The firm defaults if value of its assets (R_t) falls below a certain threshold (K) at given time t. The model was then developed based on Black-Scholess framework for European option pricing. It is possible to implement this framework into retail portfolio as well where we model defaults of private clients [20]. The main difference is that while for corporates the value of their assets is somehow derived from financial markets – interest rates, price of bonds…) this is not the case of private clients. The value of client's assets depends on individual factors $U_i(t)$ as well as on general factors $F(t)$ (which are equal for the whole segment). Therefore we model client's asset return at time t as:

$$R_i(t) = \sqrt{\rho} F(t) + \sqrt{1-\rho}\, U_i(t). \quad (2.6.)$$

We assume that the asset return change is normally distributed. Further we assume that general and individual factors are independent and there exist correlation among clients in development of general factors. As we can see from the formula, parameter ρ expresses the degree to which the asset return is linked to general (macroeconomic) factors. We assume that the threshold K depends on macroeconomic variables as:

$$K_i(t) = \beta_0 + \sum_{k=1}^{N} (\beta_{i,k} I_k(t - lag)), \quad (2.7.)$$

where I_k is the k-th macroeconomic factor and time lag incorporates the “reaction” time. When a macroeconomic factor changes, it takes some time for consumer to react on this change, therefore there is certain delay of defaults in data. Now we can express the probability of default of the i-th client:

$$PD_i(t) = P\big(R_i(t) < K_i(t)\big) = P\left(\sqrt{\rho}F(t) + \sqrt{1-\rho}U_i(t) < K_i(t)\right) = \Phi\left(\frac{K_i(t) - \sqrt{\rho}F(t)}{\sqrt{1-\rho}}\right). \quad (2.8.)$$

3. Cox Proportional Hazard Model

For the modelling purposes we decided to use Cox proportional hazard model (CPH). The reasoning is that this model directly incorporates macroeconomic variables into the model. Moreover in CPH model, macroeconomic variables are time dependent (the input of the model is a time series not a fixed value of variable at a certain time). As Crook and Banasik (2005) showed, there is an evidence that aggregate delinquency and write-offs rates vary with the state of economy, therefore we consider such a property as crucial for our purposes. Model was originally introduced by Cox and Oakes (1984), further surveyed by Hosmer and Lemeshow (1999) and later used also in financial context by Stepanova and Thomas (2001). Using CPH model in stress testing is not a novelty, it was applied by Bellotti and Crook (2008) with promising results, therefore we consider this technique as appropriate for our purpose.

3.1. Definition of CPH model

Cox proportional hazard model is based on survival analysis. Originally it predicts the time until death. In our case death means default of a client. Hence at any given time we predict the time until default of the client. Detailed description of the data used is in section 4.1. Data preparation. The initial idea of the model is quite simple [7]. Firstly, we consider T as a random variable – this is the time until default (death). Probability that the client defaults until time t is expressed as following:

$$F(t) = \Pr(T < t). \quad (3.1.)$$

$F(t)$ is known as lifetime distribution function or failure function. It is a complementary function to survival function (probability that the client survives until time t). Therefore we can write:

$$S(t) = \Pr(T \geq t) = 1 - F(t) = \int_t^\infty f(u)du, \quad (3.2.)$$

where $f(t)$ is the probability density function of T. We are interested in probability of an instant default of a client at some time t (intuitively we suppose that the client survives until time t). This is called the hazard function $h(t)$ and it is defined as following:

$$h(t) = \lim_{\delta t \to 0} \frac{P(t \leq T < t + \delta t | T \geq t)}{\delta t}. \quad (3.3.)$$

Cumulative hazard rate is defined as:

$$H(t) = \int_0^t h(u)du. \quad (3.4.)$$

The survival probability can be further expressed in terms of hazard function. From the limit above we have:

$$h(t) = \frac{f(t)}{S(t)} = \frac{\frac{d}{dt}[1 - S(t)]}{S(t)} = -\frac{\dot{S}(t)}{S(t)}. \quad (3.5.)$$

Then the survival probability can be expressed in terms of hazard function as:

$$S(t) = \exp\left(-\int_0^t h(u)du\right). \quad (3.6.)$$

However we have still not specified survival function neither hazard function nor density function $f(t)$ of the survival time T. As we are interested in probability of default, our final interest is in finding hazard function as it gives us the corresponding information. One of the most common ways is to use exponential function which is known as the Cox proportional hazard model. The model takes the following form:

$$h(t, x(t), \beta) = h_0(t)\exp\left(\beta^T x(t)\right), \quad (3.7.)$$

where

- $x(t)$ is a vector consisting of:
- client's individual variables, fixed at the time of application
- macroeconomic variables, which are dependent on time
- β is a vector of estimated coefficients
- $h_0(t)$ is estimated baseline hazard function dependent on time

3.2. Estimation of CPH model

After we have defined the model (3.7.) we need some robust methods to estimate the coefficients. In following subsection we introduce technique for estimation of the model on real data. Firstly, we estimate β by partial likelihood function, which allows us to make estimation for β without knowing baseline hazard function [5].

$$L_p(\beta) = \prod_{i=1}^{n}\left(\frac{\exp(\beta x_i t_i)}{\sum_{j\in R(t_i)}\exp(\beta x_j t_i)}\right)^{\delta_i}, \quad (3.8.)$$

n – is number of observations (clients)
δ_i - is indicator function, where $\delta_i = 1$ for defaulted client, $\delta_i = 0$ for non-defaulted client
$R(t_i) = \{j : t_j \geq t_i\}$.

For the macroeconomic variable we use the value at the time of default (t_i). It might be not so obvious from the formula that the value of t_i is the same for every observation but since the clients has different beginning, the actual value of macroeconomic variable is different. This incorporates the dynamics into the model. For the computation purposes log-likelihood function is used:

$$\log L_p(\beta) = \sum_{i=1}^{n} \delta_i \left\{ \beta x_i t_i - log \sum_{j \in R(t_i)} \exp(\beta x_j t_i) \right\}, \quad (3.9.)$$

Maximisation of log-likelihood function is accomplished by Newton-Rhapson method[1]. Even though we have estimates for $\hat{\beta}$, there is still long way to the final result. We still need to estimate the hazard function. There are several ways how this can be done, non-parametric [18] as well as parametric [19]. Firstly, we estimate cumulative baseline hazard function according to [5]:

$$\widehat{H_0}(t) = \sum_{t_i \leq t} \frac{\delta_i}{\sum_{j \in R(t_i)} \exp(\beta x_j t_i)}. \quad (3.10.)$$

Then we estimate $h_0(t)$ from the cumulative hazard function $\widehat{H_0}(t)$ in a following way [19]. Let us assume the times-to-event (variable $T2$ in an input data set) follow a Weibull distribution. We model $H_0(t)$ with respect to time t through natural logarithm:

$$z_0(t) = \ln\big(H_0(t)\big) = \gamma_0 + \gamma_1 \ln(t), \quad (3.11.)$$

where γ_0, γ_1 are transformations of the parameters of the Weibull distribution. We estimate the parameters by OLS. Then the baseline hazard function is derived from $H_0(t)$ through its derivative with respect to(t):

$$h_0(t) = \frac{dH_0(t)}{dt} = \frac{dln(H_0(t))}{dln(t)} \frac{dln(t)}{dt} \frac{dH_0(t)}{dlnH_0(t)}. \quad (3.12.)$$

And we obtain estimation of baseline hazard function:

$$\widehat{h_0}(t) = t^{-1} \widehat{\gamma_1} \exp(\widehat{\gamma_0} + \widehat{\gamma_1} \ln(t)). \quad (3.13.)$$

As Royston states in the article [19], there is no particular reason why Weibull model should fit the data, which opens room for estimates with non-linear terms in t or $\ln(t)$ respectively. However, in the case of our model simple linear transformation fits our data quite well (Appendix- Fig. 7.6.) therefore we do not go deeper in this particular case and use this estimation for $\widehat{h_0}(t)$.

There are almost no assumptions on the data entering the model. The only crucial assumption is the “proportionality”. It practically means that if an individual i is today twice as risky as an individual j, he or she must be twice as risky tomorrow and this holds for any given time we observe. Mathematically we write:

[1] Newton-Rhapson method is standardized method in SAS procedure proc phreg.

$$HR = \frac{h_i(t)}{h_j(t)} = \frac{h_0(t)\exp(\beta_1 x_1^i + \beta_2 x_2^i + \cdots + \beta_k x_k^i)}{h_0(t)\exp(\beta_1 x_1^j + \beta_2 x_2^{ij} + \cdots + \beta_k x_k^j)}$$
$$= \exp\{\beta_1\left(x_1^i - x_1^j\right) + \beta_2\left(x_2^i - x_2^j\right) + \cdots + \beta_k\left(x_k^i - x_k^j\right)\}, \quad (3.14.)$$

So the HR is independent of time.

In this chapter we introduced the mathematical background for Cox Proportional Hazard model and estimation techniques which we use in next section to estimate equation 3.7. on real data.

4. Model Estimation on Real Data

In Chapter 3 we have defined the CPH model and methods we will use for its estimation. In this Chapter we will fit the model to our data. We will work with mortgage loans portfolio. Firstly, we define the data and macroeconomic variables which we want to analyse. Secondly, we analyse candidates for the final model and lastly, we conclude the results and choose the final model.

4.1. Data preparation

We work with portfolio of mortgage loans. To create our data set for estimating the model we work with data from January 2008 to December 2015. Moreover, data from 2007 were used but only for information about rating as is clarified below. Data are collected in monthly snapshots. Following exclusion rules were applied in order to make estimation on risk-relevant and correct data:

- all non-risk relevant observations (observations with exposure <= 0)
- observations with any data missing were excluded.

We used following information as an input of the model:

- ID number of client
- Date of record (Record Date)
- T1, which denotes the time of a client since its beginning to actual date (in months)
- T2, where T2-T1 is the time period during which the explanatory variables are constant on a client level
- Default Flag with value 0 if client is not in default and 1 if client is in default
- Rating of the client one year prior to date of record (numeric 1-9, 9 is default)
- Macroeconomic values

Rating information is a part of initial data and provides complex information about the riskiness of the client based on its personal characteristics. We do not focus on analysing individual characteristics of clients and use rating information as the most relevant one according to its riskiness. According to the macroeconomic variables some of them are usually reported quarterly (see the Table 4.2.). We used linear interpolation in order to obtain monthly data. Example of the data used as an input to the model is in Table 4.1.

As a modelling environment, SAS enterprise guide was used. To estimate the coefficients in CPH model we used the “proc phreg” procedure – Fig.4.1. [21].

Tab. 4.1.: Illustration of Data used for modelling

Record Date	ID client	T1	T2	Default Flag	Rating	GDP_rlgrth
31.1.2008	4318	0	1	0	2	12,1
29.2.2008	4318	1	2	0	2	10,7
31.3.2008	4318	2	4	0	2	9,3
31.5.2008	4318	4	5	0	2	6,6
31.3.2009	54937	0	1	0	6	-5,9
30.4.2009	54937	1	2	0	6	-6,0
31.5.2009	54937	2	3	0	8	-6,0
30.6.2009	54937	3	4	1	9	-6,1

```
proc phreg data=Input_data plots(overlay)=survival simple;
id ClientID;
model T2*Def_Flag(0)= Rating_YA GDP_rlgrth /corrb risklimits entry=T1;
output out=Out survival=surv xbeta=beta resmart=Mart resdev=Dev ressch=Resch;
run;
```

Fig. 4.1.:Proc phreg procedure estimated in SAS Enterprise Guide.

4.2. Macroeconomic variables

Selection of macroeconomic variables which enter the model is crucial for our analysis. As our purpose is to find out what is the relation between the state of economy and the default rate in our portfolio we were focused mainly on commonly used economic drivers. All entering variables are listed in the table below. We used data from Statistical Office of Slovak Republic, Eurostat and National Bank of Slovakia. Predictions were calculated by internal department of the bank.

Tab. 4.2.: List of Macroeconomic variables used for the analysis

Abbreviation	Description	Data	Unit
GDP			
GDP_Nom	Nominal GDP	Quarterly	EUR mil.
GDP_PPY	GDP in constant prices on the base of the previous year	Quarterly	EUR mil.
GDP_chain	GDP chain linked	Quarterly	EUR mil.
GDP_rlgrth	Real growth of GDP	Quarterly	%
YY_GDP_PPY	Rate of change of GDP_PPY to the same month in previous year.	Quarterly	%
QQ_GDP_PPY	Rate of change of GDP_PPY to the three months prior	Quarterly	%
Trade Balance			
Export	Export of goods in mil. EUR	Monthly	EUR mil.
Import	Import of goods in mil. EUR	Monthly	EUR mil.
Trade_Bal	Trade Balance = Export - Import (in mil. EUR)	Monthly	EUR mil.
TB_YY	Rate of change of Trade Balance to the same month in previous year	Monthly	%

Abbreviation	Description	Data	Unit
TB_QQ	Rate of change of Trade Balance to the three months prior	Monthly	%
Export_YY	Rate of change of Export to the same month in previous year	Monthly	%
Export_QQ	Rare of change of Export to the three months prior	Monthly	%
Import_YY	Rate of change of Import to the same month in previous year	Monthly	%
Import_QQ	Rare of change of Import to the three months prior	Monthly	%
Prices			
Infl_CPI	Inflation rate CPI	Monthly	index
HICP	Harmonized inflation index	Monthly	index
HICP_MrCH	Monthly rate of change of HICP	Monthly	%
HICP_ArCH	Annual rate of change of HICP	Monthly	%
Other			
Unempl_rate	Unemployment rate	Quarterly	%
Rl_wage_grth	Real growth of wages	Quarterly	%
Hous_cons	Household consumption	Quarterly	EUR mil.
HC_chain	Household consumption chain linked	Quarterly	EUR mil.
HC_yy	Houcehold consumption rate of change compared to the same month in previous year	Quarterly	%
IES_1M	Index of economic sentiment	Monthly	index
EUR_USD	Exchange rate EUR/USD	Monthly	index

Before we entered the variables into the model, we plotted the values together with the default rates just to have a quick glance on the possible relation. It is not a proper statistical method but it is very useful to form our expectations about the results. As we will see later statistical results are in line with the graphical analysis. We do not provide plots for all the variables only for those which were selected as significant when entering the model.

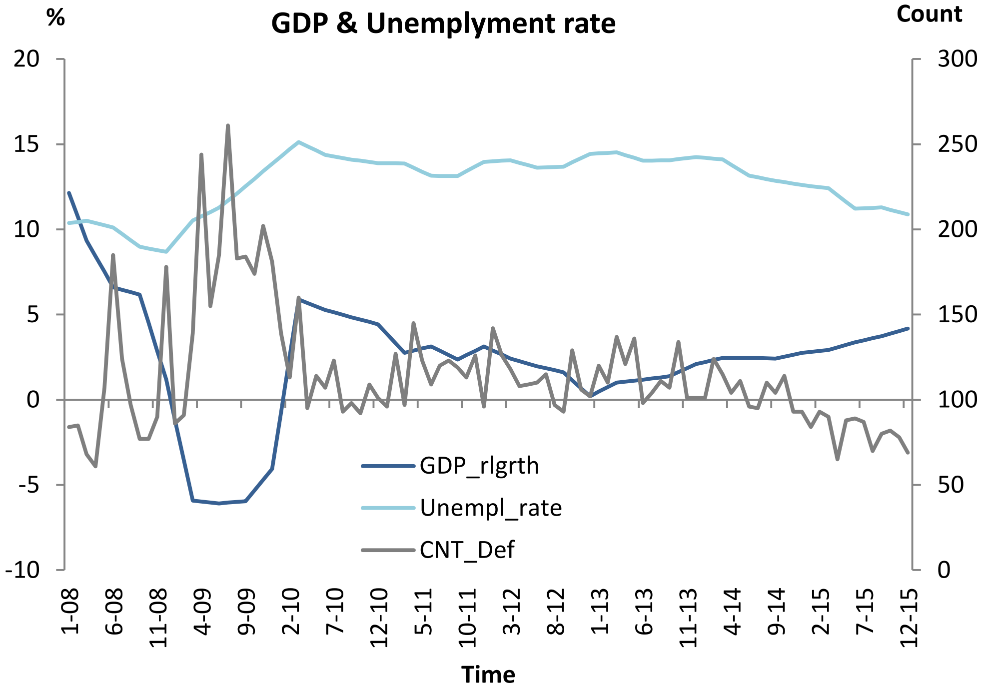


Fig. 4.2.: Development of economic drivers: real growth of GDP and Unemployment in comparison with the realised number of defaults (CNT_Def)

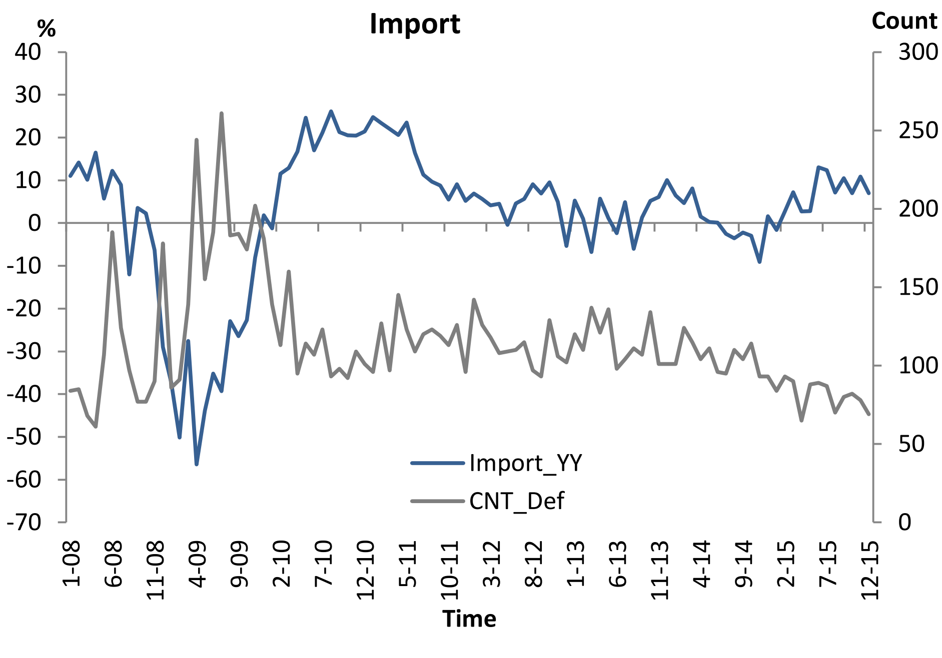


Fig. 4.3.: *Development of the annual rate of change of Import in comparison with the realised number of defaults (CNT_Def)*

4.3. Estimation of the coefficients

Firstly, we entered all of the macroeconomic variables as explanatory into the model. However, due to their correlations and non-linearity of the model there were too many influences and we came up with no significant result for any macro variable. Only the information about rating was significant. Therefore we decided to proceed the other way around. We kept the variable "Rating" and analysed all other variables one by one. We ended up with much shorter list of variables with satisfying results. The reasons for exclusion were several.

Hazard ratio (3.14.) for some of the variables was equal to unity or unity was inside of 95% confidence interval for HR. It means the variable has no explanatory power. This was followed by estimates of $\hat{\beta}$ very close to zero which is in line with intuition. This reason adopts for variables: GDP_Nom, GDP_PPY, GDP_chain, Infl_CPI, Hous_cons, HC_chain, Rl_wage growth, Trade_Bal, HICP_MrCH, HICP_ArCH, Export_QQ, Export, TB_YY, TB_QQ, Import_QQ and Import.

We obtained quite nice HR for variable EUR_USD but the confidence interval was too broad which indicates the HR is not constant over time which is in contradiction with the assumption of the model.

After this initial analysis we were left with variables GDP_rlgrth, YY_GDP_PPY, Unempl_rate, HC_yy, IES_1M, HICP_index, Export_YY and Import_YY. Some variables give similar information (GDP_rlgrth, YY_GDP_PPY) and / or are highly correlated (Import_YY, Export_YY) therefore the one with lower AIC criterion was chosen. Based on this rule YY_GDP_PPY and Export_YY were excluded (Complete table with results – Appendix - Tab. 7.2. and 7.3.).

In order to explain as much data as possible, further analysis with left variables was conducted. We have tried combinations of two and three different variables (except for rating). However, the improvement of the model and estimates of the coefficients were quite mild (Appendix – Tab. 7.4. and 7.5.). For example when combining GDP_rlgrth with other variables we came up with HR very close to 1 (or 1 in the 95% confidence limit) for all tested variables. This somehow implies that the real growth of GDP contains significant part of the "market information" and is correlated with almost all other indicators, which is again in line with intuition. For combination of three macroeconomic variables we were not able to find any model that would satisfy conditions either on HR.

As there is no right solution for a selection of the final model, we proceeded as following. We have chosen the model for single macroeconomic variable with the lowest AIC criterion (Appendix – Tab. 7.3. This left us with variable Import_YY. However, results for GDP_rlgrth were quite good as well therefore we decided to further analyse the results for this variable as well. To complete the list, we included in further analysis the model with two variables with the lowest AIC as well (Appendix – Tab.7.5.).

4.4. Model with real growth of GDP (GDP_rlgrth)

Results from the model run are summarized in the table below (Tab. 4.3.). From the table it is obvious that the variable Rating has much more explanatory power than the GDP: HR is nearly 2, whereas HR for GDP is still quite close to 1. This was the scenario for every single variable we have tested. This implies that the individual contribution to the riskiness is much higher than the risk emerging from the market. However, it must be noted that this is the result for the portfolio of the mortgage loans where the individual characteristics are the main risk indicators.

The relatively high level of standard error of estimated coefficients may be concerning, especially the one for GDP_rlgrth. However, when we look at the HR and confidence limits for HR we can see that it never crosses the value of unity. It indicates that the impact of the estimated variable on the HR is quite stable and we consider this result as acceptable.

Tab. 4.3.: Estimation of the coefficients for the model with real growth of GDP[2]

Parameter	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence	
Rating	0.66824	0.00444	22655.4069	<.0001	1.951	1.934	1.968
GDP_rlgrth	-0.05735	0.00397	208.2143	<.0001	0.944	0.937	0.952

The developed model has now the following form:

$$h(t) = h_0(t)\exp\big(0.66824xRating - 0.05735xGDP_{rlgrth}\big). \quad (4.1.)$$

In order to obtain baseline hazard function we proceeded as described in subsection 3.2. From the OLS estimation we have formula for baseline hazard function:

$$\widehat{h_0}\,(t) = t^{-1}0.85944\exp(-8.06366 + 0.85944\ln(t)). \quad (4.2.)$$

Equations 4.1. and 4.2. give us complete estimated model. The graph in Fig. 4.4. shows the realized default rates and the estimated hazard function – estimated PD.

The estimated model does not capture the monthly jumps in DR however it follows the trend quite well. We do not observe somehow delayed reaction of the model to the market development. Therefore if we are interested in forecasting number of defaults in upcoming year this model seems to be satisfying. Moreover the model captures the DR quite well also subject to distribution in rating grades (Fig.4.5.) We observe slight overestimation of the risk in the last rating grade which is actually not that disturbing property. The vice versa scenario would be undesirable.

[2] As suggested in [5] and [7] variable with interactions of Rating and GDP should enter the model as well. We conducted estimation of the model with this variable as well and provide the results in Appendix in Tab. 7.1. However, for such a variable the HR was very close to one and we consider model without this variable as more suitable. Value of AIC is slightly lower 228245.28 compare to 228249.72 which is a minor decrease.

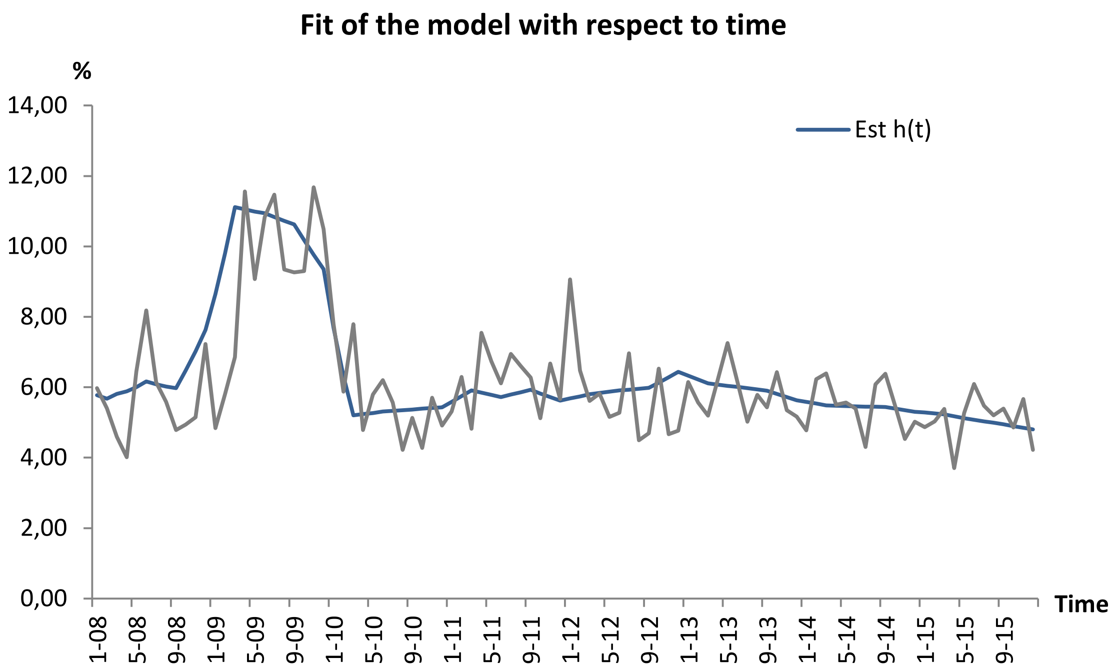


Fig. 4.4.:Estimated model with GDP in comparison with realised default rates

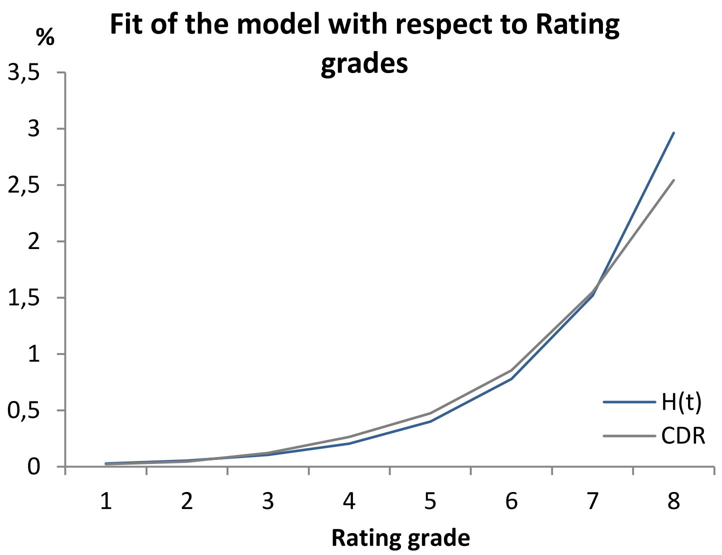


Fig 4.5.:Comparison of estimated cumulative hazard function H(t) with realised cumulative default rates.

Even though the fit of the model looks good we have to check if the HR is constant over time – assumption of proportionality, and if the variables enter the model in a correct functional form – Chapter 5.

4.5. Model with Import (Import_yy) and Unemployment rate (Unempl_rate)

We obtained good results also for variable Import_yy. As mentioned in previous section, we were looking also for a model with more than one macroeconomic variable in order to examine if it is more suitable for our purposes. As explained before, we have chosen model with variables Import and Unemployment rate. We provide the estimation of the coefficients for both models.

- Estimation of the model with Import

Tab. 4.4.: Estimation of the coefficients for the model with Import

Parameter	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
Rating	0.6772	0.00468	20964.2726	<.0001	1.968	1.95	1.986
Import_YY	-0.00951	0.000738	166.1652	<.0001	0.991	0.989	0.992

The technique used for the estimate of the final model is analogous as in subsection above. Baseline function is estimated as:

$$\widehat{h_0}\,(t) = t^{-1}1.40075\,\exp(-10.68463 + 1.40075\ln(t)). \quad (4.3.)$$

- Estimation of the model with Import and Unemployment rate

Tab. 4.5.: Estimation of the coefficients for the model with Import and Unemployment rate

Parameter	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
Rating	0.67698	0.00468	20898.2881	<.0001	1.968	1.95	1.986
Unempl_rate	0.05374	0.01271	17.8636	<.0001	1.055	1.029	1.082
Import_YY	-0.01166	0.0008996	168.0209	<.0001	0.988	0.987	0.99

The baseline hazard function:

$$\widehat{h_0}\,(t) = t^{-1}1.44732\,\exp(-11.58651 + 1.44732\ln(t)). \quad (4.4.)$$

The AIC criterion is 203 789.7 for the model with import, and 203 773.5 for the model with import and unemployment rate. Decrease in the value of AIC is mild when adding variable. However, compare to the model with GDP (AIC=228 249.7) we see considerable improvement. Though when we investigate HR and its confidence limits the results are not that optimistic. Value of HR for macroeconomic variables is in both cases close to unity, while for GDP the value is 0.944. Graphical analysis (Fig. 4.6.) confirms that the fit of the models is much worse than the one with real growth of GDP.

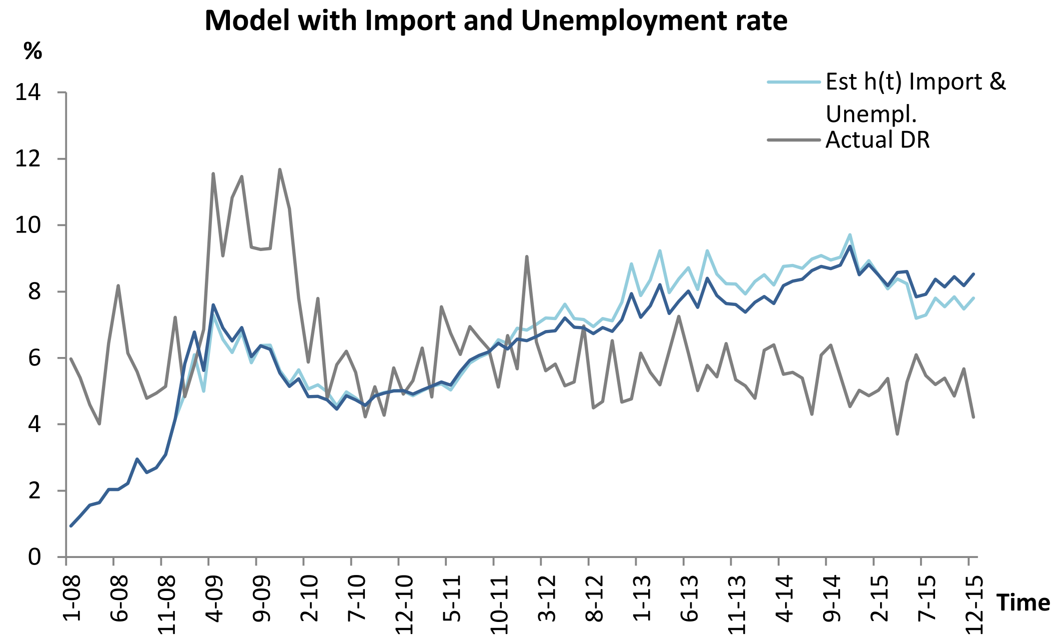


Fig.4.6.: Estimated model with Import, and model with Import and Unemployment rate in comparison with realised default rates

We can see in the Fig.4.6. that even though the estimations of the coefficients were promising, the estimated model does not fit the data neither in case of including only import nor when including also unemployment rate. As other variables were excluded before, we conclude that the real growth of GDP is the most suitable explanatory variable for our data. Therefore in further analysis of residuals we consider only the model based on GDP.

5. Check of the Fit of the Model

In this section we refer to [7]. After we have estimated model for GDP based on equations 4.1. and 4.2. we test whether the assumptions are fulfilled. The examination is mainly graphical based on plots of residuals. There are no strict rules on accepting or rejecting the model, we are left only with indications if the model is estimated properly or not. As the model itself has different properties as i.e. linear regression also the residuals are defined differently. There are several types of residuals for CPH model depending on the assumption we want to examine. Firstly, we check whether the variables are correlated and then we compute the residuals.

The correlation is below 0.1 therefore it is suitable to include both variables into the model at the same time.

Tab. 5.1.: Estimated correlation matrix

Parameter	Rating	GDP_rlgrth
Rating	1	0.0972
GDP_rlgrth	0.0972	1

5.1. Cox-Snell residuals

Cox-Snell residuals are defined simply by estimated cumulative hazard function:

$$r_{C_i} = \widehat{H}_i(t_i^*) = -log\widehat{S}_i(t_i^*), \quad (5.1.)$$

where t_i^* is survival time of i-th individual. The interpretation is based on following facts. If T (survival time) is a random variable with corresponding distribution function $S(t)$, then the random variable $Y = -\log S(T)$ has an exponential distribution with unit mean, irrespective of the form of $S(t)$. Therefore if the model fitted to the observed data is satisfactory, estimated survivor function $\hat{S}(t_i)$ will have similar values to the true value $S(t_i)$ and also similar properties. Then the residuals $r_{C_i} = -log\widehat{S}_i(t_i^*)$ will behave as n observations from a unit exponential function. After computing Cox-Snell residuals, the Kaplan Meier estimate of the survivor function of these values is found. If we then plot $-log\widehat{S}_i\left(r_{C_i}\right)$ against the residuals we should obtain straight line through the origin with unit slope. Otherwise the model is not fitted properly.

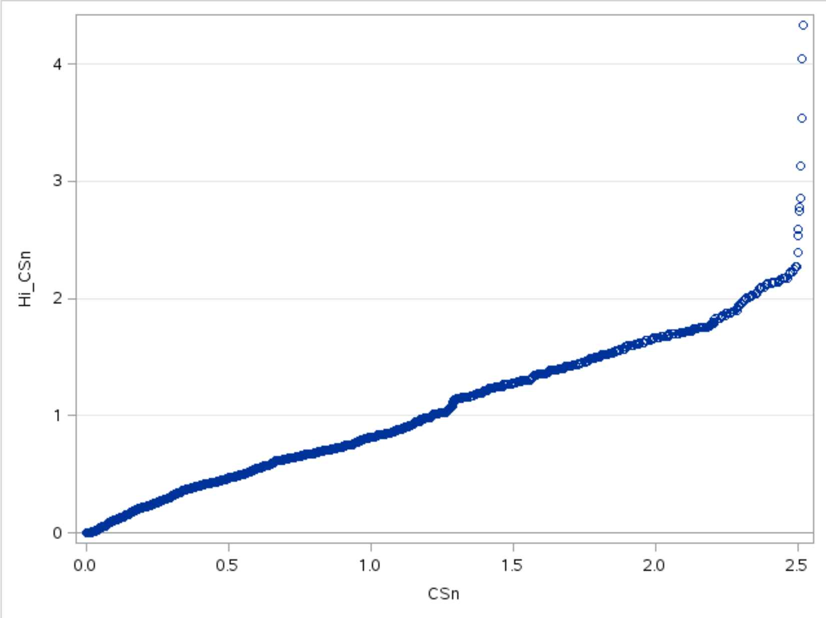


Fig. 5.1.: Cumulative hazard function for the Cox-Snell residuals respect to Cox-Snell residuals (CSn)

In this case we observe that indeed the plot of the residuals gives us a straight line with almost unit slope. There is 19 residuals with value greater than 2.5 though compare to the total number of residuals 146 400 this is a negligible number. However, Cox-Snell residuals have limitations in assessing model adequacy [7]. We can interpret a good result rather as the model is not fitted bad than that the model is good. Further investigation of other types of residuals should be performed.

5.2. Martingale residuals

Slight modification of Cox-Snell residuals is often made in order to properly distinguish the censored data also in the computation of residuals. Therefore indicator function δ_i is introduced, which takes value of zero if the observed survival time is censored and the value of unity[3] if it is uncensored. Further simple transformation leads to definition of Martingale residuals[4].

$$r_{M_i} = \delta_i - r_{C_i}. \quad (5.2.)$$

Martingale residuals take values in range $(-\infty, 1]$. The interpretation is that we measure the difference between the realization of i-th individual at the survival time, which is either zero or unity, and the estimate. Therefore the closer to zero are the observations the better fit. On the other hand, values close to unity indicate defaults which are not captured by the model. We provide an index plot for Martingale residuals (Fig. 5.2.).

[3] Crowley and Hu (1977) found that inflated the residual to too great extent. Therefore they suggested value of 0.692 – median of the unit exponential distribution.

[4] Martingale residuals can be derived also through *martingale methods* which relies heavily on probability theory and stochastic processes [22].

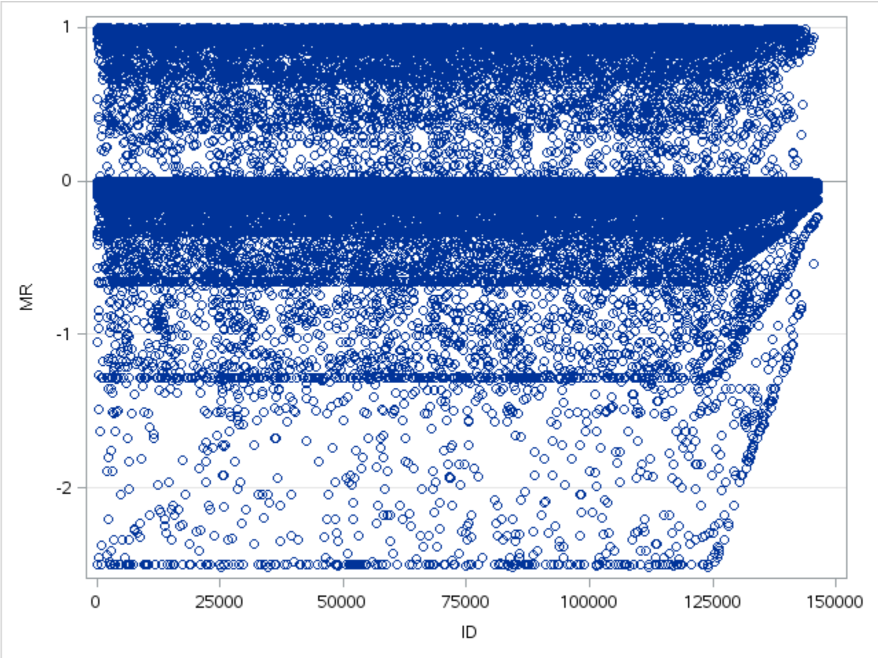


Fig. 5.2.: Index plot of Martingale residuals

We can observe lot of values close to unity on the graph above. However, analysis based only on graphs can be quite misleading due to the high number of plotted observations. There is 146 000 individuals under the study while 11 026 is uncensored (in default). Therefore much more observations are still close to zero than those close to unity. Though from the definition of Martingale residuals the result is still concerning, individuals in default have the estimated probability of failure close to zero. Indeed more than half of the uncensored observations (6567) have the value of the Martingale residual greater than 0.75 and only 2854 less than 0.5. The histogram of the residuals is in Appendix – Fig.7.9.

Since this is a crucial in the model we provide further analysis with the aim to find some pattern for the data which are not fitted well. Firstly, we define data which are not fitted well as those with value of Martingale greater than 0.5. We provide a plot of Martingale residuals (Fig. 5.3.) respect to Rating of the client for the uncensored observations (clients in default). There is observable pattern in the graph below. The better Rating has the client, the higher is the value of Martingale residual. It is quite intuitive, we do not expect client with a good rating to default, their estimated hazard rate (or probability of default) is therefore estimated as quite small. Therefore we conclude that the proposed model preserves the required properties, the reason of high values of residuals is the good rating of the clients who defaulted. This may be due to the one year time lag of the used Rating or because of an unappropriated risk assessment. However, variable Rating is taken as an input to the model and it is behind the scope of this work to analyse reasons of its possible inappropriateness.

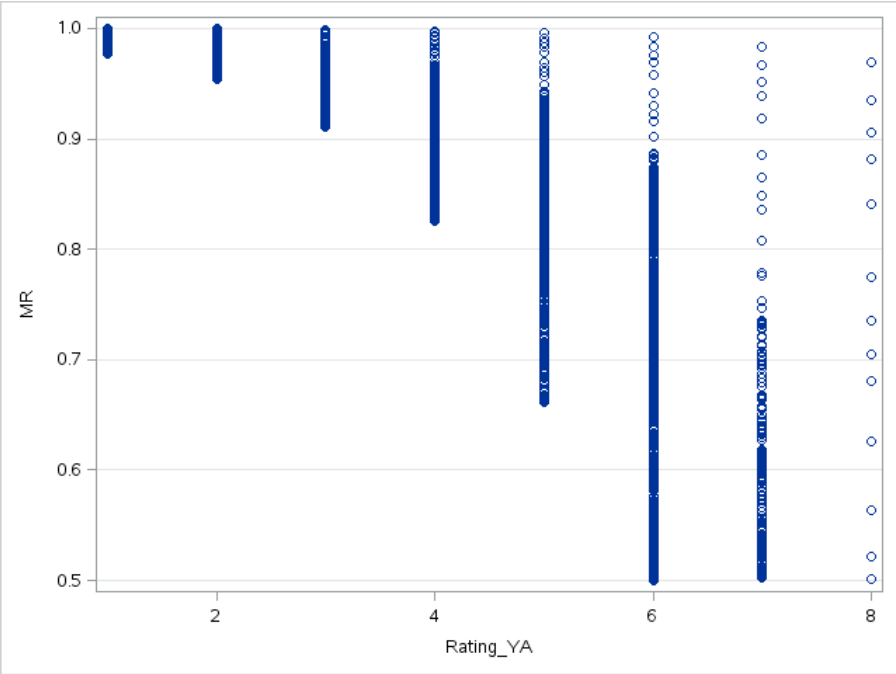


Fig. 5.3.: Plot of the Martingale Residuals respect to Rating of the client

5.3. Deviance residuals

Deviance residuals are defined as a transformation of martingale residuals in order to obtain residuals which are symmetrically distributed around zero, if the model is fitted properly.

$$r_{D_i} = sgn(r_{M_i})\,[-2\{r_{M_i} + \delta_i \log(\delta_i - r_M)\}]^{1/2}. \quad (5.3.)$$

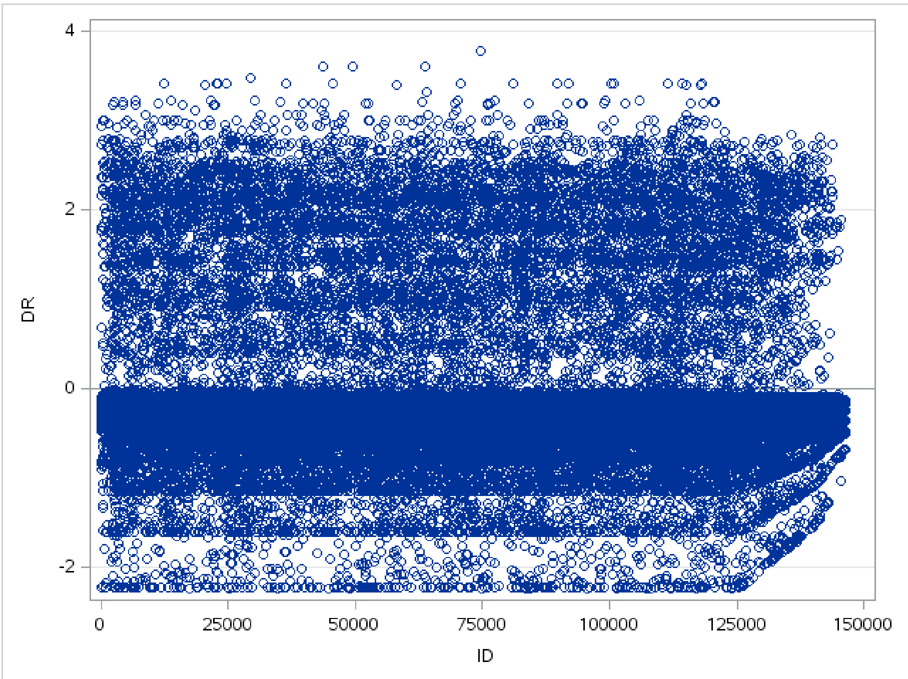


Fig. 5.4.: Index plot of the Deviance residuals

We observe two phenomena in Fig. 5.4. First one is a strong concentration from the lower side of the zero. Closer look to the expression for the deviance residuals shows that those are values for negative Martingale residuals. We can explain this concentration by domination of censored observations over uncensored. Second one is the spread in positive values. We obtain higher positive values of deviance residuals in case of Martingale residual close to unity. This is in line with the previous result, which we have already analysed.

We have some indications from the residual analysis that there is a room for a better fit of the model. One of the reasons might be that the explanatory variables do not enter the model in a correct functional form. Therefore we examine if the linear functional form is proper for variables Rating as well as GDP_rlgrth.

5.3.1. Check of the functional form of the variables

According to [7] we try to fit the model with different functional form of the variable and compare the AIC criterion. If there is a significant decrease in the value of the statistics we should consider other than linear functional form of the variable. Commonly used is logarithmic transformation but this is unsuitable since GDP_rlgrth has also negative values. We are left with the powers of the variable.

Tab. 5.2.: Characteristics of the model fitted by different functional form of GDP

	-2 Log	AIC	Hazard Ratio	95% Hazard Ratio Confidence Limits	
GDP_rlgrth	228245.72	228249.72	0.944	0.937	0.952
GDP_rlgrth^2	228413.84	228417.84	1.007	1.005	1.010
GDP_rlgrth^3	228275.80	228279.80	0.998	0.998	0.998

Statistic -2Log as well as AIC have slightly increased in both cases: GDP_rlgrth^2 and GDP_rlgrth^3 compare to the linear form of GDP_rlgrth. Also the HR is very close to unity which means that the variable has almost no explanatory power. Therefore we reject these transformations as possible enhancement of the model.

However, the non- linearity of the variable may have different form than the one we have tried. In order to identify such non-linearity we follow the [Collet] methodology and we will treat the variable as a factor. We order the values and split them into four approximately equally sized groups. Each group has then its own corresponding number – factor. As a next step we create four indicator variables corresponding to each level of factor. Then we compare the results from fitting the model with factor variable with that from fitting with indicator variables. Difference between the values of AIC is a measure of non-linearity across the levels of the variable. Creation of factor and indicator variables is illustrated in Tab. 5.3.

Tab. 5.3.: Indicator variables for GDP

Range of GDP	GDP_fac	GDP_1	GDP_2	GDP_3	GDP_4
(-6.1, 1.3)	1	1	0	0	0
(1.3, 2.6)	2	0	1	0	0
(2.6, 3.8)	3	0	0	1	0
(3.8, 12.1)	4	0	0	0	1

Value of AIC criterion is the same for both cases (228408.46). Based on this and previous result we conclude that the linear functional form of the variable GDP_rlgrth is appropriate. The variable Rating was tested in analogous way. As it is already factor variable (with nine levels) we compare only AIC of the model fitted by the factor variable with the model fitted by indicator variables. There is mild decrease in AIC (227987.49 compare to 228249.72) when the model is fitted by indicator variables. However, due to strong correlations between indicator variables we leave the variable Rating as a factor.

5.4. Schoenfeld residuals

So far we have tested the if the model satisfy basic properties and if the variables enter the model in the correct functional form. However, the crucial assumption of the model is that the HR (3.14) is constant over time. For this purpose we use Schoenfeld residuals. The main difference comparing to previous residuals is that we estimate residual for each individual and each variable in the model. The residual for i-th individual and j-th variable is given by:

$$r_{P_{ij}} = \delta_i \{x_{ji} - \widehat{a_{ji}}\}, \quad (5.4.)$$

where x_{ji} is the value of the j-th variable for the i-th individual,

$$\widehat{a_{ji}} = \frac{\sum_{l \in R(t_i)} x_{jl} \exp(\hat{\beta} x_l)}{\sum_{l \in R(t_i)} \exp(\hat{\beta} x_l)}, \quad (5.5.)$$

and $R(t_i)$ is the set of all individuals at risk at time t_i. Note that the residuals are non-zero only for uncensored observations. As shown in [23] the expected value of ij-th residual is $E\left(r_{P_{ij}}\right) \approx \beta_j(t_i) - \widehat{\beta_j}$, where $\beta_j(t_i)$ is time varying coefficient of the variable at death time t_i and $\widehat{\beta_j}$ is the estimated coefficient in fitted CPH model. Consequently, a plot of the values of $r_{P_{ij}} + \widehat{\beta_j}$ against the death times provides information about the form of the time varying coefficient of the variable. A horizontal line suggests that the coefficient of j-th variable is constant and the proportional hazard assumption is satisfied.

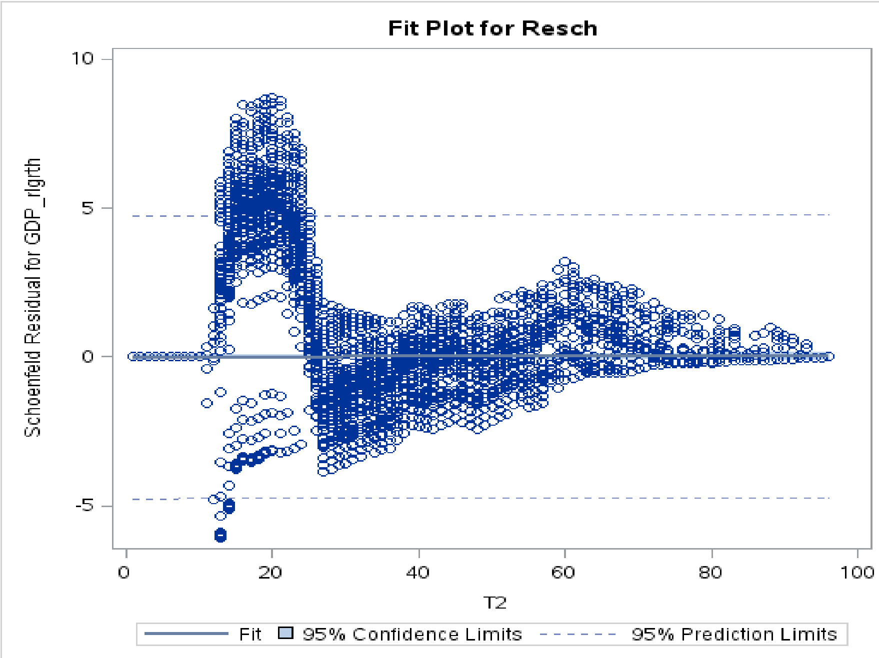


Figure 5.5.: Plot of the Schoenfeld residuals for GDP_rlgrth with a fitted straight line.

Even though the fitted line has slope equal to zero we observe non-linearity at the beginning. If we compare it with the development of GDP (Fig. 4.2.) we notice that trend of residuals corresponds to the trend of variable. The impact of the variable differs in crisis years 2008 and 2009 and the later years which are under the study. In previous sections we tested if there is a functional form for GDP which would fit the data better with negative result. However, for stress testing purposes we are interested in extreme development of GDP. Therefore the data cannot be time consistent. From this point of view, we consider this result as expectable and in accordance with intuition. From the mathematical definition of the model and its assumptions we interpret this result as a warning. At present state the data are fitted quite well but re-estimation of the coefficient for GDP might be needed, especially if there will be noticeable change in trend.

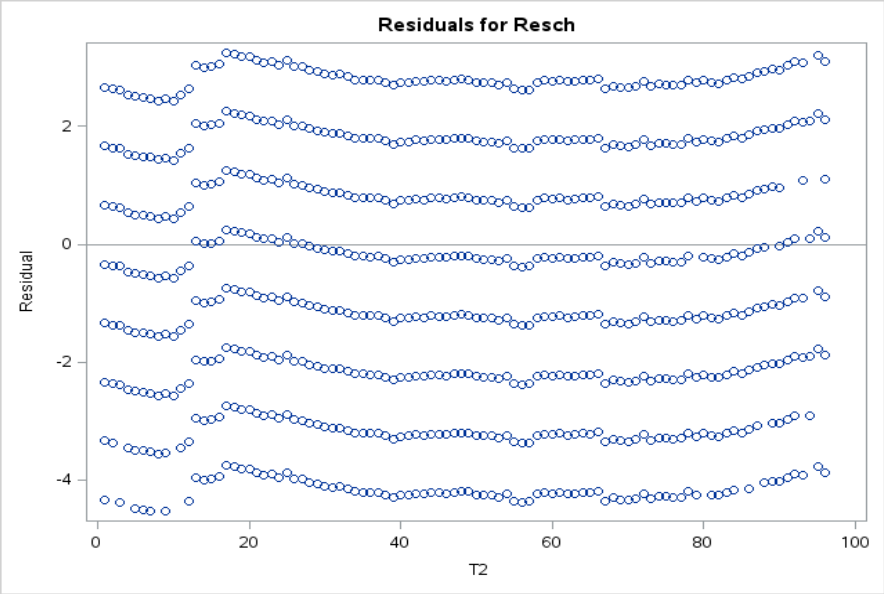


Figure 5.6.: Plot of the Schoenfeld residuals for Rating

For the variable Rating we observe eight patterns, these correspond to eight levels of the variable (last ninth level is for the default and its observable only at the time of the event therefore there is no representation for it in the residuals). Each of the pattern is almost straight line with zero slope. Therefore we conclude that for the explanatory variable Rating the value of HR is the same over time and the assumption is fulfilled.

6. Interpretation of Results

After we have analysed which variables have significant influence on the curve of default rates, we have chosen our preferred model. As a final predictors we take Rating and real growth of GDP. We have estimated the coefficient for these two variables. Now we are interested how we can interpret these estimates and how big is the influence on the predicted default rates when these parameters change.

From the definition of the hazard function (3.7.) we can write that if the j-th variable increases by one the hazard of the i-th individual changes $\exp(\beta_i)$ times which is the value of the estimated HR (Tab.4.3.). In case the sign of the coefficient is positive hazard increases, if the sign is negative it decreases. For our estimated coefficients $\hat{\beta}$ we obtain following result: if the variable Rating increases by one, the probability of default increases 1.95 times, in case real growth of GDP increases by 1, PD decreases 0.94 times.

However, the final interest is in the capital which has to be allocated in the bank to cover possible losses. The regulatory capital requirement is given by capital adequacy ratio:

$$Capital\ ratio = \frac{Regulatory\ capital}{Risk\ weighted\ assets} \geq\ 8\%. \quad (6.1.)$$

Risk weight for mortgage loans according to Basel II is given by [14]:

$$RW\ = 12.5 \times 1.06 \times \left[LGD \times \phi\left(\frac{1}{\sqrt{1-R}} \times \phi^{-1}(PD) + \sqrt{\frac{R}{1-R}} \times \phi^{-1}(0.999) \right) - PD \times LGD \right] \text{[5]} \ (6.2)$$

and the total risk weighted assets are obtained as product of risk weight and amount of loans [3]:

$$RWA = RW \times EAD. \quad (6.3.)$$

As our main purpose is to measure impact of macro-economic changes we analyse changes of GDP only. We compare three scenarios of the development of GDP. We consider months May 2016 until December 2016. First scenario is standard prediction of GDP, second one replicates the strong decrease from crisis months (August 2008 – April 2009) and the last one is constant decrease by 20% every month. We are interested in changes of regulatory capital amount. As this is only modelling example we change only value of GDP and consequently PD while all other parameters remain unchanged. Let us set the amount of loans (EAD) to € 1 mil. What is the amount of capital requirement under the three scenarios?

The table below illustrates development of GDP, development of PD's based on GDP values and amount of capital required by regulator in order to have adequate capital ratio.

[5] Parameter R is defined by regulator [3] and for mortgage loans is set to R=0.15, LGD = 45% and Φ is the cumulative distribution function of standard normal distribution.

Tab. 6.1.: Scenarios for development of GDP, estimated PD and impact on capital requirement

Time	Real growth of GDP (%)			Probability of Default (%)			Capital Requirement (mil. €)		
	Scen. 1	Scen. 2	Scen. 3	Scen. 1	Scen. 2	Scen. 3	Scen. 1	Scen. 2	Scen. 3
5-16	3.59	6.17	3.97	4.93	4.25	4.82	0.1248	0.1154	0.1234
6-16	3.62	4.51	3.17	4.92	4.67	5.4	0.1246	0.1213	0.1262
7-16	3.54	2.85	2.54	4.93	5.13	5.22	0.1248	0.1273	0.1285
8-16	3.46	1.2	2.3	4.94	5.63	5.37	0.125	0.1336	0.1304
9-16	3.39	-1.17	1.62	4.96	6.44	5.49	0.1252	0.1427	0.1318
10-16	3.35	-3.54	1.3	4.96	7.37	5.58	0.1252	0.152	0.133
11-16	3.32	-5.91	1.4	4.97	8.43	5.66	0.1253	0.1614	0.1339
12-16	3.28	-5.97	0.83	4.97	8.45	5.72	0.1253	0.1615	0.1346

As an illustration we compare the capital requirements in December 2016. Firstly we compare scenario 1 with scenario 2. Real growth of GDP decreases by 9.25%, PD increases by 3.48% and the capital requirement increases by nearly 29%. When comparing scenario 1 with 3 we obtain decrease in growth of GDP by 2.45%, increase of PD by 0.75% and the capital requirement increases by 7.4%. To sum up capital requirement is lowest in the first scenario with € 125.3 K, the highest in the "deep crisis" scenario € 161.5 K and € 134.6 K in the last scenario.

We showed how the model is applied on real life example for calculation of regulatory required capital reserve under different macroeconomic conditions. The next steps could be estimation of the model for other products in the retail portfolio and then on corporate portfolio. On the portfolio of mortgage loans the impact of the macroeconomic variable is rather mild. This might not be the case on different portfolios, especially on corporate portfolio. As the estimation of the model is quite demanding - input data set contains all observations for all clients - it might be interesting to compare the results with i.e. Merton model (subsection 2.1.5.) where the data are firstly grouped by months only afterwards is the model estimated.

Conclusion

We have defined the legislative framework for the assessment of credit risk and stress testing as its integral part. We introduced several methodologies which are commonly used as the modelling techniques for stress testing. Our proposed model is based on survival analysis and we chose the Cox proportional hazard model as the final modelling technique (3.7.). The most attractive property of the selected approach is that we can introduce the macroeconomic variables as time – varying and not as the fixed value during the observation period. Moreover this methodology treats also for the different beginnings of the single clients in the input data set and adjust for their individual time duration influences.

Most commonly used macroeconomic variables (Tab 4.2.) were tested for their influence on default rate on mortgage loans portfolio. As the final candidates we chose three models: model with real growth of GDP (eq. 4.1. and 4.2.), model with macroeconomic variable Import (eq. 4.3.) and the model with combination of Import and Unemployment rate (eq. 4.4.). We compared the fit of all three models and based on graphical analysis the model with GDP was chosen as the final (Fig.4.4.). Afterwards we tested the adequacy of our proposed model in Chapter 5. As a result of these analyses we conclude that the model is estimated properly, nevertheless we must be aware of two aspects. If the input variable Rating is not appropriately assigned the predicted probability of default is also misleading. Secondly, the estimated coefficient for real growth of GDP is likely to change over time, therefore regular re-estimation of the model is necessary in order to obtain realistic predictions. However, this is common routine hence we consider it as an expected result.

Finally, we have applied our proposed model on real life example and estimated the regulatory capital requirement under three scenarios of development of real growth of GDP. The first scenario is the forecast until the end of the year 2016. We have shown that indeed impact of macroeconomic changes is observable in portfolio of mortgage loans and we have quantified this impact. As a result of this work we propose the model defined by equations 4.1. and 4.2., which we consider as suitable for stress testing.

Except for the model proposal this work provides the mathematical background for the Cox proportional hazard model, commonly used estimation techniques, the implementation procedure in SAS guide environment and the extensive background for the check of the model adequacy. The room for further research is in estimating the model on different portfolios and examining whether there is difference in the impact of the macroeconomic changes, or using different technique (i.e. Merton model 2.1.5.) and comparison of the obtained results.

References

[1] ANDERSON, T.W. – GOODMAN, L.A. 1957. Statistical inference about Markov Chains. *Annals of Mathematical Statistics* [online]. 1957, vol. 28, no. 1, [ref. 2016-04-16]. Available on internet: < http://projecteuclid.org/download/pdf_1/euclid.aoms/1177707039>

[2] BANASIK, J. – CROOK J. 2005. Explaining Aggregate Consumer Delinquency. University of Edinburgh. . [ref. 2016-04-16]. Available on internet: <https://www.researchgate.net/publication/215991102_Explaining_Aggregate_Consumer_Delinquency_Behaviour_over_Time>

[3] BASEL II: 2006 International Convergence of Capital Measurement and Capital Standards: A revised Framework – Comprehensive Version. [ref.: 2016-04-16] Available on internet: <http://www.bis.org/publ/bcbs128.pdf >

[4] BASEL III: 2010 A Global Gegulatory Framework for More Resilient Banks and Banking Systems. [ref.: 2016-04-16] Available on internet: <http://www.bis.org/publ/bcbs189.pdf>

[5] BELLOTTI T. – CROOK, J. 2008. Credit scoring with macroeconomic variables using survival analysis. In *Journal of the Operational Research Society* [online]. 2008, vol. 60, no. 12, [ref. 2016-04-16]. Available on internet: <http://www.research.ed.ac.uk/portal/files/8628798/Credit_scoring_with_macroeconomic_variables_v10_4_1.pdf>

[6] BORIO, C. – FURFINE, C. – LOWE, P. 2001. Procyclicality of the financial system and financial stability: issues and policy options. Technical Report No. 1. *Bank for International Settlements*. [Ref. 2016-04-16]. Available on internet: <http://www.bis.org/publ/bppdf/bispap01a.pdf **>**

[7] COLLET, D. 2003. *Modelling Survival Data in Medical Research.* [online]. London: Chapman and Hall/Crc, 2003. 391 p. [ref. 2016.04.16.] Available on internet: <http://evunix.uevora.pt/~pinfante/eb1011/Collett%20David%20-%20Modelling%20survival%20data%20in%20medical%20research%20(Chapman&Hall-CRC,%202004)(ISBN%201584883251).pdf >. ISBN 1-58488-325-1.

[8] COX, D.R. – OAKES, D. 1984. Analysis of Survival Data. London: Chapman and Hall, 1984. p. 201 ISBN 041224490X

[9] CROWLEY, J. – HU, M. 1997. Covariance Analysis of Heart Transplant Data. In *Journal of the American Statistical Association* [online]. 1977, vol. 72, [ref. 2016-04-16] Available on internet: <http://amosdevelopment.com/webhelp/crowley__hu_1977.htm>

[10] DIRECTIVE 2013/36/EU on access to the activity of credit institutions and the prudential supervision of credit institutions and investment firms. [ref.: 2016-04-16] Available on internet: <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32013L0036&from=EN>

[11] FORSTER, J.J. – SUDJIANTO, A. 2013. Modelling time and vintage variability in retail credit portfolio: The decomposition approach. *Lloyds Banking Group* [online]. 2013 [ref. 2016-04-16]. Available on internet: < http://arxiv.org/pdf/1305.2815.pdf >

[12] HOSMER, D. W. – LEMESHOW, S. 1999. Applied Survival Analysis: Regression Modeling of Time-to-Event Data. Second Edition. NewYork: Wiley, 2008. p. 416. ISBN 978-0-471-75499-2

[13] IBFed response to Basel Committee consultation on a revised standardised approach to credit risk: 2015. [ref. 2016-04-16] Available on internet: <www.ibfed.org/ibfed/download/9830>

[14] JURČA, P. Lectures: Quantitativ Methods in Risk Management. Spring term 2015/2016. [ref. 2016-04-16] Available on internet: <http://www.iam.fmph.uniba.sk/institute/jurca/qrm/Chapter5.pdf>

[15] Principles for the Management of Credit Risk: 2000. Basel Committee on Banking Supervision. [ref.: 2016-04-16] Available on internet: <http://www.bis.org/publ/bcbs75.pdf>

[16] REGULATION (EU) No 575/2013 on prudential requirements for credit institutions and investment firms. [ref.: 2016-04-16] Available on internet: <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32013R0575&from=EN>

[17] Revision to the Standardised Approach for credit risk. Consultative Document: 2014. Basel Committee on Banking Supervision. [ref.: 2016-04-16] Available on internet: <http://www.bis.org/bcbs/publ/d307.pdf>

[18] RODRIGUEZ, G. 2005. Non-Parametric Estimation in Survival Models. Princeton University. [ref. 2016-04-16] Available on internet: <http://data.princeton.edu/pop509/NonParametricSurvival.pdf>

[19] ROYSTON, P. 2011. Estimating a smooth baseline hazard function for the Cox model. Research report No. 314. *London: University College London, Department of Statistical Science*. [ref. 2016-04-16]. Available on internet: <https://www.ucl.ac.uk/statistics/research/pdfs/rr314.pdf>

[20] RÖSCH, D. - SCHEULE, H. 2008. Stress Testing for Financial Institutions. London: Risk Books, 2008. p. 457. ISBN 978-1-906348-11-3

[21] SAS Institute Inc., SAS/STAT User’s Guide, Version 8, Cary, NC:SAS Institute Inc., 1999. [ref. 2016-04-16] Available on internet: <http://www.okstate.edu/sas/v8/saspdf/stat/chap49.pdf>

[22] THERNEAU, T.M. – GRAMBSCH, P.M. – FLEMING T.R. 1990. Martingale-Based Residuals for Survival Models. In *Biometrika* [online]. 1990, vol. 77, no.1 [ref. 2016-04-16] Available on internet: <http://www.escarela.com/archivo/anahuac/03o/residuals.pdf>

[23] THERNEAU, T.M. – GRAMBSCH, P.M. 1994. Proportional Hazards Test and Diagnostics Based on Weighted Residuals. In *Biometrika* [online]. 1994, vol. 81, no 3 [ref. 2016-04-16] Available on internet:
<http://www.escarela.com/archivo/anahuac/03o/residuals.pdf>

[24] THOMAS, L.C. – STEPANOVA, M. 2001. Phab scores: proportional hazards analysis behavioural scores. In *Journal of the Operational Research Society* [online]. 2001, vol 52, no. 9 [ref. 2016-04-16] Available on internet:
<http://www.jstor.org/stable/822780?seq=1#page_scan_tab_contents>

Appendix

Tab. 7.1.: Estimated model for Rating_YA, GDP_rlgrth and the variable of interactions: Rating_YA x GDP_rlgrth

Parameter	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
Rating_YA	0.6632	0.00484	18761.2726	<.0001	1.941	1.923	1.96
GDP_rlgrth	-0.07315	0.0073	100.3764	<.0001	0.929	0.916	0.943
Inter_RatG	0.00307	0.0012	6.04	0.0105	1.003	1.001	1.005

Tab. 7.2.: Estimated coefficients of variables for model with one macroeconomic variable

Variable	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
YY_GDP_PPY	-0.07863	0.00563	195.1303	<.0001	0.924	0.914	0.935
GDP_rlgrth	**-0.05735**	**0.00397**	**208.2143**	**<.0001**	**0.944**	**0.937**	**0.952**
Unempl_rate	-0.04152	0.00985	17.7699	<.0001	0.959	0.941	0.978
HC_yy	-0.03466	0.00876	15.6579	<.0001	0.966	0.949	0.983
HICP_index	-0.02014	0.00323	38.8197	<.0001	0.98	0.974	0.986
IES_1M	-0.01795	0.00149	144.3557	<.0001	0.982	0.979	0.985
HICP_MrCH	-0.01483	0.03479	0.1818	0.6698	0.985	0.92	1.055
Export_YY	-0.00881	0.0008141	117.0161	<.0001	0.991	0.99	0.993
Import_YY	**-0.00951**	**0.000738**	**166.1652**	**<.0001**	**0.991**	**0.989**	**0.992**
HICP_ArCH	-0.00841	0.00698	1.92	0.2287	0.992	0.978	1.005
Rl_wage_grth	-0.00738	0.00614	1.17	0.2299	0.993	0.981	1.005
QQ_GDP_PPY	-0.00428	0.00206	4.95	0.0372	0.996	0.992	1
GDP_Nom	-0.0000839	0.00000927057	81.8732	<.0001	1	1	1
GDP_PPY	-0.0000753	0.00000920409	66.9623	<.0001	1	1	1
GDP_chain	-0.0000962	0.0000107	80.4764	<.0001	1	1	1
Hous_cons	-0.0002194	0.0000279	61.7367	<.0001	1	1	1
HC_chain	-0.0002932	0.0000594	24.3404	<.0001	1	1	1
Trade_Bal	0.0000699	0.000059	1.46	0.236	1	1	1
Export	-0.0001034	0.0000135	58.4977	<.0001	1	1	1
TB_YY	0.0000852	0.0000215	15.6463	<.0001	1	1	1
TB_QQ	0.00000730393	0.00000498163	2.1497	0.1426	1	1	1
Import	-0.0001405	0.0000155	81.8101	<.0001	1	1	1
Import_QQ	0.0003486	0.0008156	0.1827	0.669	1	0.999	1.002
Export_QQ	0.00111	0.000693	2.46	0.1079	1.001	1	1.002
Infl_CPI	0.00748	0.00741	1.0193	0.3127	1.008	0.993	1.022
EUR_USD	1.12901	0.13124	74.0033	<.0001	3.093	2.391	4

Tab. 7.3.: Statistic -2 log(L) and AIC criterion for model with one macroeconomic variable

Variable	-2 LOG L	AIC
Import_YY	203785.65	203789.65
Export_YY	203836.75	203840.75
TB_YY	203939.78	203943.78
Rl_wage_grth	223387.91	223391.91
Export_QQ	223878.57	223882.57
TB_QQ	223879.03	223883.03
Import_QQ	223880.98	223884.98
GDP_rlgrth	228245.72	228249.72
YY_GDP_PPY	228260.06	228264.06
IES_1M	228312.82	228316.82
GDP_nom	228379.29	228375.29
Import	228375.77	228379.77
GDP_chain	228376.93	228380.93
EUR_USD	228381.76	228385.76
GDP_PPY	228390.73	228394.73
Hous_cons	228396.2	228400.2
Export	228399.87	228403.87
HICP_index	228420.15	228424.15
HC_chain	228435	228439
Unempl_rate	228441.83	228445.83
HC_yy	228443.72	228447.72
QQ_GDP_PPY	228455.21	228459.21
HICP_ArCH	228458.09	228462.09
Trade_Bal	228458.13	228462.13
Infl_CPI	228458.52	228462.52
HICP_MrCH	228459.35	228463.35

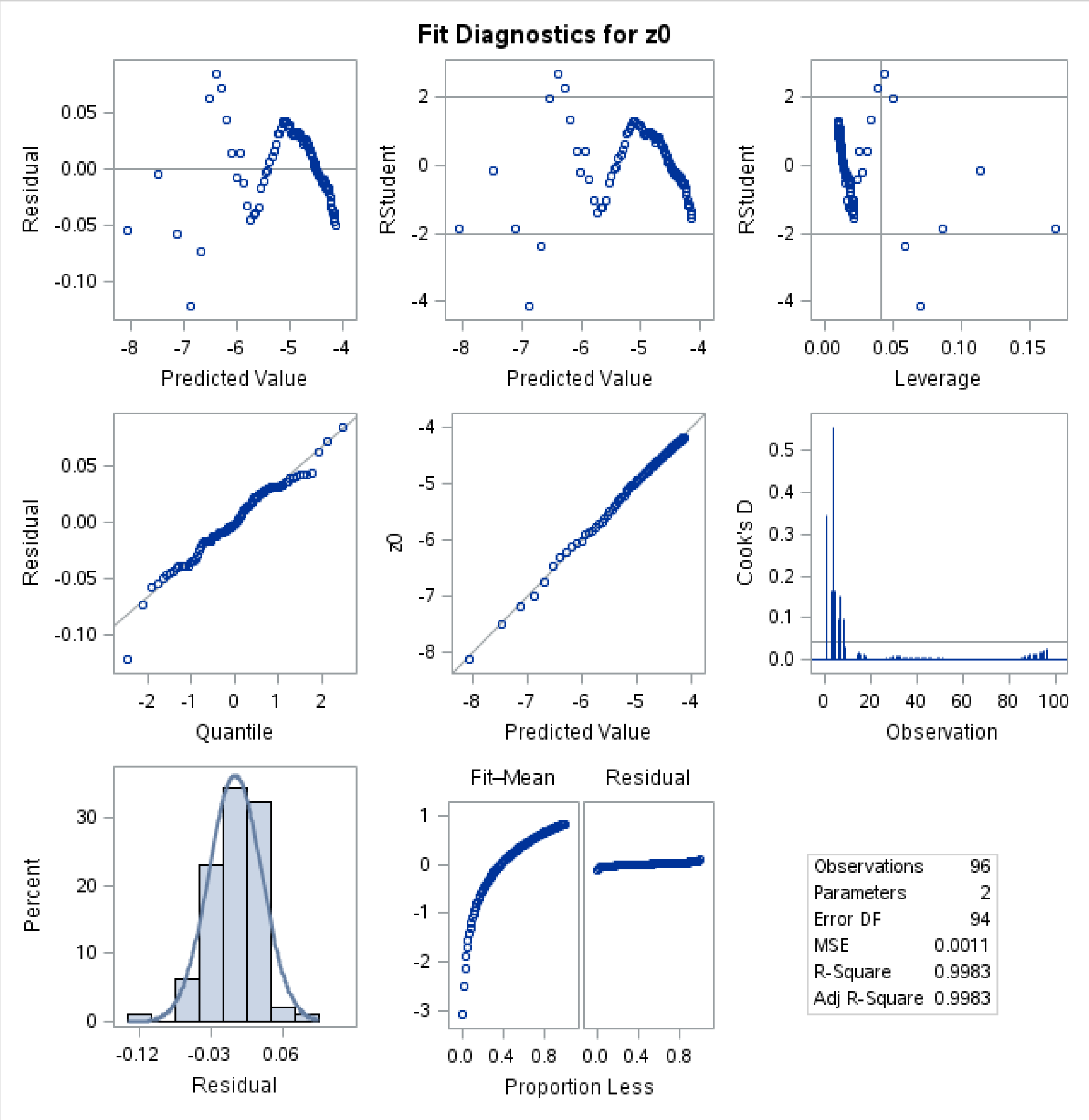
Tab. 7.4.: Estimated coefficients of variables for model with two macroeconomic variables

Parameter	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
Unempl_rate	0.05374	0.01271	17.8636	<.0001	1.055	1.029	1.082
Import_YY	-0.01166	0.0008996	168.0209	<.0001	0.988	0.987	0.99
Unempl_rate	0.05664	0.0136	17.3462	<.0001	1.058	1.3	1.087
Export_YY	-0.01168	0.00107	119.1993	<.0001	0.988	0.986	0.99
Unempl_rate	-0.0924	0.01205	58.7915	<.0001	0.912	0.89	0.934
HC_yy	-0.07994	0.01066	56.2033	<.0001	0.923	0.904	0.943
Unempl_rate	-0.03509	0.01003	12.15	0.0005	0.966	0.947	0.985
HICP_index	-0.01879	0.00325	33.3897	<.0001	0.981	0.975	0.988
HC_yy	-0.04477	0.0091	24.2237	<.0001	0.956	0.939	0.973
Import_YY	-0.00966	0.0007396	170.4763	<.0001	0.99	0.989	0.992
GDP_rlgrth	-0.06218	0.00759	67.0451	<.0001	0.94	0.926	0.954
IES_1M	0.00213	0.00286	0.5562	0.4558	1.002	0.997	1.008
GDP_rlgrth	-0.05929	0.00428	191.7186	<.0001	0.942	0.935	0.95
Unempl_rate	0.01344	0.011	1.27	0.2218	1.014	0.992	1.036
GDP_rlgrth	-0.05757	0.00415	192.8627	<.0001	0.944	0.936	0.952
HC_yy	0.00173	0.00924	0.035	0.8516	1.002	0.984	1.2
GDP_rlgrth	-0.05637	0.00432	169.9875	<.0001	0.945	0.937	0.953
HICP_index	-0.00202	0.00352	0.3301	0.5656	0.998	0.991	1.005
HC_yy	-0.02325	0.00916	6.76	0.0111	0.977	0.96	0.995
HICP_index	-0.0181	0.00333	29.6144	<.0001	0.982	0.976	0.988
Unempl_rate	-0.00421	0.01063	0.157	0.6919	0.996	0.975	1.017
IES_1M	-0.01775	0.00158	126.6616	<.0001	0.982	0.979	0.985
HC_yy	-0.00654	0.00925	0.501	0.4791	0.993	0.976	1.012
IES_1M	-0.01766	0.00155	129.4526	<.0001	0.982	0.98	0.985
HICP_index	-0.00787	0.00341	5.325	0.021	0.992	0.986	0.999
IES_1M	-0.01675	0.00158	111.6955	<.0001	0.983	0.98	0.986
GDP_rlgrth	-0.04905	0.00762	41.4857	<.0001	0.952	0.938	0.966
Import_YY	-0.00172	0.00141	1.59	0.2229	0.998	0.996	1.001
GDP_rlgrth	-0.0605	0.00634	90.979	<.0001	0.941	0.93	0.953
Export_YY	0.0009366	0.00131	0.5142	0.4733	1.001	0.998	1.004
HICP_index	-0.00863	0.00335	6.52	0.0099	0.991	0.985	0.998
Import_YY	-0.00902	0.0007625	139.8396	<.0001	0.991	0.99	0.993
IES_1M	-0.0057	0.00248	5.67	0.0219	0.994	0.989	0.999
Import_YY	-0.00728	0.00122	35.3179	<.0001	0.993	0.99	0.995

Tab. 7.5: Statistic -2 log(L) and AIC criterion for model with two macroeconomic variables

Parameter	-2 LOG L	AIC
GDP & Imp_yy	203743.94	203749.94
GDP & Exp_yy	203744.91	203750.91
HC & Imp	203761.04	203767.04
Unempl & Imp_yy	203767.45	203773.45
HICP & Imp	203778.97	203784.97
IES & Imp_yy	203780.36	203786.36
Unempl & Exp_yy	203819.06	203825.06
GDP_rlgrth & Unempl	228244.22	228250.22
GDP_rlgrth & IES_1M	228245.17	228251.17
GDPrlgrth_HICP index	228245.39	228251.39
GDP_rlgrth & HC_yy	228245.69	228251.69
HICP_IES	228307.46	228313.46
HC_yy & IES	228312.31	228318.31
Unempl & IES	228312.66	228318.66
Unempl & HC_yy	228384.36	228390.36
Unempl & HICP	228407.97	228413.97
HC_yy & HICP	228413.69	228419.69

Fig. 7.6.: Residuals for estimation of $\widehat{\gamma}_0$ and $\widehat{\gamma}_1$ for the model with GDP_rlgrth



Observations	96
Parameters	2
Error DF	94
MSE	0.0011
R-Square	0.9983
Adj R-Square	0.9983

Fig. 7.7.: Residuals for estimation of $\hat{\gamma}_0$ and $\hat{\gamma}_1$ for the model with Import_yy

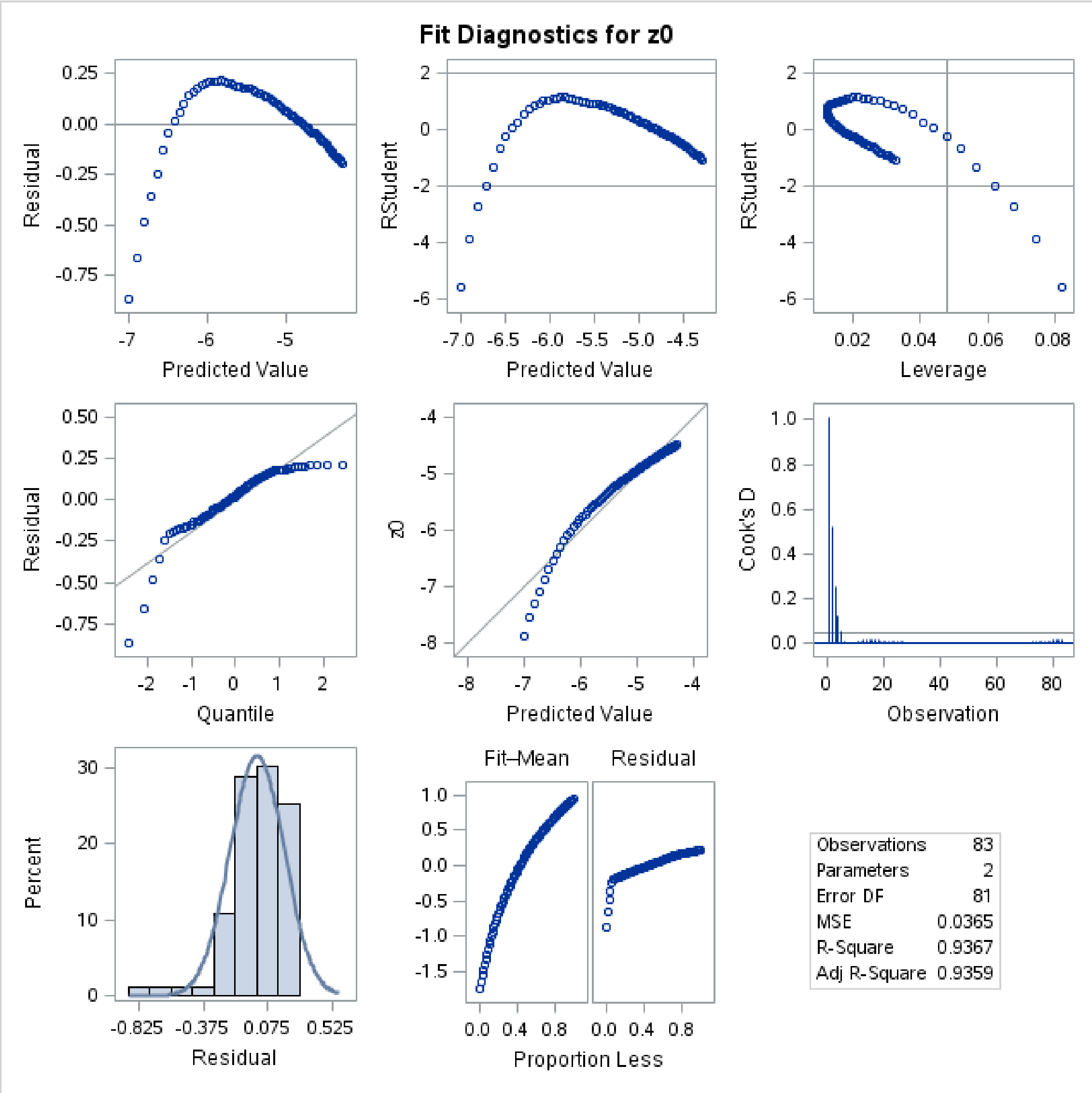


Fig. 7.8.: Residuals for estimation of $\hat{\gamma}_0$ and $\hat{\gamma}_1$ for the model with Import_yy and Unempl_rate

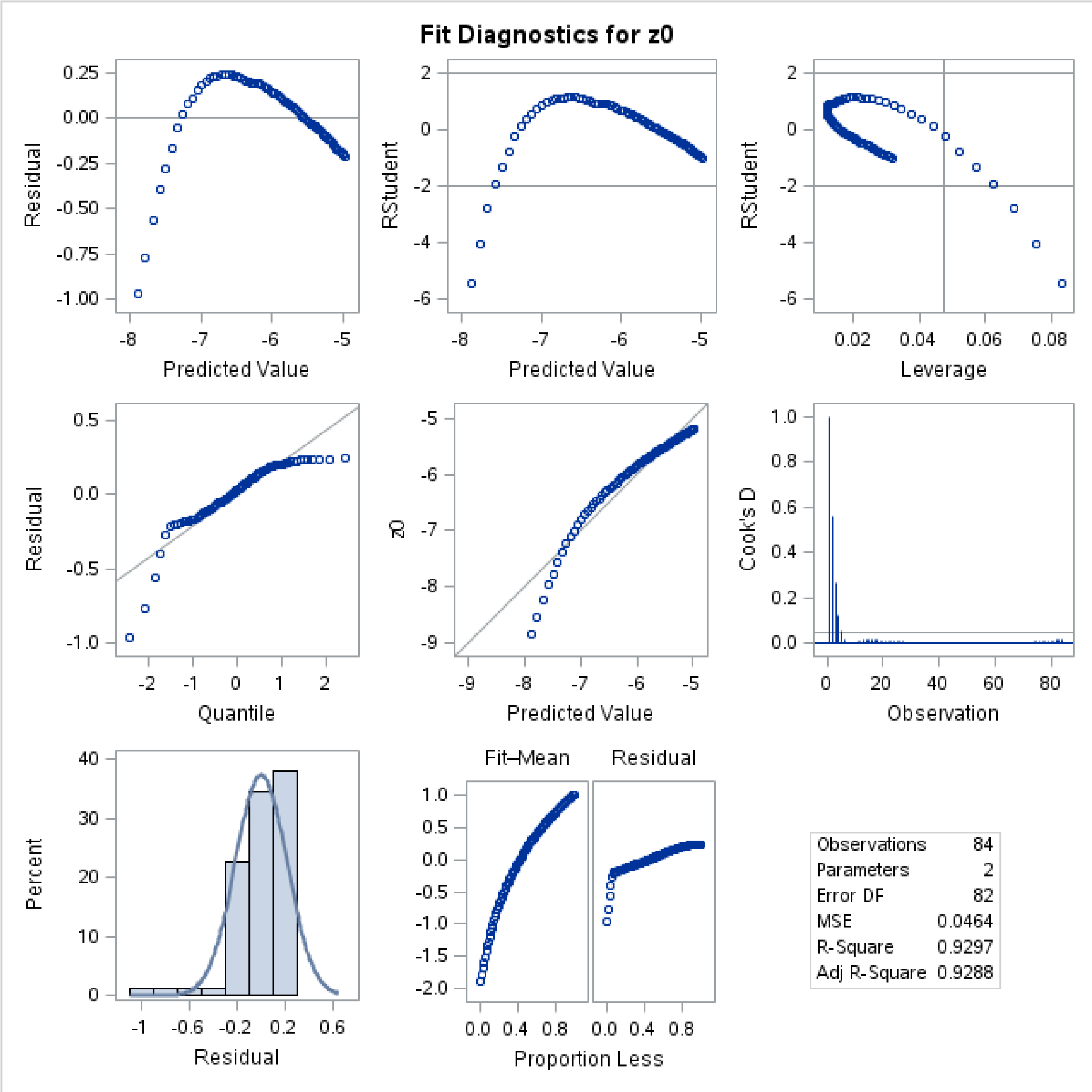


Fig. 7.9.: Histogram of Martingale residuals (MR)