Erasmus University Rotterdam
School of Economics – Department of Econometrics
Master Thesis in Quantitative Finance

Maturity of Savings Deposits

B.X.J.S. Klijn – 348953

June 30, 2014

Abstract

Liquidity management, the management of funds in order to have sufficient short term assets to pay off short term debts, is an important aspect of risk management. According to Frauendorfer and Schürle (2007), duration matching is a widely used technique to manage liquidity risk. However, duration matching is difficult for savings deposits, because they have no fixed maturity. Hence, the goal of this research is to determine the maturity of savings deposits, so that duration matching can be applied. Instead of a single maturity, I divide the total deposit volume into several parts, or buckets, with their own maturity. This research uses future deposit volume simulations in combination with Value-at-Risk and the liquidity constraint of Bardenhewer (2007) to create the buckets. The final model is evaluated by comparing it with a much simpler model, the random walk. I conclude that the final model only provides additional value for products with an attractive client rate. In contrast, for less attractive products both models result in nearly identical volume simulations and buckets.

Keywords: Maturity; savings deposits; non-maturing liabilities; duration matching

Academic supervisor:	Dr. H.J.W.G. Kole
Co-reader:	Dr. M. Jaskowski
External supervisor:	P. Klompe

Contents

Notation

This section lists the abbreviations and variables used in this research. The abbreviations are accompanied with their full name and the variables with a short description. Also the page on which they are mentioned for the first time is displayed.

Abbreviations

Notation	Full name	Page
AR	Auto Regressive	10
CIR	Cox Ingersoll Ross	11
DNB	De Nederlandsche Bank	6
HJM	Heath Jarrow Morton	11
ILAAP	Internal Liquidity Adequacy Assessment Process	6
LogL	Log Likelihood	26
OLS	Ordinary Least Squares	26
SIC	Schwarz Information Criterion	14
VAR	Vector Auto Regressive	36
VaR	Value-at-Risk	7

Variables

Variable	Description	Page
D	Deposits	16
D_{jan}	Dummy for January	42
D_{may}	Dummy for May	42
D_{mon}	Dummy for Mondays	30
D_{na}	Dummy for when deposits/withdrawals not available	26
D_{ol}	Dummy for outlier (February 23, 2010)	28
D_{rel}	Relative deposits	16
D_{tr}	Transformed deposits	16
Δr_{nw}	Log difference of nationwide client rate	40
R_{cl}	Client rate	19
R_{nw}	Nationwide client rate	19
S	Spread	19
S_{rel}	Relative spread	19
$S_{rel,ba}$	Relative spread before alteration	24

Variable	Description	Page
$S_{rel,dr}$	Relative spread drop	24
W	Withdrawals	46
W_{rel}	Relative withdrawals	45
W_{tr}	Transformed withdrawals	18
Δv_{nw}	Log difference of nationwide deposit volume	26
V	Deposit volume	16
Δy_{3m}	Log difference of 3-month Euribor rate	40
Y_{3m}	3-month Euribor rate	40

1 Introduction

Liquidity management, the management of funds in order to have sufficient short term assets to pay off short term debts, is an important aspect of risk management. Without proper liquidity management, even successful companies that make profits in the long run, can go bankrupt due to insufficient liquidity in the short run. This also applies to banks and other financially orientated companies. To ensure adequate liquidity management at these companies, De Nederlandsche Bank (DNB (2012)) states rules and guidelines in their Internal Liquidity Adequacy Assessment Process (ILAAP) manual.

According to Frauendorfer and Schürle (2007), duration matching is a widely used technique to manage liquidity risk. This technique is also used by Allianz Nederland Asset Management, which is where this research is conducted. Duration matching aims to match the magnitude and timing of incoming and outgoing cash flows, for example, by investing the companies' assets in funds that have the same maturity as the companies' liabilities.

However, duration matching is difficult for savings deposits, because they have no fixed maturity. In the literature and from a bank's perspective, savings deposits are therefore also called non-maturing liabilities. It is unclear for the bank when it needs to repay the clients' money, because clients have the right to withdraw any amount at any time. Technically, savings deposits have a zero contractual maturity (DeWachter, Lyrio, and Maes (2006)), and in theory it is possible that all funds disappear overnight. Fortunately for the banks, due to the number of clients the vast majority of individual decisions cancels out each other. Consequently, the overnight disappearance of all savings deposits is not very likely, but the actual maturity profile is still unclear.

Hence, the goal of this research is to determine the maturity of savings deposits, so that duration matching can be applied in order to minimize liquidity risk and maximize profit. However, minimal risk and maximal profit are conflicting and cannot be achieved at the same time. To prevent any liquidity issues, banks could invest the clients' savings in products with a very short maturity. However, these products are typically less profitable than products with longer maturities and therefore less appealing. Vice versa, the more profitable products with long maturities entail more liquidity risk. The optimal solution, a moderate profit against a reasonable risk, can only be achieved by dividing the total deposit volume into several parts, also called buckets. These buckets are thereafter invested in products with a corresponding maturity.

The literature contains several options to divide the total volume in buckets, namely the non-maturation theory, the replicating portfolio approach, and future deposit volume simulations in combination with statistical measures. The third method is preferred, because it has a close connection to the volume and the future simulations make the method robust and compliant with ILAAP. Furthermore, the magnitude and maturity of the buckets can easily be determined with several measures, such as Value-at-Risk (VaR) and the liquidity constraint of Bardenhewer (2007). The other two methods are not sufficient. The non-maturation theory divides the volume in only a few buckets and is not able to specify the maturity at the same time. The replicating portfolio approach fulfils the shortcomings of the former, but it has little connection to the deposit volume.

In order to be able to create buckets with the third method, realistic future scenarios must be created. In this research, the future volume is created indirectly with separate models for the deposits and withdrawals. This enables us to observe the effect of important variables and events on deposits and withdrawals individually, instead of the net effect on the volume. The relative spread, which is a measure for the attractiveness of the deposit product, is one of the important variables. This measure takes into account the difference in the client rate of banks and also the state of the economy. The client rate is the interest rate offered to the client by the bank. Other important variables are related to the nationwide deposit volume and a client rate change.

For each of the important auxiliary variables, realistic simulations must be created as well. Based on the work of Diebold and Li (2006), the yield curve is simulated in order to create the nationwide client rate. In turn, that is used to create the nationwide deposit volume. The Allianz client rate model has a special purpose; it is possible to adjust several parameters to obtain the desired level of client rate attractiveness. Then, all the individual models are put together, also referred to as the framework, and they are used to create the deposit volume simulations. These simulations are used to construct buckets with VaR and the liquidity constraint of Bardenhewer (2007).

Due to the different construction of the VaR and liquidity constraint buckets, they have different properties. The liquidity constraint buckets are stricter, because multiple time periods are evaluated, instead of a single one for the VaR buckets. The cumulative weight of buckets up to a certain maturity are always lower than those of the VaR method. Also, due to the observation of multiple periods, the liquidity constraint buckets are more consistent across different samples. The advantage of the VaR buckets is that the effect of a specific client rate scenario can be evaluated. This can potentially result in much more profit, but it is also more labour intensive and risky.

The final model of this research is evaluated by comparing it with an alternative and much simpler model, the random walk. Given the differences between the final model and the random walk model for two separate products, I conclude that the additional value of my model depends on the nature of the product. For products that are marginally dependent of the relative spread, my model has no clear additional value. For these products, the simulations and buckets are almost identical to those of the random walk. For products that are more attractive and do depend on the relative spread, my model is advantageous. My model takes into account the attractiveness of the product, while the random walk does not take into account anything. Despite this fact, the buckets of my model and the random walk appear similar. However, this is just due to the lack of a clear client rate scenario. The properties of the client rate model in the evaluation are set such that they are similar to the historical sample. However, these are generic and the regime can therefore not be regarded as attractive or unattractive. The differences are much larger, if the random walk buckets are compared with a clearly attractive client rate, which confirms the additional value of the model.

2 Literature

This section gives an overview of several methods that are used to divide the total volume in buckets. Namely, the non-maturation theory, the replicating portfolio approach, and future deposit volume simulations in combination with statistical measures.

2.1 Non-maturation theory

The non-maturation theory normally divides the total deposit volume into at least two parts; a stable core part and a more volatile part. The core part is assigned a long maturity, while the volatile part gets a shorter maturity. The distinction between being core or volatile can be made in a variety of ways. Oesterreichische Nationalbank (OeNB (2008)) assigns a volume of two standard deviations below the mean volume to the core part. What remains is the volatile portion. Optionally, a volume of two standard deviations above the mean volume can be characterized as extremely volatile, which is then invested in very short maturities.

This method is easy to apply, but there is a drawback. The core deposits become negative if the standard deviation is very large relative to the mean. Consequently, this implies that all funds are volatile, which is equivalent to assuming just a single maturity and that is insufficient. Furthermore, under normal circumstances, this method divides the deposit volume into just two parts. Ideally, the volume should be divided into more parts, so that we can use many maturities. Additionally, the non-maturation theory does not give insight in the exact maturities that should be used for the buckets. Nevertheless, according to Kalkbrener and Willing (2004), this approach is used by many banks.

2.2 Replicating portfolio

Another approach is the replicating portfolio, which was first used by Wilson (1994) for non-maturing products. The idea behind this method is that products with equivalent yields should have the same cash flows, or at least approximately. The replicating portfolio is formed from several interest rate paying instruments, with the purpose to mimic the client rate as closely as possible, except for a certain margin. This margin represents the profit of the bank, so ideally it should be positive and not too volatile. The weight of each interest rate paying instrument is determined with an optimization criterion. Often, the criterion is either to minimize the standard deviation of the margin or to maximize the Sharpe ratio of the margin. The Sharpe ratio is defined as the mean of the margin, divided by the standard deviation of the margin. The sum of the weights must be equal to one, and often they are also required to be greater than zero individually. Finally, the total deposit volume is invested in the maturities according to the weights of the replicating portfolio. Potentially, this method can be combined with the non-maturation theory. For example, Maes and Timmermans (2005) invest only a part of the deposits in the replicating portfolio, instead of the total deposit volume.

It should be noted that the replicating portfolio method determines the weights solely by the client rate and that the deposit volume plays no role. Therefore, this method seems not too promising, although it is often used by large international banks, according to Wielhouwer (2003), DeWachter et al. (2006) and Frauendorfer and Schürle (2007).

Bardenhewer (2007) extends the general replicating portfolio approach with a liquidity constraint and moving averages of the interest rates instead of the rate of a single point in time. Of these extensions the liquidity constraint is the most promising, because it provides a connection between the deposit volume and the weights of the replicating portfolio. It determines the maximum percentage decline of the total deposit volume for each maturity, which is used as a lower bound for the weight of the corresponding maturity in the replicating portfolio. This makes sure that there is always enough liquidity, at least, for the examined historical period.

So far, I have only considered the static replicating portfolio. The static replicating portfolio assumes that the weight vector is constant for a long period of time and it is determined over a single realisation, namely the history. Frauendorfer and Schürle (2007) are critical on these assumptions and propose a dynamic replicating portfolio, which estimates the weight vector more frequently. Furthermore, instead of the weight vector being estimated over a single historical realisation, it is estimated using multiple simulated future scenarios. These scenarios usually include simulations for the deposit volume, the client rate and the yield curve. Because the weights of the dynamic method are estimated over a range of possible future scenarios it is more robust than the static approach. Frauendorfer and Schürle (2003) use a multistage stochastic programming model that determines a dynamic replicating portfolio. They show that this method results in a more stable and higher margin than with the static replicating portfolio.

2.3 Framework

The development of non-maturing liabilities is often not considered separately, but mostly in the context of entire valuation frameworks, for example in Hutchison and Pennacchi (1996), Jarrow and Van Deventer (1998), and O'Brien (2000). Ellis and Jordan (2001) state that "*any method that is adopted for valuing NMDs* [Non-Maturity Deposits] *should not be viewed in isolation, but rather in the context of an overall asset and liability management framework.*" These frameworks consist of a valuation method and models for the yield curve, client rate and the deposit volume. The various different valuation models will not be explored, as that is not the primary goal of this research. However, I do need models for the yield curve and the client rate, because they affect the deposit volume. The concept of a framework is promising, because it provides future deposit volume simulations, that can be analysed in a variety of ways. For example with the dynamic replicating portfolio method, but also with the liquidity constraint of Bardenhewer (2007) or the more traditional VaR. In addition, a framework of simulation models is in accordance with the ILAAP guidelines of DNB (2012), because it takes into account future cash flows.

2.3.1 Volume

According to Ellis and Jordan (2001), the volume model should incorporate "*the deposit rate, the Treasury rate and other factors such as lagged deposit balances.*" Especially the autoregressive (AR) term, the lagged deposit balance, is mentioned throughout the entire literature. This is not surprising since the deposit volume of tomorrow is probably close to today's volume. Paraschiv and Schürle (2010) combine an AR(1) model with a variable that accounts for the spread between the market and client rates. Kalkbrener and Willing (2004) mention that it might be interesting to include macro economic variables, but at the same time note that "*forecasting macro economic developments over longer periods is not a trivial task.*" The latter clearly limits the number of possible variables to include. Several studies find a relation between the deposit volume and the market or client rate. Surprisingly, Kalkbrener and Willing (2004) note that the relation between the market rates and the deposit volume is not particularly high in the German market.

2.3.2 Yield curve

The yield curve, frequently referred to as market rate in the literature, includes the stochastic component in most studies. Kalkbrener and Willing (2004) are the exception because they include an additional stochastic component in the volume model. The frameworks use the yield curve to value the cash flows. Although this research does not focus on the valuation, the market rate is also the driving factor for the (nationwide) client rate. For example, when the European Central Bank lowers the market rates, banks are likely to follow and not vice versa.

The foundations for yield curve models are laid by Vasicek (1977), Cox, Ingersoll, and Ross (1985) (CIR) and Heath, Jarrow, and Morton (1992) (HJM). There is no clear favourite in the context of non-maturing deposits. Hutchison and Pennacchi (1996) and Nyström (2008) use a one-factor Vasicek model; O'Brien (2000) models the market rates with a CIR model; Jarrow and Van Deventer (1998) incorporate the HJM model; and Kalkbrener and Willing (2004) implement "*two classes of two-factor HJM models: two-factor Vasicek models and non-parametric HJM models with piecewise constant volatility functions.*"

A different approach to model the yield curve, which is not mentioned in the non-maturing liabilities literature, is the use of a Nelson-Siegel model after Nelson and Siegel (1987). The advantage of the Nelson-Siegel model is that it usually has three factors compared to the one-factor Vasicek and CIR model. The yield curve can be fitted better with three factors. Additionally, the Nelson-Siegel model incorporates an interpretable structure on the loadings, which can be identified with the level, slope and curvature of the yield curve. Diebold and Li (2006) forecast the factors of their three-factor Nelson-Siegel model with an AR(1) model, so that the future yield curve can be constructed. They state that the short-term forecasts are not better than several competing models, but "*the 1-year-ahead results are much superior*". Diebold and Li (2006) might be too optimistic, because they test on a favourable subsample. De Pooter (2007) tests several variations of the Nelson-Siegel model and he finds that a four-factor Nelson-Siegel model, estimated with a Kalman filter, outperforms most models across different horizons and subsamples.

2.3.3 Client rate

The client rate model must at least have a connection to the yield curve model. When market rates fall, banks quickly adjust the client rate in the same direction in order to prevent a negative spread and losses. On the other hand, when the market rates rise, banks are reluctant to change the client rate as a little delay in the adjustment earns them some extra money. This asymmetric behaviour is called (upward) stickiness in the literature. Despite this fact, Hutchison and Pennacchi (1996) and Jarrow and Van Deventer (1998) assume a symmetric relation between market and client rates. O'Brien (2000) uses an equilibrium model with different adjustment speeds to account for the stickiness phenomenon. Paraschiv and Schürle (2010) develop an error correction model with an AR term, market rates and a variable threshold to allow for asymmetric behaviour. The methods of O'Brien (2000) and Paraschiv and Schürle (2010) are studied in Section 4.3.

2.4 Summary

A suitable technique divides the deposit volume in multiple parts and also specifies the corresponding maturity. Furthermore, it is preferable to use a method that is based on future simulations instead of one historical realisation. Firstly, this complies better with the ILAAP guidelines of DNB (2012), because it takes into account future cash flows. Secondly, the estimation of the magnitude and maturity of the buckets over multiple scenarios is more robust. A robust approach has less risk and lower transaction fees due to smaller changes in the investment positions.

The non-maturation theory is the simplest of the methods mentioned in this section and has the least in common with the ideal method. This approach divides the volume in just two or three parts, and possibly even only one if the deposit volume is volatile. Furthermore, it gives no insight in the maturity of the buckets.

The more complex replicating portfolio might improve upon the former approach as it potentially divides the deposit volume over many maturities. Furthermore, by construction it gives the magnitude and the maturity of the parts. A large drawback is the lack of connection to the deposit volume; the method only depends on the client rate. Although, this can possibly be solved by the liquidity constraint of Bardenhewer (2007).[1] Other deficiencies in the static replicating portfolio, according to Frauendorfer and Schürle (2007), are the assumption that the weight vector is constant for a long period of time and that it is being determined over a single historical realisation. According to the same authors, their dynamic replicating portfolio improves upon the basic method, by estimating the weights more frequently and more robustly over many future simulations.

Besides the dynamic replicating portfolio, much simpler methods can be used to analyse the simulations. For example, VaR or the liquidity constraint of Bardenhewer (2007). A combination of the framework with one of these methods, gives insight in both the maturity and the magnitude of the buckets. Furthermore, the simulations make the method robust and compliant with ILAAP. Since this framework is built from scratch, another convenient feature is that the volume, yield curve, and client rate model can be easily extended or partly adjusted in the future due to the modular approach. Overall,

[1] Appendix A discusses the static replicating portfolio and the extensions of Bardenhewer (2007) in more detail.

the framework is preferred over the non-maturation theory and the static replicating portfolio.

3 Data

The data of Allianz consists of daily transaction details (date, amount and mutation type) per client and daily client rates of 32 savings products over the period of December 19, 2008 to March 28, 2013 (1115 days). The savings accounts are not only for saving money, but also for the acquisition of Allianz investment funds. Due to the latter, besides deposits and withdrawals, also eleven other mutation types are present, such as the acquisition and disposal of investment funds. To acquire a fund, money must first be deposited into a savings account, and thereafter it is withdrawn to be invested in one or several funds. Once a client discontinues his position in a fund, the money is first transferred to his Allianz savings account. Thereafter he is free to leave it there, or transfer it to a paying account at another bank. It is infeasible to inspect every single transaction manually. So, in order to separate money that is meant for saving, from money that is meant for speculation, I set up some rules. Firstly, I split all transactions in just three types: deposits, withdrawals, and accrued interest payoffs. Secondly, for every single client, I adjust deposits for withdrawals and vice versa, if they occur within five days of each other. For example, a €100 deposit on day one, followed by a €60 withdrawal on day three, is treated as a €40 deposit on day one. Finally, I summarize the transactions of all clients. The accrued interest, in fact also a deposit, is not adjusted for withdrawals, so that it can be modelled separately.

In order to make the analysis more feasible, I merge products that offer the exact same interest rate over the entire observation period. This enables me to include savings products that have ceased to exist and were merged with others and products that have a very low volume and a small number of transactions. The focus of this research lies on two of the resulting aggregate products that represent about 95% of the total deposit volume on average.

3.1 Volume

Figure 1a shows a time series plot of the savings product that represents about 67% of the total volume on average. This product will be referred to as product #1. The volume is normalized to have a starting value of €1 for confidentiality reasons. The most striking is the increase of about 380% over a five-month-period during the beginning of 2010. Most likely, the attractive client rate that Allianz offered at that time, lies at the root of this, see Figure 1b. Section 3.3 discusses the client rate and measures for its attractiveness, such as the rank (Section 3.3.1) and the (relative) spread (Section 3.3.2). Section 4.1.1 goes into more detail about the relation between the deposits and the spread, which is the difference between the client rate of Allianz and competitors.

Further, today's volume is very closely related to that of yesterday, because of the large number of customers and the fact that a lot of their individual transactions cancel out each other. This results in an AR(1) coefficient that is very close to one, which corresponds to the findings in the literature, where models often include an AR factor. An Augmented Dickey Fuller test with optimal lag length based on Schwarz Information Criterion (SIC), an intercept and without a trend, results in p-value of 0.236, so that it cannot reject the null hypothesis of a unit root. To get a better view of the data I take log differences of the volume series. The log volume differences in Figure 1b clearly show the accrued interest payoffs in the beginning of each year and again the enormous increase in early 2010. Other large log differences are due to client rate reductions or wealthy clients who withdraw or deposit a large sum of money. The effect of client rate reductions, especially the ones from February and December 2010, are studied in Section 4.1.2. Also note that the mean volume of product #1 is smaller than twice the standard deviation of the volume, see Table 1. This clearly makes the non-maturation theory in combination with the approach of OeNB (2008) insufficient.

Product #2, the second largest aggregate savings product, represents on average 27% of the total balance. The characteristics of product #1 and #2 are much the same, but the former is an internet account and the latter is not. This means that transactions must be made via the telephone or by letter. The telephone is 24 hours a day accessible, so there are no significant time savings or losses. This product charges slightly higher transaction costs for investment orders, but most importantly, the client rate of product #2 is always 50 basis points lower than that of product #1. The intention is to stimulate clients to get an internet account, which is cheaper and more efficient for the bank. A comparison of product #1 and #2, with Figure 1 and Table 1 reveals several other differences and similarities.

First of all, the extreme volume increase of product #1 in early 2010, is not present for product #2. This might be due to the lower client rate, which at the same time, might be the reason for the obvious downward trend. Further, Figure 1d shows an extremely large observation on February 23, 2010 which is due to the discontinuation of an investment fund. The entire balance is deposited to the savings accounts. This is a rare event and a dummy can be incorporated in the model to account for this outlier. Lastly, due to the lower deposit volume and thus the increased influence of individual transactions, the volume time series of product #2 is less gradual.

Similar to product #1, Figure 1d clearly shows the accrued interest payoffs in the

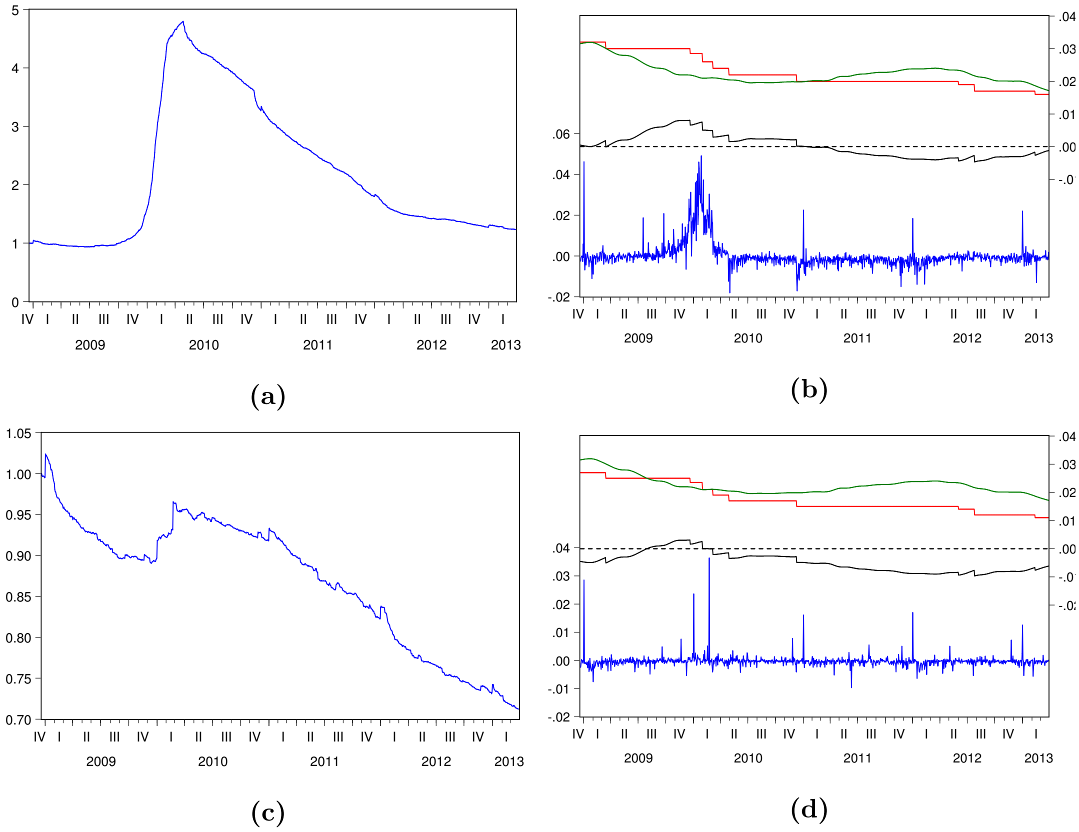


Figure 1: Panel (a) and (c) show a time series of the deposit volume. Panel (b) and (d) show the log volume differences (blue), the client rate (red), the nationwide client rate (green) and the spread (black) between the client rate and the nationwide client rate. Panel (a) and (b) belong to product #1 and panel (c) and (d) to product #2.

Table 1: Summary statistics for the deposit volume with a starting volume (V) of €1 and the log difference of the deposit volume (Δv), for both product #1 and #2. The first order autocorrelation is designated with ρ_1. The sample period runs from December 19, 2008 to March 28, 2013 (1115 observations).

	V #1	V #2	Δv #1	Δv #2
Mean	2.1563	0.8689	0.0002	-0.0003
Median	1.5813	0.8966	-0.0009	-0.0003
Minimum	0.9345	0.7123	-0.0180	-0.0096
Maximum	4.7974	1.0238	0.0491	0.0364
Std. Dev.	1.1764	0.0793	0.0065	0.0022
Skewness	0.7932	-0.5148	3.4840	8.8156
Kurtosis	2.2639	2.0227	19.771	128.41
ρ_1	0.9999	0.9997	0.7027	0.0500

beginning of each year and the presence of a unit root. The Augmented Dickey Fuller test results in a p-value of 0.978, so that the null hypothesis of a unit root cannot be rejected. Both log volume difference series have high positive skewness and excess kurtosis. This points to distributions with long tails, especially the right one, more mass on the left side of the distribution and high peaks.

3.2 Data transformation

The presence of a unit root in the deposit volume suggests that the deposit volume must first be differenced, before estimating an AR model as the literature suggests. An entirely different option is to model the deposits, withdrawals and the accrued interest payoffs individually, and later combine them to construct the volume increment. Individual models capture the effects of important variables better. For example, consider the effect of a client rate reduction (Section 4.1.2). When the effect is estimated on the log volume differences, only the net effect is visible. When modelled individually, the effect on both the deposits and the withdrawals is clear. This also gives a more accurate view on future cash flows and consequently complies better to the ILAAP manual of DNB (2012). When estimating the models for the deposits and withdrawals, it is advantageous to work with normally distributed series. This is achieved with the transformation in (2), which also greatly preserves the shape of the time series. The entire procedure is described below in more detail.

First, to make the series easier to compare over time, the ordinary deposits (D) are transformed to relative deposits (D_{rel}) by dividing the deposits with the deposit volume (V) of the previous day,[2]

$$D_{rel,t} = \frac{D_t}{V_{t-1}}. \tag{1}$$

Still, even after this transformation, Figure 2a shows that the relative deposits are much larger in early 2010 than in other periods, similar to the log volume difference. Also, the time series of the relative deposits in Figure 2a and the corresponding histogram in Figure 2c show that the series is log-normally distributed, due to the non-negative nature. I apply the following transformation on the relative deposits to obtain transformed deposits (D_{tr}),

$$D_{tr,t} = \frac{-1}{\ln(D_{rel,t})} = \frac{1}{\ln(V_{t-1}) - \ln(D_t)}. \tag{2}$$

Use of solely the natural logarithm, on the domain of the relative deposits (0 to 0.05), results in negative values and an inverted shape in absolute sense. So, I specifically use -1 as numerator and the natural logarithm as denominator to reverse this. Appendix B discusses the natural logarithm as alternative transformation in more detail.

[2] The time subscript t will be left out in future equations, when the time subscripts are the same and when it does not cause confusion.

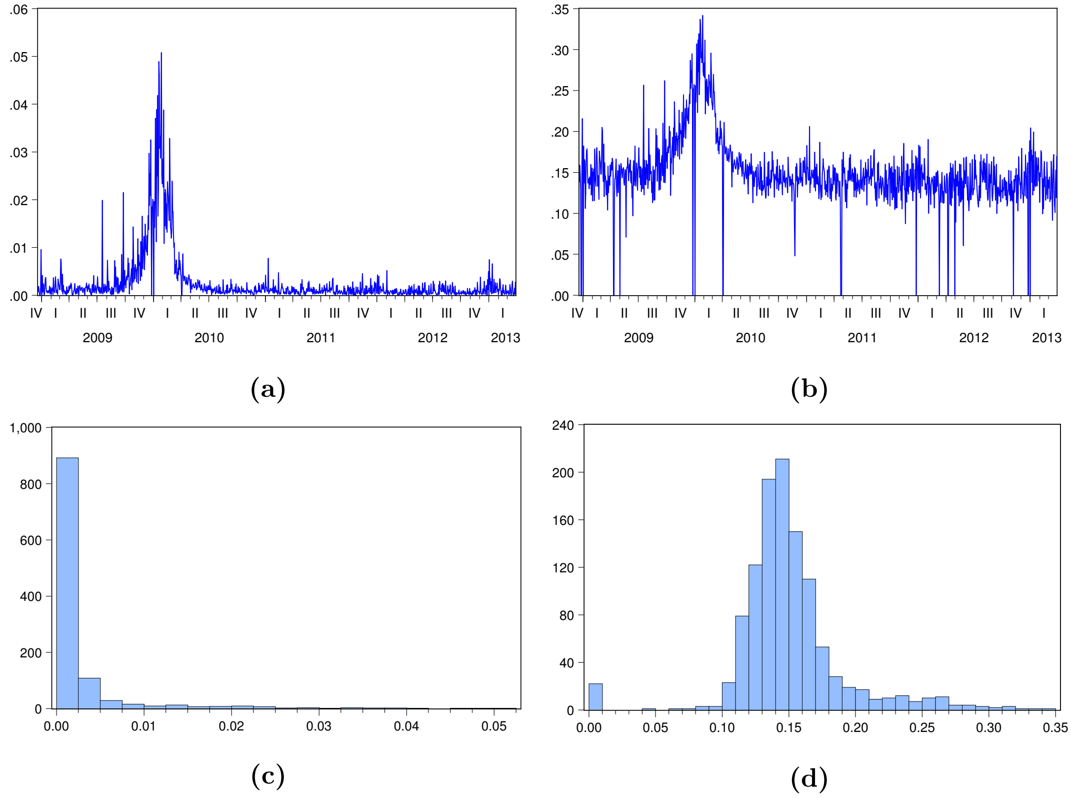


Figure 2: Panel (a) shows a time series of the relative deposits, which are deposits divided by the volume of the previous day, and panel (b) shows the transformed deposits after applying (2). Panel (c) and (d) show a histogram of the relative deposits and the transformed deposits, respectively.

Table 2: Summary statistics for the relative deposits (D_{rel}) and the transformed deposits (D_{tr}) by (2). The first order autocorrelation is designated with ρ_1. The sample period runs from December 19, 2008 to March 28, 2013 (1115 observations).

	D_{rel}	D_{tr}
Mean	0.0027	0.1498
Median	0.0010	0.1445
Minimum	0.0000	0.0000
Maximum	0.0508	0.3355
Std. Dev.	0.0057	0.0412
Skewness	4.5104	0.4159
Kurtosis	26.989	8.2612
ρ_1	0.8081	0.5783

Both transformations result in more normally distributed series (Figure 2d), but the transformation in (2) is favourable because it preserves the shape better (Figure 2b) and it behaves better for days with no transactions. All these facts are convenient for the estimation of the model. Normality of the distribution in Figure 2d is still convincingly rejected given the p-value of 0.000 for the Jarque-Bera statistic. However, the skewness and kurtosis of 0.42 and 8.26, respectively, comply better with a normal distribution compared to the skewness of 4.51 and kurtosis of 26.99 in Figure 2c. Just like the deposits, withdrawals are always non-negative and they are transformed in the exact same way, which yields the transformed withdrawals (W_{tr}).

3.3 Client rate

The main reason for consumers to deposit money is to receive interest. Interest is directly influenced by the client rate and as such, the client rate is an important variable when choosing a savings product. Moreover, the client rate is easily observable and straightforward to compare. This in contrast to other characteristics that might play a role when choosing a savings product, such as the image, the trustworthiness, and the service of the issuing bank. Similarly, when making a model for the deposit volume, the client rate is a logical starting point. Albeit, a single client rate without any reference is of little value. The client rate of a single bank must be compared with others in order to characterize it as attractive or not. The next sections discuss the rank of the client rate of Allianz and the (relative) spread between the client rate of Allianz and competitors as measures for attractiveness.

3.3.1 Rank

Neglecting all other characteristics of a savings product, a savings product with a higher client rate is considered to be more attractive than one with a lower client rate. Therefore, a straightforward way of comparing savings products is to sort them on their client rate. The product with the highest client rate receives rank #1, the one with the second highest client rate receives rank #2, etcetera. Ties have the same rank. To account for the fluctuating number of similar products, and to reflect that rank #1 of 100 is better than rank #1 of 50, the rank is transformed to a rank score,

$$\text{rank score} = 1 - \frac{\text{rank Allianz}}{\text{number of similar products}}. \tag{3}$$

The rank is an ordinal measure, so the magnitude of the difference between client rates does not matter, only the fact that there *is* a difference. This is no big deal when all the differences are more or less equal. However, if this is not the case, it might be more accurate to use a measure that does account for the magnitude of the difference.

Additional problems are caused by the construction of the rank, which requires the client rate of every savings product in the Netherlands, similar to that of Allianz. Currently, more than 100 exist and new products constantly come and go. Moreover, the historical values of all these client rates are not available to me.

3.3.2 Spread

Another measure to compare client rates, which will mostly be used in this research, is the relative spread (S_{rel}). To obtain this measure, firstly, the spread (S) is constructed as the difference of the client rate (R_{cl}) and the nationwide client rate (R_{nw}),

$$S = R_{cl} - R_{nw}. \tag{4}$$

Here, R_{nw} is the average client rate of Dutch savings accounts similar to that of Allianz. DNB provides this on a monthly basis and I transform it into a daily series with cubic spline interpolation. Secondly, the spread is divided by the level of the nationwide client rate to obtain the relative spread,

$$S_{rel} = \frac{S}{R_{nw}}. \tag{5}$$

The spread itself can also be used as a measure, but it does not discriminate between periods with different nationwide client rates, while the relative spread does. In periods with a low nationwide client rate, a spread of a certain magnitude is larger in percentage terms, and thus more attractive. The advantage of the (relative) spread over the rank score is that the measure accounts for the magnitude of the difference, and the data is more easily accessible.

3.3.3 Comparison

Another company, MoneyView, was kind enough to supply me with the weekly rank (not the client rate) of all Dutch banks. I transformed the rank to the rank score and used cubic spline interpolation to obtain daily data, so that I was able to compare the rank score with the (relative) spread. Together with the client rate of product #1 and the nationwide client rate, they are depicted in Figure 3 and the statistical properties of the measures are shown in Table 3.

First of all, note that the client rate of a given day is mostly the same as the day before. Just nine negative changes occur in the entire sample. Hence, the client rate is not able to explain the volume increase in early 2010.

On the other hand, the three measures, the spread, relative spread, and rank score, do show an increase in early 2010, which makes them potentially valuable as explanatory

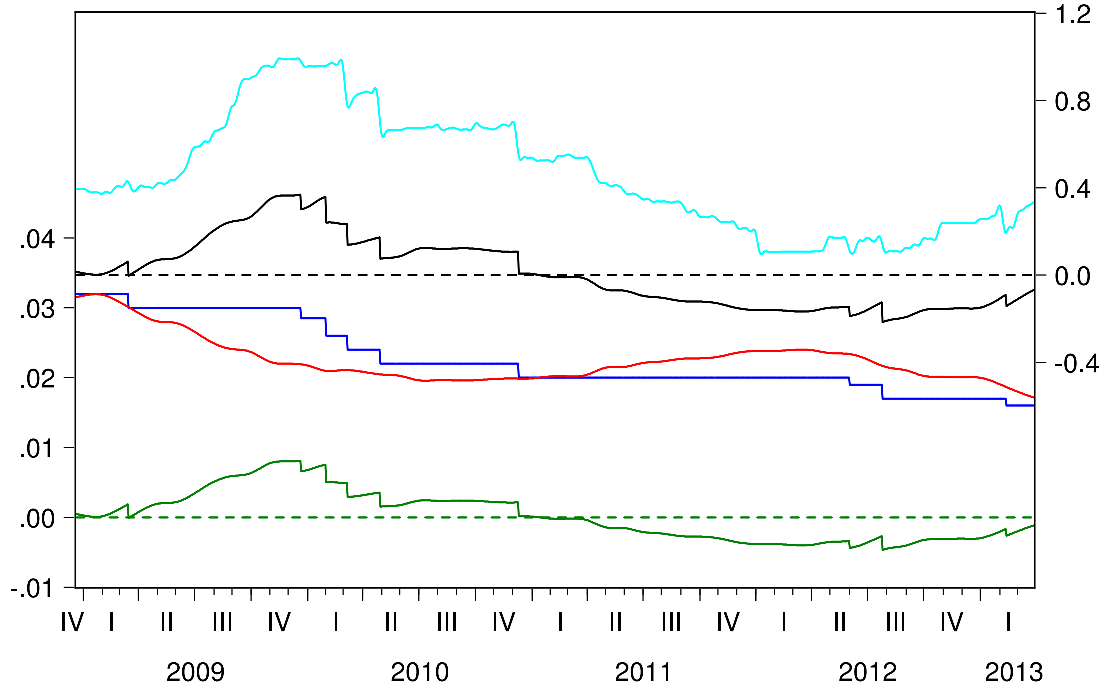


Figure 3: Time series of the client rate of product #1 (blue, left y-axis), nationwide client rate (red, left y-axis), spread (green, left y-axis), relative spread (black, right y-axis), and rank score (cyan, right y-axis). The sample period runs from December 19, 2008 to March 28, 2013 (1115 observations).

Table 3: Summary statistics for the client rate (R_{cl}), rank score, spread (S) and relative spread (S_{rel}). The first order autocorrelation is designated with ρ_1. The sample period runs from December 19, 2008 to March 28, 2013 (1115 observations).

	R_{cl}	Rank score	S	S_{rel}
Mean	0.0227	0.4618	0.0002	0.0074
Median	0.0200	0.4070	-0.0002	-0.0085
Minimum	0.0160	0.0958	-0.0047	-0.2155
Maximum	0.0320	0.9928	0.0081	0.3689
Std. Dev.	0.0049	0.2662	0.0035	0.1590
Skewness	0.6663	0.4439	0.5857	0.5978
Kurtosis	2.0235	2.0971	2.3024	2.3080
ρ_1	0.9994	0.9997	0.9989	0.9988

variable. Just like the client rate, the measures have a very high first order autocorrelation. Further, the distributions have a fairly low kurtosis (2.1 to 2.3), which means that they are relatively flat, compared to a normal distribution. The positive skewness, indicates a somewhat longer right tail and more mass on the left side of the distribution. Possibly, the bank attracts consumers with a very attractive client rate for a short period of time, and thereafter sets a less attractive client rate for a longer time period.

The nine client rate reductions are clearly visible in the three measures. Although, the third decrease, at February 1, 2010, seems not to be present in the rank score. Either, this decrease had simply no impact on the rank score, which is not likely, or the client rate reduction is not observed.

Despite the different construction of the spreads and the rank score, the time series of the measures are very similar, not taking into account the scaling. The correlation between the rank score and the spread is 0.95, the correlation between the rank score and the relative spread is 0.96, and the correlation between the spreads is even 0.99.

Concluding, we can create several logical measures, which increase in the same period that the volume increases sharply. The rank score only puts the client rates in the correct order, whereas the (relative) spread does also account for the magnitude of the difference. The relative spread even considers different nationwide client rate regimes. Also from a practical point of view, the (relative) spread is more attractive than the rank score. For the latter, sufficient historical data is not present, which makes it difficult to formulate adequate models, and the collection and processing of future data is more labour intensive. However, the similar statistical properties and the high correlations, show that the three measures are not that different after all. Taking into account all factors, the relative spread is the best option and it will be used further on in this research.

4 Methods and results

Section 2 shows that the framework provides the best basis for dividing the deposit volume in buckets. Its simulations make the method robust and forward looking and in combination with the liquidity constraint or VaR, the magnitude and maturity of the buckets can easily be determined.

The framework to create these simulations consists of several modules for important variables. The modules for the deposit volume, yield curve, nationwide client rate, nationwide deposit volume and the client rate of Allianz are discussed in Section 4.1 to 4.5 respectively. The Allianz client rate has a special purpose and serves as an input variable to observe the effect of different client rate regimes on the buckets. The models of Section 4.1 to 4.5 are combined in Section 4.6 to create simulations. Also the creation of buckets with VaR and the liquidity constraint of Bardenhewer (2007) is discussed. The final model is evaluated in Section 4.7 by comparing the buckets with the buckets of an alternative and much simpler model, the random walk.

4.1 Deposit volume model

The deposit volume model consists of separate submodels for the deposits, the withdrawals and the accrued interest payoffs. This section describes the construction of these submodels and the influence of important variables.

4.1.1 Deposits and spread

The relation between the relative spread (S_{rel}) and the transformed deposits (D_{tr}) as discussed in Section 3.3 can be captured with

$$D_{tr} = c + \beta_s \cdot S_{rel} \cdot f\left(S_{rel}\right) + \varepsilon. \tag{6}$$

Here, $f(S_{rel})$ is a function of the relative spread. For example, with $f(S_{rel}) = 1$, (6) becomes a simple linear model, but also more advanced models can be implemented. Inspection of the time series and scatter plot in Figure 4 gives more context to the relation between the relative spread and the transformed deposits and enables us to specify a correct form for $f(S_{rel})$. The figure contains evidence for a positive relation between both, just as outlined in the previous sections, but only when the relative spread is higher than about 0.1. For a relative spread below 0.1, no relation is present at all. Apparently, for some consumers the client rate is more or less irrelevant and they deposit money no matter what the relative spread is. Other clients are more selective and they only deposit money when the relative spread is high. Only for the latter clients a positive relation with the relative spread exists. Also, note that the positive relation becomes more volatile for an increasing spread.

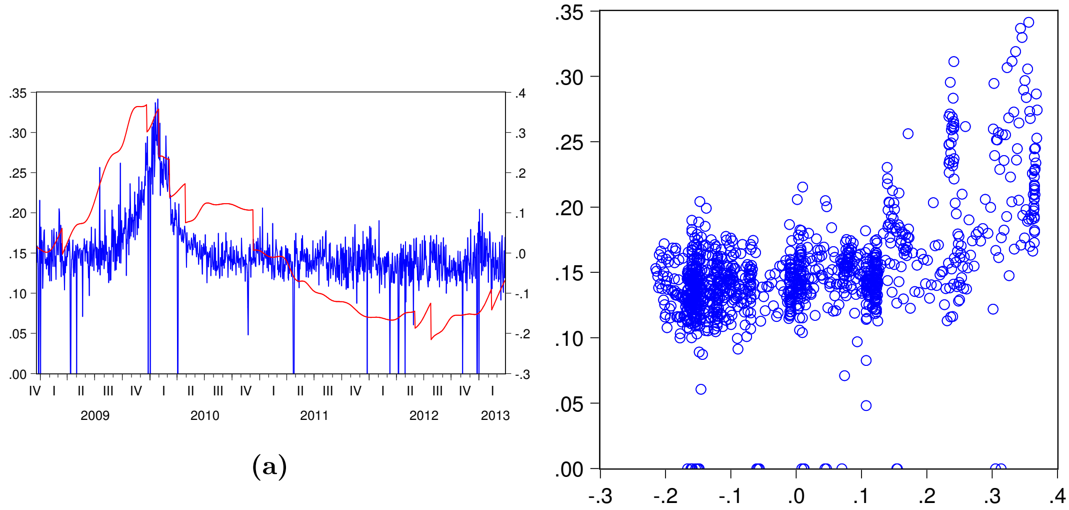
At least two approaches exist to incorporate this ambiguous relation in the model in (6) for the transformed deposits. Namely, a simple threshold,

$$f\left(S_{rel}\right) = \mathbb{1}_{[S_{rel} > \beta_{th}]}, \tag{7}$$

and a logit threshold,

$$f\left(S_{rel}\right) = \frac{\exp\left(\beta_m\left(S_{rel} - \beta_{th}\right)\right)}{1 + \exp\left(\beta_m\left(S_{rel} - \beta_{th}\right)\right)}. \tag{8}$$

In (7), $\mathbb{1}_{[A]}$ is the indicator function, which is 1 if A holds, and 0 otherwise. The logit threshold in (8) is less strict than (7) and can also attain values between 0 and 1. It can be seen that (7) is actually nested in (8). If β_m goes to infinity, (8) converges to the



(a) **(b)**

Figure 4: Panel (a) shows a time series of the transformed deposits (blue, left y-axis) and the relative spread (red, right y-axis) and panel (b) shows a scatter of the relative spread (x-axis) and the transformed deposits (y-axis).

simple dummy in (7). So, when β_m is large, the difference between both approaches is marginal, and it might be better to use the simple dummy, because it uses one parameter less. On the other hand, many regressions are used to estimate the value of β_{th} in (7), whereas for (8) just a single regression is sufficient, which makes it more practical. In Section 4.1.4, models are estimated with (7), (8) and $f(S_{rel}) = 1$.

For the relation with the transformed withdrawals, we can expect a similar but opposite behaviour. Some clients might not care about what the relative spread is and withdraw money any time. Other clients withdraw more money when the client rate is unattractive, and vice versa. For these clients a negative relation can be expected with the relative spread and possibly even a similar ambiguous relation as for the transformed withdrawals. In that case, β_s should be negative and β_m and β_{th} more or less equal to that of the deposit model.

4.1.2 Client rate adjustment

Besides the differences in the level of the client rate of Allianz and competitive banks, an extra effect might be present when a bank changes its client rate. In this research only a client rate adjustment of Allianz is considered, because client rate data of other individual banks is not available.

Consider the case where Allianz reduces its client rate. This degrades the competitive advantage, and assuming that the other characteristics remain the same, a logical consequence would be a drop in the deposit volume, caused by an increase in withdrawals. However, not all clients might be aware of the reduction or may be in the position to withdraw money immediately. As a result, money is still withdrawn some days after the client rate reduction. The above scenario, a shock in the transformed withdrawals (W_{tr}) that (slowly) dies off, can be captured by an exponential decline,

$$W_{tr} = c + \beta_{ed} \cdot \exp(\beta_j \cdot j) + \varepsilon. \tag{9}$$

Note that a lower client rate is also incorporated in the relative spread, but the latter variable does not capture the shock of the client rate decline. In this equation, j stands for the number of days since the last client rate reduction and it is zero on days that a new client rate is effective for the first time. To comply with the outlined scenario above, β_{ed} must be positive and β_j must be negative. Additionally, $\frac{ln(0.5)}{\beta_j}$ indicates after how many days only 50% of the initial shock in the transformed withdrawals is left. I will refer to this as the half life time. A reasonable value for the half life time ranges from a few days to weeks, but not months or even years.

For clarity, I examine the client rate reductions of April and December 2010, which have a clear impact. Figure 5 shows the exponential decline of (9) fitted on the transformed withdrawals of April 28, 2010 to April 21, 2011 (257 days). The estimated values for c, β_{ed} and β_j are respectively $0.164, 0.048$ and -0.050, all having a p-value of 0.000. The estimated coefficients are intuitive in the sense that β_{ed} is positive and β_j is negative and corresponds to a half life time of about 14 days, which is quite long but still acceptable.

The exponential decline model can be complemented with two additional variables. To start with, the level of the relative spread before alteration ($S_{rel,ba}$) might be used as a threshold. For the same reason that the relative spread only affects customers above some threshold, a decline might only have effect if the previous spread was above a certain level. Secondly, the difference in the relative spread before and after the client rate change, the drop of the relative spread ($S_{rel,dr}$), can be used to modify the magnitude of the shock. The reasoning behind this variable is that larger declines have a greater effect. With these additions in their simplest form, (9) becomes

$$W_{tr} = c + \beta_{ed} \cdot \mathbb{1}_{[S_{rel,ba} > \tau_{ba}]} \cdot S_{rel,dr} \cdot \exp(\beta_j \cdot j) + \varepsilon, \tag{10}$$

where τ_{ba} is the threshold. However, since only nine client rate reductions take place in the entire sample period, also only nine different observations are available for $S_{rel,ba}$ and $S_{rel,dr}$. This makes it hard to estimate a relation, as will become clear in Section 4.1.4.

Besides, the clients that leave the bank and withdraw their money, also no longer deposit their money. The withdrawal of their entire wealth is quite noticeable as can be seen in Figure 5. However, assuming that the regular deposits only involve a small

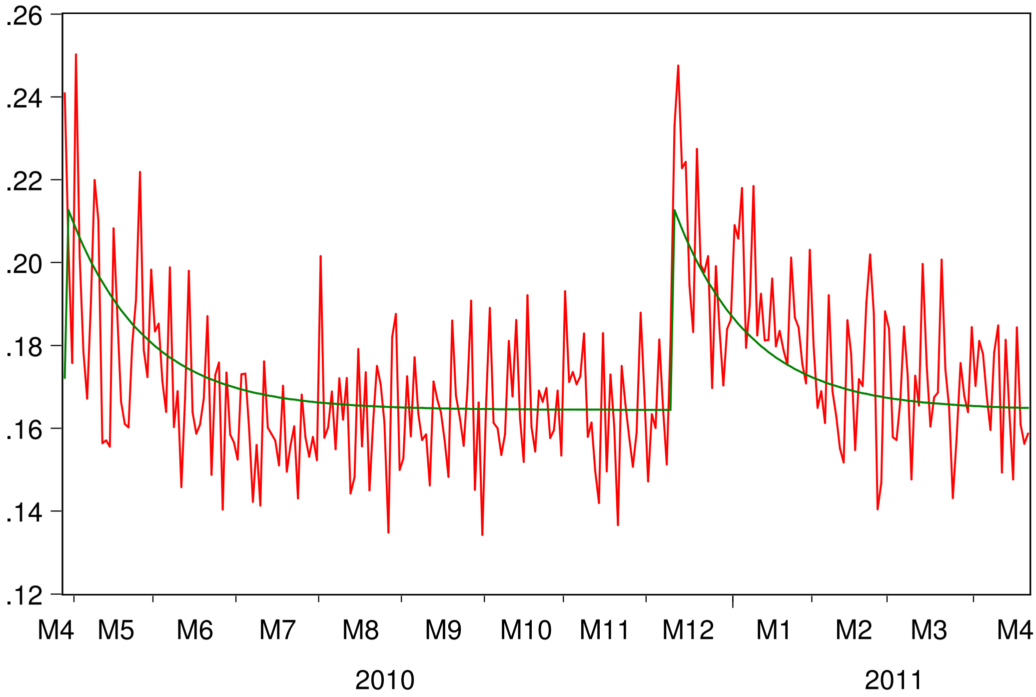


Figure 5: A time series of the transformed withdrawals (red), together with a fitted exponential decline (green) as in (9). The sample period runs from April 28, 2010 to April 21, 2011 (257 days).

percentage of their wealth, the discontinuation of the regular deposits has a marginal influence. Therefore, it is not necessary to include a variable that accounts for a client rate reduction in the model for the transformed deposits. A simplified numerical example in Appendix C illustrates this.

Besides a client rate reduction, a client rate increase might as well take place. The effect of a client rate increase might be visible as a positive shock in the deposits, because new consumers join the bank. However, consumers who are not (yet) a client of the bank, do not receive a notification of this increase. Furthermore, they need to monitor the client rate of many banks, and not just a single one. It is therefore likely that the consumers react slower compared to the decline in client rate. On the other hand, consumers that want to switch are likely to switch immediately when they note the increase. Concluding, an exponentially declining positive shock might be present in the deposits. Compared to the client rate decline shock in the withdrawals, the initial effect might be delayed and the half life time somewhat longer. Additionally, the relative spread *after* alteration and the magnitude of the change might be valuable.

Similar to the fact that a client rate reduction barely influences the deposits, a client rate increase is not likely to influence the withdrawals much. Customers that were already satisfied with the former and lower client rate, are certainly content with the new one. Unfortunately, I am not able to verify these relations, because there is no client rate increase present in the sample period.

4.1.3 Nationwide deposit volume

Lastly, I consider the change of the nationwide deposit volume to be an important variable. Every change in the deposit volume of a nation comes from changes in the total deposit volume of all individual banks together and vice versa. Assuming that the percentage change in the nationwide volume can be attributed to each account of every individual client equally, each individual bank should experience the same percentage change. This assumption is somewhat simplistic and banks with the best characteristics will experience a lesser decline or a firmer increase than the nationwide average. Still, the deposits of every single bank are likely to be positively correlated with the log difference of the nationwide deposit volume (Δv_{nw}) and the withdrawals negatively.

4.1.4 Estimation

Table 4 shows several models (A to F) for the transformed deposits (D_{tr}) of savings product #1. Besides the estimated coefficients, with p-values in parentheses, also the R^2, the SIC and the Log-Likelihood (LogL) are included to compare the models on their (adjusted) goodness of fit.

Model A is the most basic variant, which includes just a constant (c), a dummy for national holidays for which deposits are not available (D_{na}), and the relative spread (S_{rel}). The coefficients are significant and yield the correct sign. The R^2 is relatively high, given the simplicity, but a comparison with model B shows the shortcomings of estimation with Ordinary Least Squares (OLS). Model A is identical to B, except for the fact that the latter accounts for heteroskedasticity with the GARCH variance estimator,

$$h_t = \omega + \alpha \cdot \varepsilon_{t-1}^2 + \beta \cdot h_{t-1} + \gamma \cdot \mathbb{1}_{[S_{rel,t} > 0.170]}, \tag{11}$$

without the γ part. The GARCH terms are highly significant and the SIC and LogL values show that model B is better in terms of adjusted goodness of fit. This leads to the conclusion that the coefficients and significance of model A might be wrong, which seems particularly true for the coefficient of the relative spread. The GARCH variance estimator is also used for the other models, because heteroskedasticity is clearly present.

In addition to the former models, C includes the log differences of the nationwide deposit volume (Δv_{nw}). The positive sign confirms the hypothesis that it correlates positively to the transformed deposits. However, the variable is only marginally significant on the 90% confidence level and not on the 95% level. The evaluation criteria are not

Table 4: Estimated coefficients with p-values in parentheses for regressions of the transformed deposits (D_{tr}) of savings product #1. The explanatory variables from top to bottom are a constant (c), a dummy for days with no deposits (D_{na}), the log difference of the nationwide deposit volume (Δv_{nw}), the relative spread (S_{rel}), and β_m and β_{th} belong to the threshold relation in (7) and (8). Row seven to ten contain the estimates of the GARCH parameters of (11). The bottom section displays the R^2, SIC and LogL value for evaluation. The sample period runs from December 19, 2008 to March 28, 2013 (1115 days).

	A	B	C	D	E	F
c	0.152 (0.000)	0.146 (0.000)	0.145 (0.000)	0.140 (0.000)	0.140 (0.000)	0.140 (0.000)
D_{na}	−0.149 (0.000)	−0.141 (0.000)	−0.141 (0.000)	−0.145 (0.000)	−0.144 (0.000)	−0.145 (0.000)
Δv_{nw}			0.093 (0.092)	0.137 (0.003)	0.147 (0.001)	0.155 (0.001)
S_{rel}	0.131 (0.000)	0.069 (0.000)	0.064 (0.000)	0.160 (0.000)	0.161 (0.000)	0.171 (0.000)
β_m					388.3 (0.487)	388.3 (0.474)
β_{th}				0.128	0.127 (0.000)	0.127 (0.000)
ω $\left[\times 10^{-5}\right]$		2.891 (0.000)	2.665 (0.000)	2.936 (0.000)	2.814 (0.000)	4.370 (0.000)
α		0.105 (0.000)	0.101 (0.000)	0.105 (0.000)	0.104 (0.000)	0.092 (0.000)
β		0.839 (0.000)	0.847 (0.000)	0.834 (0.000)	0.838 (0.000)	0.791 (0.000)
γ $\left[\times 10^{-4}\right]$						1.306 (0.013)
R^2	0.523	0.442	0.447	0.592	0.593	0.602
SIC	−4.266	−4.742	−4.738	−4.810	−4.800	−4.808
LogL	2389	2665	2666	2706	2708	2716

unanimous on whether or not it improves the model, but model D to F show that the variable becomes more significant for more advanced models.

Model D and E incorporate the threshold relation with the relative spread that is discussed in Section 4.1.1. Model D uses the simple dummy in (7), with a threshold value of 0.128. This threshold is chosen from a range of values (0 to 0.3) with steps of 0.001, because it results in the highest Log Likelihood value. Due to the fixed value in each of the 301 regressions, the p-value is not available. The threshold (β_{th}) of model E relates to (8) and is estimated in a single regression. The p-value of the β_m-coefficient indicates that the estimated value is not significantly different from zero. However, the β_m value of model D is actually ∞ and combined with the almost identical β_{th}-coefficient, the threshold relation of both models is about the same. This gives confidence that model E is estimated correctly. Also the other coefficients and the evaluation criteria of both models are similar. Model E is not significantly better than model D, but from a practical perspective it is more convenient to use model E. It is more flexible in the estimation and every coefficient can be estimated in a single regression, whereas 301 regressions are needed to determine the threshold for model D.

The last model for the deposits includes an extra variance parameter as in (11), to reflect the extra uncertainty when the relative spread is above 0.170. This extra uncertainty is already mentioned in Section 4.1.1 and stems from Figure 4b. The value of 0.170 is chosen, because it results in the highest Log Likelihood value when varying the threshold between 0 and 0.3 with steps of 0.001. The coefficient is positive and significant, which confirms the hypothesis of extra uncertainty, and it also improves the model significantly. Model F will be used in the simulation.

Similarly, the estimation results for the transformed deposits (D_{tr}) of savings product #2 are shown in Table 5. As can be seen in Section 3, savings product #2 offers a far less attractive client rate than savings product #1, and as a consequence, many variables that are discussed in the previous sections are not valuable for this model. For example, the threshold relation between deposits and the relative spread. People that are interested in a high client rate, and thus a high relative spread, will deposit in savings product #1 and not in #2, because the client rate of the latter is always 50 basis points lower. Additionally, the models for product #2 include a dummy for a single outlier (D_{ol}) that is caused by the discontinuation of an investment fund.

Model E, F and G, are the models with a threshold and confirm that it is indeed not valuable. The threshold of model E is determined by testing a range of values, and thresholds smaller than -0.450 all result in the same maximum likelihood. In perspective, the minimum value of the relative spread in this sample is -0.446. So, the dummy is effectively always 1, which is the reason why the coefficients are exactly the

Table 5: Estimated coefficients with p-values in parentheses for regressions of the transformed deposits (D_{tr}) of savings product #2. The explanatory variables from top to bottom are a constant (c), a dummy for a single outlier (D_{ol}), a dummy for days with no deposits (D_{na}), the log difference of the nationwide deposit volume (Δv_{nw}), the relative spread (S_{rel}), and β_m and β_{th} belong to the threshold relation in (7) and (8). Row eight to ten contain the estimates of the GARCH parameters of (11) without the γ-part. The bottom section displays the R^2, SIC and LogL value for evaluation. The sample period runs from December 19, 2008 to March 28, 2013 (1115 days).

	A	B	C	D	E	F	G
c	0.123 (0.000)	0.123 (0.000)	0.123 (0.000)	0.123 (0.000)	0.123 (0.000)	0.117 (0.000)	0.117 (0.000)
D_{ol}	0.181 (0.000)	0.180 (0.705)	0.181 (0.000)	0.181 (1.000)	0.181 (0.000)	0.187 (0.000)	0.187 (0.999)
D_{na}	−0.117 (0.000)	−0.117 (0.000)	−0.117 (0.000)	−0.117 (0.000)	−0.117 (0.000)	−0.117 (0.000)	−0.117 (0.014)
Δv_{nw}	0.015 (0.772)	0.020 (0.700)					
S_{rel}	0.025 (0.000)	0.025 (0.000)	0.025 (0.000)	0.025 (0.000)	0.025 (0.000)	0.089 (0.000)	0.093 (0.000)
β_m						456.4 (0.899)	456.4 (0.912)
β_{th}					−0.450	0.038 (0.044)	0.038 (0.171)
ω $\left[\times 10^{-4}\right]$		6.695 (0.040)		1.875 (0.606)			1.814 (0.333)
α		0.020 (0.320)		0.013 (0.476)			0.029 (0.207)
β		−0.502 (0.474)		0.572 (0.481)			0.576 (0.170)
R^2	0.498	0.498	0.498	0.498	0.498	0.488	0.488
SIC	−4.835	−4.816	−4.841	−4.822	−4.841	−4.808	−4.790
LogL	2713	2713	2713	2713	2713	2701	2702

same as of model C. Model F and G are estimated with the logit threshold variable in (8) with OLS and GARCH respectively. The LogL values are worse than that of all other models and it is therefore not valuable to include the threshold relation.

Furthermore, the GARCH parameters in model B, D and G are far from significant and sometimes even negative. From model A and B it appears that also the log difference of the nationwide volume cannot contribute to the model. The fact that this variable is significant for product #1 corresponds with the assumption that more attractive savings products have a stronger relation with this variable. Model C has the best overall statistics, the best SIC value and significant coefficients; this model will therefore be used in the simulation.

Table 6 shows the estimation results for the transformed withdrawals (W_{tr}) of savings product #1. Model A and B contain the same variables, but A is estimated with OLS and B with the GARCH variance estimator. The estimated coefficients of both methods are very much alike, and the sign of the coefficients corresponds with the expectations. In addition to the previously explained variables, a dummy for Mondays (D_{mon}) is included. On these days, more withdrawals take place, because they are postponed from the weekend. A similar dummy variable is not included in the deposits model, because there it yields the wrong sign. Despite the strong similarities between A and B, the LogL values and the high significance of the GARCH parameters indicate that it is better to estimate with a GARCH model.

In addition to A and B, model C contains the basic exponential decline as in (9). The addition of this variable is worthwhile, because the LogL value indicates a large improvement. Moreover, the coefficients are significant and their sign corresponds with the expectation. The β_j-coefficient corresponds with a half life time of three working days, which is on the short side but credible.

The threshold relation as discussed in Section 4.1.1, is captured with (7) and (8) in model D and E respectively. The β_{th}-coefficient of D is chosen from a range of values (0 to 0.3) with steps of 0.001, because it gives the highest Log Likelihood value. The estimated coefficients of both models are virtually identical. The reason for this is the very large value of the β_m-coefficient in model E. A β_m-coefficient of ∞ corresponds with the simple dummy of (7) as in model D. The fewer parameters in D result in a lower SIC value, but the fit is practically identical given the R^2 and LogL value. From a practical point of view, it is easier to use the logit relation of E as in (8), because it can be estimated with one regression. That is also why it is used to estimate a more advanced exponential decline.

As described in Section 4.1.2, the exponential decline can be complemented with the level of the relative spread before alteration ($S_{rel,ba}$) as a threshold, and the mag-

Table 6: Estimated coefficients with p-values in parentheses for regressions of the transformed withdrawals (W_{tr}) of savings product #1. The explanatory variables from top to bottom are a constant (c), a dummy for days with no withdrawals (D_{na}), a dummy for Mondays (D_{mon}), the log difference of the nationwide deposit volume (Δv_{nw}), β_{ed} to β_j belong to the exponential decline as in (9) and (10), the relative spread (S_{rel}), and β_m and β_{th} belong to the threshold relation in (7) and (8). Row thirteen to fifteen contain the estimates of the GARCH parameters of (11) without the γ-part. The bottom section displays the R^2, SIC and LogL value for evaluation. The sample period runs from December 19, 2008 to March 28, 2013 (1115 days).

	A	B	C	D	E	F	G	H
c	0.162	0.162	0.162	0.164	0.164	0.165	0.157	0.157
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
D_{na}	−0.164	−0.164	−0.164	−0.165	−0.165	−0.165	−0.164	−0.163
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
D_{mon}	0.018	0.019	0.019	0.019	0.019	0.019	0.018	0.018
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Δv_{nw}	−0.259	−0.271	−0.314	−0.284	−0.284	−0.303	−0.095	−0.086
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.045)	(0.070)
β_{ed}			0.019	0.019	0.019			
			(0.000)	(0.000)	(0.000)			
$\beta_{ed} \cdot S_{rel,dr}$						−0.250		
						(0.000)		
$\beta_{ed} \cdot \mathbb{1}_{[A]}$							0.016	
							(0.000)	
$\beta_{ed} \cdot S_{rel,dr} \cdot \mathbb{1}_{[A]}$								−0.159
								(0.000)
β_j			−0.225	−0.186	−0.186	−0.130	−0.004	−0.004
			(0.010)	(0.008)	(0.008)	(0.001)	(0.001)	(0.001)
β_s	−0.010	−0.011	−0.011	−0.036	−0.036	−0.040	−0.054	−0.049
	(0.013)	(0.000)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
β_m					1125	189.0	800.6	2408
					(0.952)	(0.782)	(0.922)	(0.970)
β_{th}				0.103	0.103	0.100	0.090	0.097
					(0.000)	(0.001)	(0.000)	(0.001)
ω $\left[\times 10^{-5}\right]$		2.135	1.633	1.427	1.427	1.196	1.541	1.733
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
α		0.103	0.106	0.089	0.089	0.082	0.071	0.072
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
β		0.840	0.852	0.872	0.872	0.886	0.882	0.875
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.585	0.585	0.592	0.606	0.605	0.609	0.636	0.637
SIC	−5.076	−5.143	−5.147	−5.173	−5.160	−5.166	−5.214	−5.212
LogL	2848	2895	2905	2919	2919	2919	2946	2945

nitude of the drop of the relative spread ($S_{rel,dr}$) as a multiplier. Only nine client rate reductions are present in the sample, so many thresholds (0.062 to 0.107) result in the same maximum likelihood. The mean value is used in the estimation so, $\mathbb{1}_{[A]}$ in Table 6 corresponds with $\mathbb{1}_{[S_{rel,ba}>0.085]}$. Model F and G include the additional variables individually and model H combines them. The sign of the exponential decline in model F becomes negative, but this is due to the negative values in the drop of the relative spread. Although, the LogL value of model G and H show a large improvement, the half life time that corresponds with the β_j-coefficient is far from credible, namely about six months. The half life time of model F corresponds with about five working days, which is certainly credible. However, there is only a marginal and non-significant improvement from E to F, so I use model E for the simulation. The statistics of D are actually better than E, but E is more practical and flexible in the estimation.

Table 7 reports the estimation results for the transformed withdrawals (W_{tr}) of savings product #2. Because this product is far less attractive, the threshold relation and the exponential decline are not included in this model. Clients who are sensitive for changes in the client rate, do not invest in a product with a client rate that is 50 basis points lower. The remaining variables are used in every possible combination, estimated with OLS as well as with the GARCH variance estimator.

To start with, model A, B, E and F include the log difference of the nationwide deposit volume. This variable is expected to be negatively correlated with the withdrawals, but here the estimated coefficients are positive. This is not intuitive and therefore the variable should be excluded. The effect of the relative spread is very small, and far less than in the previous withdrawal and deposit models. Moreover, it is far from significant and thus also this variable should not be included. What remains is a very simple model, with just two dummies and a constant. The evaluation criteria of the GARCH variant are better than that of the OLS model, although the estimates of both models are almost identical. Additionally, the α and β parameter of the GARCH model are highly significant and the ω just not on the 95% confidence level, so model H, the GARCH variant, will be used in the simulation.

Lastly, the accrued interest payoffs should be modelled; these are not included in the deposits model. I multiply the client rate with the deposit volume of each day, and because money cannot be transferred in the weekends, each Friday is multiplied with $\frac{3}{365}$ and other days with $\frac{1}{365}$. The calculated accrued interest should match the real accrued interest. Since the data sample runs from December 19, 2008, the accrued interest over 2008, that is paid out in 2009, cannot be determined. A regression of the real accrued interest on the calculated accrued interest of the other four years, results in a coefficient of 0.990. With a standard deviation of 0.001, this is significantly different

Table 7: Estimated coefficients with p-values in parentheses for regressions of the transformed withdrawals (W_{tr}) of savings product #2. The explanatory variables from top to bottom are a constant (c), a dummy for days with no deposits (D_{na}), a dummy for Mondays (D_{mon}), the log difference of the nationwide deposit volume (Δv_{nw}), and the relative spread (S_{rel}). Row six to eight contain the estimates of the GARCH parameters of (11) without the γ-part. The bottom section displays the R^2, SIC and LogL value for evaluation. The sample period runs from December 19, 2008 to March 28, 2013 (1115 days).

	A	B	C	D	E	F	G	H
c	0.130 (0.000)	0.131 (0.000)	0.134 (0.000)	0.133 (0.000)	0.133 (0.000)	0.133 (0.000)	0.134 (0.000)	0.134 (0.000)
D_{na}	−0.136 (0.000)	−0.135 (0.000)	−0.135 (0.000)	−0.135 (0.001)	−0.136 (0.000)	−0.135 (0.001)	−0.135 (0.000)	−0.135 (0.001)
D_{mon}	0.004 (0.002)	0.004 (0.003)	0.004 (0.002)	0.004 (0.003)	0.004 (0.002)	0.004 (0.002)	0.004 (0.002)	0.004 (0.003)
Δv_{nw}	0.157 (0.000)	0.106 (0.025)			0.102 (0.007)	0.058 (0.150)		
S_{rel}	−0.010 (0.021)	−0.007 (0.092)	−0.002 (0.606)	−0.002 (0.553)				
ω $\left[\times 10^{-5}\right]$		2.659 (0.094)		2.768 (0.068)		2.734 (0.078)		2.606 (0.064)
α		0.038 (0.008)		0.043 (0.005)		0.042 (0.005)		0.044 (0.004)
β		0.886 (0.000)		0.879 (0.000)		0.881 (0.000)		0.883 (0.000)
R^2	0.456	0.455	0.450	0.449	0.453	0.452	0.449	0.449
SIC	−5.075	−5.070	−5.070	−5.072	−5.077	−5.073	−5.076	−5.078
LogL	2847	2854	2841	2852	2844	2853	2841	2852

from the ideal value 1. The difference between the estimated and real series comes from clients who terminate their account during the year, and receive their accrued interest earlier. Despite the difference, the accrued interest that is used for the simulation, is determined as above.

4.2 Yield curve

This section presents the yield curve module, which is mostly based on the work of Diebold and Li (2006). I follow the steps of their research to show that my data is comparable and so that I can use their method. Then the Kalman filter is introduced, as a more advanced and accurate estimation method, for which Hamilton (1994), Kim and Nelson (1999) and De Pooter (2007) are used.

As has been stated in the literature section, there are many approaches to model and simulate the yield curve. An advantage of the Nelson-Siegel model over the others is that it poses structure on the loadings, which can be identified with the level, slope and curvature of the yield curve.

4.2.1 Data

The data used to replicate the approach of Diebold and Li (2006) comes from DNB and Bloomberg. It consists of end-of-month yields of 1, 3 and 6-month Euribor rates and 1, 2, 3, 5, 10 and 30-year Euro Swaps over the period of January 2003 to March 2013 (123 months). The yields of more maturities are available, but there is a discrepancy between the 12-month Euribor rate and the 1-year Euro Swap. In order to have a yield curve that can be fitted correctly, I include just these yields. It would be even better to use either Euribor rates or Euro Swaps, but the former only exist for the short end of the yield curve and latter only for the long end.

Occasionally, yields become marginally negative, such as the yield of Dutch 3-month government bonds during the summer of 2013, but generally it can be assumed that yields are non-negative. The data used for this research, does not violate that assumption. To prevent future yield curve simulations from becoming negative, the initial yields are transformed with the natural logarithm. These yields will be referred to as transformed yields. They are transformed back after the simulation, which results in simulated yields that are non-negative.

Table 8 contains the mean, standard deviation and 1, 12 and 30-month autocorrelation of the transformed yields and it shows that most stylized facts of a yield curve are present. Namely, the mean shows that the average yield curve is increasing, high autocorrelations show that the yield dynamics are persistent and the standard deviations show that the long end of the yield curve is less volatile than the short end. However,

Table 8: The mean, standard deviation and 1, 12 and 30-month autocorrelation of the transformed yields. Maturities are given in months.

Maturity	Mean	Std. Dev.	ρ_1	ρ_{12}	ρ_{30}
1	-4.261	1.015	0.964	0.376	0.320
3	-4.075	0.818	0.964	0.376	0.235
6	-3.942	0.663	0.963	0.386	0.152
12	-3.904	0.649	0.962	0.425	0.197
24	-3.814	0.601	0.956	0.459	0.252
36	-3.720	0.546	0.954	0.464	0.265
60	-3.565	0.430	0.950	0.458	0.263
120	-3.361	0.285	0.946	0.468	0.240
360	-3.258	0.246	0.950	0.516	0.259

the table does not confirm that the long end is more persistent than the short end, because ρ_{30} does not increase for longer maturities. Though, the last stylized fact does hold for the untransformed yields.

4.2.2 Diebold and Li

To represent the yield curve, Diebold and Li (2006) use the following three-factor Nelson-Siegel model,

$$y_t(\tau_i) = \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda_t\tau_i}}{\lambda_t\tau_i}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda_t\tau_i}}{\lambda_t\tau_i} - e^{-\lambda_t\tau_i}\right) + \varepsilon_{i,t}, \tag{12}$$

where y_t is the bond yield with maturity τ_i in months, and $\beta_{1t}, \beta_{2t}, \beta_{3t}$, and λ_t need to be estimated. Diebold and Li (2006) assume that λ_t is constant over time, so that only the beta's need to be estimated. The fixation of λ_t greatly simplifies the estimation, because all loadings become constant, and the factors $(\beta_{1t}, \beta_{2t}, \beta_{3t})$ can then easily be estimated with OLS. There is no consensus on the value of λ_t in the literature. Diebold and Li (2006) suggest to choose it such that the loading of the factor β_{2t} achieves its maximum at a maturity of 30 months. In that case, the correct value of λ_t is 0.0598.[3]

The time series of the factors $(\beta_{1t}, \beta_{2t}, \beta_{3t})$ in (12) are estimated by OLS, with 0.0598 for all λ_t. Diebold and Li (2006) show that the factors are closely related to the level (l_t), slope (s_t) and curvature (c_t). Throughout the literature, the level is defined as the constant 1, but there is some variation in the definition of the slope and curvature. My definition of the slope and curvature is equivalent to that of Diebold and

[3]Diebold and Li (2006) use 0.0609 for λ_t, but this maximizes the loading for β_{2t} at approximately 29.45 months.

Li (2006) in the sense that the former is defined as the yield of the longest maturity minus the shortest maturity in the data set. Similarly, the curvature is constructed as twice the yield of the median maturity, minus the yield of the longest and the shortest maturity. The correlation between (β_{1t}, l_t), (β_{2t}, s_t), and (β_{3t}, c_t) is 0.98, -0.99, and 0.99 respectively, which is comparable to Diebold and Li (2006) who find 0.97, -0.99, and 0.99 respectively. For a more detailed elaboration on the factors $(\beta_{1t}, \beta_{2t}, \beta_{3t})$ and the level, slope and curvature I refer to their work.

Then, Diebold and Li (2006) use OLS to estimate an AR model for each factor. It is also possible to use a vector autoregressive (VAR) model, in which case the factors are also dependent on the lag of the other two factors. This is beneficial if the cross-correlation between all factors is high. However, that is not the case. Moreover, a VAR model is more complicated to estimate due to the larger number of parameters. Additionally, this might result in overfitting and consequently poor forecasts. Table 9 shows the estimated coefficients for the AR(1) model,

$$\beta_{i,t} = c_i + \phi_i \cdot \beta_{i,t-1}. \tag{13}$$

The estimated coefficients and p-values in parentheses provide some important insights. First of all, the AR-coefficient ϕ_2 is larger than one. This is bad, since it stimulates future β_2's to become bigger and bigger; it diverges. Secondly, the estimated constants are insignificant. Instead, the value of ϕ_2 should be smaller than one, such that the series converges to the unconditional mean, which is $\frac{c_i}{1-\phi_i}$. This is also the reason why a constant must be incorporated in the model, even if it is not significant. The incorporation of a constant makes sure that the future beta's converge to a value different from zero, provided that ϕ_i is smaller than 1.

So, for my data set, it is not sufficient to solely use OLS to estimate the AR model for the factors. De Pooter (2007) uses the Kalman filter, which is a more general estimation method. It is able to estimate the fixed value for λ_t and restrictions can be imposed in the numerical estimation program, such that the AR-coefficient is smaller than 1. Just like De Pooter (2007), I use the OLS estimates as starting values for the Kalman filter.

4.2.3 Kalman filter

An alternative way of estimating (12) and (13), is to represent both in a state space model and solve it with the Kalman filter. The Kalman filter is a recursive method to obtain an optimal forecast of the state vector B_t, using all information up to time $t-1$. Contrary to OLS, which estimates the equations separately, this method does it simultaneously.

In order to do this, first (12) and (13) are formulated in a state space model, which

Table 9: Estimated coefficients with p-values in parentheses for the AR(1) model in (13). The sample period runs from January 2003 to March 2013 (123 months).

i	c_i	ϕ_i
1	-0.091 (0.272)	0.974 (0.000)
2	-0.010 (0.603)	1.008 (0.000)
3	-0.040 (0.222)	0.807 (0.000)

consists of a state and a measurement equation. The latter is

$$Y_t = XB_t + \varepsilon_t, \tag{14}$$

which is the matrix notation of (12). Y_t is a $[9 \times 1]$ vector with yields, X is a $[9 \times 3]$ matrix with the factor loadings and B_t is a $[3 \times 1]$ vector with the factors, $(\beta_{1t}, \beta_{2t}, \beta_{3t})'$. The state equation

$$B_t = \mu + \Phi B_{t-1} + \nu_t \tag{15}$$

is the matrix notation of the AR(1) model in (13), where μ is a $[3 \times 1]$ vector of constants, $(c_1, c_2, c_3)'$, and Φ is a $[3 \times 3]$ diagonal matrix, with ϕ_1, ϕ_2, and ϕ_3 on the diagonal. The residual vectors ε_t and ν_t are assumed to be independently normally distributed with mean 0 and variance R and Q respectively,

$$\varepsilon_t \sim \mathcal{N}(0, R), \tag{16}$$

$$\nu_t \sim \mathcal{N}(0, Q), \tag{17}$$

where R is a $[9 \times 9]$ diagonal matrix and Q is a $[3 \times 3]$ diagonal matrix.

For now we assume that the nineteen parameters in $\Theta(\lambda, \mu, \Phi, R, Q)$ are known. The state vector, B_t, can then be estimated by:

$$B_{t|t-1} = \mu + \Phi B_{t-1|t-1}; \tag{18}$$

$$P_{t|t-1} = \Phi P_{t-1|t-1} \Phi' + Q; \tag{19}$$

$$\eta_{t|t-1} = Y_t - Y_{t|t-1} = Y_t - XB_{t|t-1}; \tag{20}$$

$$f_{t|t-1} = XP_{t|t-1}X' + R; \tag{21}$$

$$B_{t|t} = B_{t|t-1} + P_{t|t-1}X' f_{t|t-1}^{-1} \eta_{t|t-1}; \tag{22}$$

$$P_{t|t} = P_{t|t-1} - P_{t|t-1}X' f_{t|t-1}^{-1} XP_{t|t-1}. \tag{23}$$

To initialize these equations, I use the unconditional mean,

$$B_{0|0} = (I - F)^{-1} \mu, \tag{24}$$

and the unconditional covariance matrix of the state vector, which is derived in Kim and Nelson (1999),

$$\operatorname{vec}\left(P_{0|0}\right) = \left(I - F \otimes F\right)^{-1} \operatorname{vec}\left(Q\right). \tag{25}$$

With the initial values for $B_{0|0}$ and $P_{0|0}$, a prediction for time period 1 can be made, $B_{1|0}$ and $P_{1|0}$, and consequently also a prediction of the yield curve, $Y_{1|0}$. In time period 1, the prediction of the yield curve can be evaluated with the real value, $Y_{1|1}$, to create the prediction error, $\eta_{1|0}$. In turn, the prediction error is used to adjust the prediction of the state vector to get a better estimate of the state vector, $B_{1|1}$, with a better fit of the yield curve. This recursive process is repeated until every observation in the sample is used.

However, the parameter set (Θ) is *not* known in advance and must be estimated as well. This is done in an iterative process. It starts with the assumption of an initial parameter set, $\Theta^{(0)}$, which is used to run the Kalman filter. Along the way, the prediction error ($\eta_{t|t-1}$) and the conditional variance of the prediction error ($f_{t|t-1}$) are created, which are used to calculate the Log Likelihood,

$$\mathcal{L} = \sum_{t=1}^{T} \left[-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\left(\left|f_{t|t-1}\right|\right) - \frac{1}{2}\eta'_{t|t-1} f^{-1}_{t|t-1} \eta_{t|t-1} \right]. \tag{26}$$

Then, a new parameter set, $\Theta^{(1)}$, is used to repeat this process, with the goal to obtain a better Log Likelihood. Eventually, it converges to a point where the Log Likelihood does not become any better, such that you obtain the maximum Log Likelihood parameter set, Θ_{ML}.

With 19 parameters to estimate, it is not a trivial task to find the optimal parameter set, and the maximum Log Likelihood value might actually be a local maximum instead of a global maximum. It is therefore advisable to perform the iterative process with a range of starting parameter sets. Adequate starting parameters are also important from a time management perspective. Similar to De Pooter (2007), I use many of the OLS estimates in $\Theta^{(0)}$. Furthermore, the iterative process in Matlab is able set upper and lower bounds, such that the AR-coefficients can be forced to be between 0 and 1.

The Kalman filter obtains -0.270, -0.015, -0.055, 0.915, 0.993 and 0.953 for the non-zero coefficients in μ and Φ, respectively. These estimates are quite similar to the OLS estimates with the difference that all AR(1) parameters are smaller than one in this case. Also the Kalman estimate for λ_t (0.0591) is close to the value used in the OLS approach (0.0598).

4.3 Nationwide Client rate

This section describes the module for the nationwide client rate, which is used to create the relative spread. Two models from the literature will be considered to simulate this variable, namely O'Brien (2000) and Paraschiv and Schürle (2010).

An important and interesting phenomenon, that both models aim to capture, is (upward) client rate stickiness. This means that the (nationwide) client rate reacts stronger to a decline in the market rate, than to an increase in the market rate. From the perspective of a bank this makes perfect sense. When the market rate goes down, banks adjust the client rate in a similar way to prevent losses. On the other hand, when the market rate goes up, banks can make a little bit more profit by delaying the increase of the client rate. However, both actions can only be done to some extent. Clients might switch to another bank if the decrease of the client rate is too rapid or they might invest in other savings products with a more market conform yield if the increase of the client rate is too slow.

4.3.1 Data

The data used in this section consists of the previously used yield curve data in Section 4.2 and additionally it uses the nationwide client rate that also comes from DNB. The latter is an end-of-month volume weighted average of the client rates of all Dutch savings products with a maturity shorter than three months. A similar variable for savings products with no maturity at all is not available. Equivalent to the yield curve section, the sample period runs from January 2003 to March 2013 (123 observations) and the market and nationwide client rate are transformed with the natural logarithm to prevent negative yields.

4.3.2 Paraschiv and Schürle

Paraschiv and Schürle (2010) construct an AR(1) model for the differences of the nationwide client rate,

$$\Delta R_t = \kappa_1 + \kappa_2 \Delta R_{t-1} + \kappa_3 \Delta r^{short}_{t-1} + \kappa_4 \Delta r^{long}_{t-1} + \kappa_5 EC_{t-1} + \kappa_6 \overline{\omega}_{t-1} + \varepsilon_t. \tag{27}$$

Besides the AR specification, this model includes the difference of the 3-month Libor rate (Δr^{short}_{t-1}), the difference of the 5-year Euro Swap rate (Δr^{long}_{t-1}), an error correction term (EC_{t-1}), and a threshold variable ($\overline{\omega}_{t-1}$),

$$\overline{\omega}_{t-1} = \omega_{t-1} \mathbb{1}_{[\omega_{t-1} \leq \tau]}. \tag{28}$$

Paraschiv and Schürle (2010) include an error correction term, because their nationwide client rate and 5-year Euro swap rate are cointegrated. The threshold variable makes it possible to estimate a threshold relation for either the short rate, the long rate or the error correction term; ω_{t-1} takes the role of the variable of interest. For example,

Table 10: Estimated coefficients and corresponding p-values in parentheses of (29). The R^2, SIC and LogL are used for evaluation. The sample period runs from January 2003 to March 2013 (123 observations).

κ_2	κ_3	R^2	SIC	LogL
0.499 (0.000)	0.041 (0.013)	0.319	−5.267	313

when it is used for the short rate and ω_{t-1} is below the threshold (τ), the effect becomes $\kappa_3 + \kappa_6$ and otherwise it is just κ_3.

When this model is estimated with data from this research, many variables become insignificant. First of all, all yields are integrated of order 1, but there is no cointegration relation present. Secondly, none of the combinations of a short Euribor rate with a long Euro Swap rate is significant. Regressions with either Euribor or Euro swap rates, show that it is better to use the Euribor rates, because those are significant and the Euro swaps are not. Also intuitively it makes more sense to include a Euribor rate, because they have a shorter maturity and that matches better with the contractual zero maturity of deposits. Also the threshold variable is excluded, because it is only marginally significant on the 95% confidence level and for extreme values only, such that it affects just about 2% of the observations. It should affect a large share of the negative values, so a reasonable value would be close to zero.

After all insignificant variables and incorrect relations are removed, an AR(1) model for the log differences of the nationwide client rate ($\Delta r_{nw,t}$), with the log differences of 3-month Euribor rate ($\Delta y_{3m,t-1}$) remains,

$$\Delta r_{nw,t} = \kappa_2 \Delta r_{nw,t-1} + \kappa_3 \Delta y_{3m,t-1} + \varepsilon_t. \tag{29}$$

The estimated coefficients with p-values and statistical measures are depicted in Table 10. This model still consists of the variables that are most often mentioned in the literature, an AR(1) model with a short rate, but it is not able to explain (upward) client rate stickiness.

4.3.3 O'Brien

Another option for the nationwide client rate module is the approach of O'Brien (2000), who specifies an asymmetric partial adjustment model. The model assumes that the client rate reverts to an equilibrium rate, which is a function of the 3-month market rate (Y_{3m}), with a speed that depends on whether the client rate is above or below the equilibrium rate.

Some minor adjustments are made to the model of O'Brien (2000). Firstly, he

Table 11: Estimated coefficients and corresponding p-values in parentheses of (30). In model A, λ^- and λ^+ can be estimated freely, while model B forces them to be equal. The R^2, SIC and LogL are used for evaluation. The sample period runs from July 2003 to March 2013 (117 observations).

model	λ^+	λ^-	ζ_1	ζ_2	R^2	SIC	LogL
A	0.419 (0.008)	5.887 (0.001)	0.764 (0.000)	−0.014 (0.000)	0.465	−5.426	327
B	2.638 (0.000)	2.638 (0.000)	0.409 (0.000)	−0.014 (0.000)	0.340	−5.257	315

uses the model for the client rate of many individual banks, but it actually suits the nationwide client rate better. The former has mostly discrete steps and it changes only occasionally, whereas the latter can have a variety of step magnitudes and it always changes. Secondly, O'Brien (2000) models the ordinary difference of the client rate, whereas I model the log difference to prevent a non-negative nationwide client rate.

Table 11 shows the estimated coefficients, with corresponding p-values in parentheses, for two specifications of the following model,

$$\Delta r_{nw,t} = \left(\lambda^+ I_t + \lambda^- (1 - I_t)\right)(\zeta_1 Y_{3M,t} - \zeta_2 - R_{nw,t-1}) + \varepsilon_t, \tag{30}$$

$$I_t = \begin{cases} 1, & \text{if } (\zeta_1 Y_{3m,t} - \zeta_2 - R_{nw,t-1}) > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{31}$$

In model A, both λ^- and λ^+ are estimated freely, while model B forces both to be equal. Model A clearly shows that λ^- is significantly larger than λ^+, which perfectly corresponds to the assumption that downward adjustments are made a lot quicker than upward adjustments. Also, the evaluation criteria show that model A is significantly better than model B. Compared to the model of Section 4.3.2, model A not only has a more elegant interpretation, but also the statistical properties are better. The model of O'Brien will be used in the simulation.

4.4 Nationwide Deposit Volume

This section discusses the module for the log differences of the nationwide deposit volume, which is used as an explanatory variable for the deposits and the withdrawals. The data used in this module, the nationwide deposit volume and the nationwide client rate, come from DNB. Monthly values for both are obtained for the sample period of January 2003 to March 2013 (123 observations).

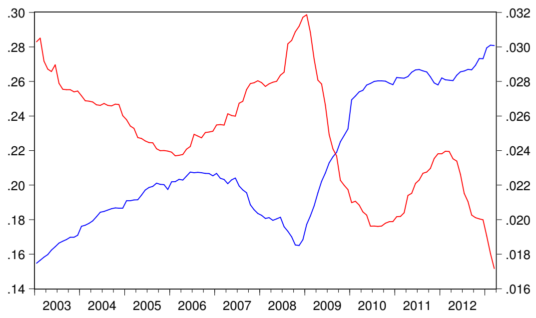
The relations between the client rate and the deposit volume are quite different on a national level than for individual banks. For the latter, generally a higher and more competitive client rate corresponds with a higher deposit volume. On a national level however, the financial situation of a country is important. This is influenced by monetary and interest rate policies, which aim to have a stable economy with moderate growth. So, when the economy flourishes and people spend a lot of money, the goal is to dampen this to some extent. This can be done by a high market rate, with consequently a high client rate. This stimulates people to save part of their wealth for the future, instead of spending it all now. Some people might be affected by this policy and save for later, but the majority does not. As a consequence, the nationwide deposit volume is low when the nationwide client rate is high.

The other side of the spectrum is a recession. In such a period, people are uncertain about the future and they are reluctant to spend money. In order to stimulate spending, governments decrease the market rates and again the client rates follow. It depends on the insecurity of the consumers, which is related to the severity of the recession, whether the low client rate has effect. In a recession, when the nationwide client rate is low, it is likely that the nationwide deposit volume is high.

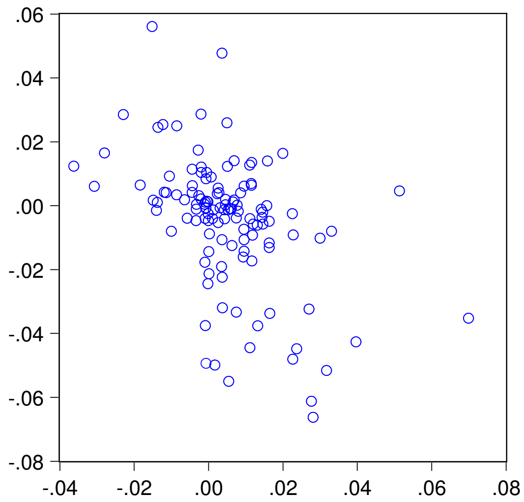
The level of supply and demand of deposit volume strengthens the effect of the financial situation. When a lot of people want to deposit money, i.e. a large supply, banks do not need to be very competitive and it is sufficient to offer a relatively low client rate, which enforces a lower client rate. Alternatively, when the supply of deposit money is low, banks need to offer competitive client rates in order to convince consumers to deposit money with them. This enforces a higher client rate.

Figure 6 shows the negative relation between the nationwide deposit volume and client rate and Table 12 shows several models for the log differences of the nationwide deposit volume. Next to the log differences of the nationwide client rate, also dummies for the months January (D_{jan}) and May (D_{may}) are included. In January and May, consumers have extra money due to the end of the year bonus and holiday allowance, respectively.

Model A is estimated with OLS and includes a constant which is insignificant. Without the constant and estimated with OLS (model B), all coefficients are significant. When the same model is estimated with a GARCH variance estimator (model C), this results in the best LogL value, but the dummy for May becomes insignificant, just like the GARCH ω-coefficient. Without the dummy for May and still estimated with GARCH (model D), all variables are significant, but even more GARCH parameters become insignificant. In addition, the LogL value of model D drops to about the same level of B, and the SIC and R^2 drop even more. Thus, model B obtains the best values for the evaluation criteria and is used in the simulation.



(a)



(b)

Figure 6: Panel (a) shows a time series of the nationwide deposit volume (blue, left y-axis) and the nationwide client rate (red, right y-axis). The left y-axis is in trillions (10^{12}) of €. Panel (b) shows a scatter of the log difference of the nationwide deposit volume (x-axis) and the log difference of the nationwide client rate (y-axis). The sample period runs from January 2003 to March 2013 (123 observations).

Table 12: Estimated coefficients and p-values in parentheses for the model of the log difference of the nationwide deposit volume. The R^2, SIC and LogL are used for evaluation. The sample period runs from July 2003 to March 2013 (117 observations).

	A	B	C	D
c	-0.001 (0.156)			
D_{jan}	0.025 (0.000)	0.024 (0.000)	0.023 (0.000)	0.024 (0.000)
D_{may}	0.013 (0.000)	0.012 (0.000)	0.011 (0.143)	
$\Delta\ln(r_{nw,t})$	-0.200 (0.000)	-0.196 (0.000)	-0.126 (0.002)	-0.147 (0.003)
$\Delta\ln(v_{nw,t-1})$	0.424 (0.000)	0.400 (0.000)	0.341 (0.000)	0.370 (0.000)
ω $[\times 10^{-6}]$			2.973 (0.227)	3.821 (0.229)
α			0.159 (0.052)	0.120 (0.120)
β			0.827 (0.000)	0.853 (0.000)
R^2	0.596	0.588	0.568	0.529
SIC	-6.322	-6.345	-6.410	-6.299
LogL	382	381	392	383

4.5 Client rate

This section covers the module for the Allianz client rate. Unlike other variables, the client rate is an internal process, so Allianz has total control over it. Therefore, the goal of this module is not to estimate a certain relation with external variables, but to assume a management process that serves as input variable. The real client rate adjustment process is based on investment valuations, profit perspectives and internal expert opinion. This research simplifies this process, by assuming that the client rate is only adjusted to maintain the relative spread in a certain region.

More specifically, the future client rate (R_{cl}) remains unadjusted, until the relative spread (S_{rel}) exceeds a lower or upper bound for a given number of days ($t_n \geq 1$) in a row. The lower and upper bound are symmetric around a reference point (o) at a distance of $-b$ and $+b$ respectively. The new value of the client rate is determined with the simulated nationwide client rate (R_{nw}) and a rewritten version of (5) and such that the resulting relative spread is equal to a lower or upper rebound value. Just like the lower and upper bound, the lower and upper rebound are symmetric around the reference point (o) at a distance of $-b_r$ and $+b_r$ respectively. Here, it must obviously hold that $0 < b_r < b$. In formula form this yields:

$$R_{cl,T+h} = \begin{cases} (1+o+b_r) \cdot R_{nw,T+h} & \text{if } \left(\sum_{i=1}^{t_n} \mathbb{1}_{[S_{rel,T+h-i} > o+b]}\right) = t_n; \\ (1+o-b_r) \cdot R_{nw,T+h} & \text{if } \left(\sum_{i=1}^{t_n} \mathbb{1}_{[S_{rel,T+h-i} < o-b]}\right) = t_n; \\ R_{cl,T+h-1} & \text{otherwise;} \end{cases} \tag{32}$$

where the parameters (o, b, b_r, t_n) are to be determined by the user. Here, $\mathbb{1}_{[A]}$ is the indicator function, which is 1 if A is true, and 0 otherwise. Further, T is the end of the sample period, and $h > 1$ is a number of days in the future.

4.6 Simulation

This section combines the estimated models of Section 4.1 to 4.5 and uses the stationary block bootstrap of Politis and Romano (1994) to create 1000 simulations for five years (1306 days) into the future. The creation of buckets with VaR and the liquidity constraint of Bardenhewer (2007) is discussed in Section 4.6.1. Section 4.6.2 contains a simulation example and shows the influence of the client rate model parameters on the buckets.

The residuals of some models contain autocorrelation, and in order to capture this, the stationary block bootstrap of Politis and Romano (1994) is used. This procedure samples blocks of residuals with a varying random length and a random starting value of the block.[4] Additionally, the models based on monthly observations (the yield curve, na-

[4]Although, the stationary bootstrap only captures the autocorrelation to some extent, it performs

tionwide client rate and nationwide volume model) use the residuals of identical periods, in order to account for possible cross-correlation in their residuals.

The simulation starts with the yield curve model, estimated with the Kalman filter as in Section 4.2.3. Because the Kalman filter uses starting values for $B_{0|0}$ and $P_{0|0}$, the first few estimates might be biased and not close to the real values. For this reason, the first six observations, which is about 5% of the sample, are not used in the Log Likelihood computation. This is in accordance with the approach of De Pooter (2007), who also leaves out 5% of his observations in the determination of the Log Likelihood. The remaining 117 residual observations in combination with the estimated AR(1) models are used to simulate future values for the state vector. The simulated state vector in combination with the Nelson-Siegel model in (12) with $\tau_i = 3$ months and $\lambda_t = 0.0591$ and the corresponding residuals are used to simulate the 3-month yield. Other yields can be constructed as well, but only the 3-month yield is used in the nationwide client rate model.

Thereafter, model A of Table 11 and model B of Table 12 are used to create the nationwide client rate and nationwide log volume difference simulations respectively. Note that these models are also estimated with 117 observations, while 123 observations are available, to incorporate the cross and autocorrelation correctly. Lastly, the monthly simulations are transformed into daily simulations with cubic spline interpolation, similar to how they were transformed for the estimation of the models.

Then, the end user specifies the parameters (o, b, b_r, t_n) of the client rate model, as outlined in Section 4.5. Together with the 3-month market rate simulation, this results in the relative spread and client rate. The chosen models from Table 4 to 7 use the simulated relative spread and the other explanatory variables to construct the transformed deposits (D_{tr}) and withdrawals (W_{tr}). In turn these are transformed back to relative deposits (D_{rel}) and withdrawals (W_{rel}) with

$$D_{rel} = \exp\left(\frac{-1}{D_{tr}}\right). \tag{33}$$

In order to retrieve meaningful relative deposits and withdrawals, the transformed variables must be greater than zero. Almost all simulations satisfy this restriction, but generally a small part $\left(\frac{6}{1306000} \approx 0.0005\%\right)$ does not. For simulations with increasingly attractive client rate regimes, this part becomes larger $\left(\frac{298}{1306000} \approx 0.023\%\right)$, but still remains relatively small.[5] To overcome this problem, negatively simulated transformed

much better than the ordinary bootstrap, simple block bootstrap (non-overlapping blocks) and moving block bootstrap (overlapping blocks with fixed length). For more details about the stationary bootstrap, see Politis and Romano (1994).

[5]Here, $\frac{6}{1306000}$ corresponds with $o = \mu - 1.4\sigma$ to $o = \mu$ and $\frac{298}{1306000}$ corresponds with $o = \mu + 1.4\sigma$ in Table 13 respectively.

deposits and withdrawals are set to zero, such that they represent no (relative) deposits and withdrawals respectively. Then, the deposit volume (V), normal deposits (D) and withdrawals (W) can be calculated recursively with the last volume observation, and the relative deposits and withdrawals:

$$D_t = D_{rel,t} \cdot V_{t-1}; \tag{34}$$

$$V_t = V_{t-1} + D_t - W_t = V_{t-1} \cdot (1 + D_{rel,t} - W_{rel,t}). \tag{35}$$

Also, on the first day of each new year, the (sum of the daily) accrued interest of the previous year is added to the volume. The daily accrued interest is calculated as the product of the daily deposit volume and client rate multiplied with $\frac{3}{365}$ for Fridays and $\frac{1}{365}$ for other days, see Section 4.1.4.

4.6.1 Buckets

With the volume simulations at hand, the buckets can be created with various measures. Two relatively simple measures are described in this section, namely VaR and the liquidity constraint of Bardenhewer (2007).

The VaR specifies the maximum loss for a certain horizon and with a predetermined probability. To be clear, the horizon associated with VaR can be interpreted in two ways, the maximum loss at precisely that horizon, or the maximum loss in the period up to that horizon. This research uses the latter, because the bank must be able to pay back money in the intermediate period as well.

To construct the buckets with VaR, for each simulation an auxiliary series is constructed as the minimum of the volume up to that time. Thereafter, the 5% quantile of all auxiliary minimum series is evaluated at the horizons of interest to construct the VaR values. The bucket with the greatest maturity is constructed as the starting volume minus the VaR value of the greatest maturity. The other maturity buckets are constructed as the difference between the VaR value of that maturity and one maturity level higher.

To illustrate the VaR procedure, assume that the deposit volume today is €100, that the buckets with the greatest maturity are 11 and 12 months and that the 95% 12-month VaR is €80 and the 95% 11-month VaR is €70. This means that we are 95% confident that over a 12-month period the volume does not drop below €20. In order to prevent liquidity issues, the 12-month bucket must not exceed €20 and to maximize profit with an increasing yield curve, the 12-month bucket must be as large as possible. So, the 12-month bucket must be €20. Similarly, the 11-month bucket must not exceed €30. However, €20 of the €30 are already included in the 12-month bucket, such that the 11-month bucket consists of just €10. The remaining buckets are constructed in a similar recursive manner.

To / From	100	102	99	98	96
100	0	R12 +2.0%	R13 -1.0%	R14 -2.0%	R15 -4.0%
102	-	0	R23 -2.9%	R24 -3.9%	R25 -5.9%
99	-	-	0	R34 -1.0%	R35 -3.0%
98	-	-	-	0	R45 -2.0%

Figure 7: Illustration of how the most negative relative volume change is obtained for various maturities (green = 1 day, yellow = 2 days, red = 3 days and blue = 4 days).

The other method, the liquidity constraint, is used by Bardenhewer (2007) to restrict the weights of the replicating portfolio, such that there is always enough liquidity to pay back clients in the historical estimation period. This method can also be used without the replicating portfolio approach and for many simulations instead of a single historical realisation.

The construction of the liquidity constraint buckets consists of the following steps. Firstly, for each simulation the relative volume change from time period to time period must be determined. Secondly, for each maturity we select the most negative relative volume change over the horizon corresponding with the maturity. This is the liquidity constraint for some horizon in a single simulation. Thereafter, we take the 5% quantile of all the simulations of the liquidity constraint for each maturity, which gives an estimate of the percentage of deposits that is at risk for that maturity. Finally, the buckets are created similar as in the VaR approach.

The procedure is illustrated in Figure 7 for a 5-day simulation and with buckets with maturities of 0 to 4 days. In the figure, R_{kl} represents the relative change in volume from day k to day l and the colors represent a maturity. For the 1-day maturity, we only consider the green circled values; the most negative one is -2.9%. Similarly, for the $2, 3$ and 4-day maturity we obtain $-3.9\%, -5.9\%$ and -5.9%. Smaller horizons are a subset of longer horizons, so the worst relative change becomes more negative with an increasing horizon. Here, we only observe a single simulation, but in case we would have more simulations, we could now determine the 5% quantile for each maturity. Finally, the buckets are created similar to the VaR approach. This results in weights of $2.9\%, 1.0\%, 2.0\%$ and 0.0% for the $0, 1, 2$ and 3-day bucket, respectively. The remaining 94.1% is put in the 4-day bucket, because that is the longest maturity in this example.

The difference between the VaR and liquidity constraint approach is that the former only evaluates the volume between the start of the simulation and some time into the future, for example a month. On the other hand, the liquidity constraint evaluates multiple time periods with a length of a month. So, not only the start of the simulation until a month in the future, but also the period of two to three months in the future, and many others. Since the liquidity constraint evaluates more periods, it is stricter than the VaR approach. For example, the VaR approach would only consider R_{12}, R_{13}, R_{14} and R_{15} in Figure 7. This would result in $0.0\%, 1.0\%, 1.0\%, 2.0\%$ and 96.0% for the $0, 1, 2, 3$ and 4-day buckets, respectively, which is less strict than the liquidity constraint buckets. Consequently, the VaR approach always results in an equally high or higher average duration compared to the liquidity constraint.

4.6.2 Client rate scenario analysis

This section analyses the influence of the parameters in the client rate model and the difference in effect on product #1 and #2. It also confirms that the liquidity constraint creates stricter buckets than VaR.

Table 13 and 14 show the estimated VaR and liquidity constraint buckets, for product #1 and #2 respectively, for five different parameter sets of the client rate model. Three of four parameters are fixed. The boundary ($b = \sigma$) and the rebound ($b_r = 0.7\sigma$) are both a function of σ, the standard deviation of the relative spread in the historical sample. The number of days (t_n) that the boundary must be exceeded in a row is chosen to be 10. On the one hand, this value must not be too small, because that might trigger a client rate change too fast. The relative spread might revert to a 'normal' level by itself and a constantly changing client rate is not customer friendly. On the other hand, t_n must not be too large, because then it changes too slow and perhaps jeopardizes the strategy of the management. $t_n = 10$ corresponds with two weeks, which seems a nice compromise. For the reference point (o), I use $\mu - z \cdot \sigma$, where μ and σ are the mean and the standard deviation of the relative spread in the historical sample. I vary z between -1.4 and 1.4 with steps of 0.7, which corresponds to very unattractive and very attractive client rate regimes. The reference point can be characterised as the targeted relative spread and the boundaries as the acceptable bandwidth.

Across methods and tables, the bucket with the shortest maturity generally has the largest weight and buckets for increasing maturities become smaller. This pattern is visible for the monthly buckets (0 to 11 months) as well as for the yearly buckets ($\leq$ 11 and 12 to 48 months). The clear exception is the bucket with a maturity of 48 months. This is the greatest maturity available in this example, and therefore it also contains money for longer maturities. The fact that buckets become smaller for an increasing horizon directly corresponds with the decreasing uncertainty increments. The uncertainty at two time periods into the future, is larger than the uncertainty at just one time period into the future, but it is less than twice that uncertainty. This is due to the randomness in the simulation and the imperfect correlation.

Secondly, both tables confirm that the liquidity constraint is stricter than VaR. The cumulative weight of buckets of the liquidity constraint is at least as high as that of the VaR buckets. This must be the case, since the VaR observes less time periods than the liquidity constraint does. Consequently, the resulting average duration of the liquidity constraint is lower.

Thirdly, the effect of the different client rate regimes is intuitive, for both product #1 and #2. Clients prefer a higher client rate, which corresponds with a higher z, and thus withdraw money less quickly and deposit more money if the client rate is higher. The result of an increasingly attractive client rate is an increased volume, with less negative changes and a higher average duration. The higher average duration means that the bank can invest in bonds with a greater maturity that yield more profit. However, not considering the most attractive client rate regime, the differences in average duration are small. At most one month in difference for product #1 ($13.9 - 14.9$) and less than two months for product #2 ($32.6 - 34.5$). The bank should evaluate whether the extra profit of the increased average duration outweighs the increased cost of a higher client rate.

The tables also show that the very attractive client rate regime, where $z = 1.4$, has the most effect on product #1. This is in accordance with the assumption that the clients of product #2 care less about the level of the client rate. The pronounced effect on product #1 is due to the incorporation of the threshold relation with the relative spread. This also explains the large difference in average duration, about six months, between a moderately attractive client rate ($z = 0.7$) and a very attractive client rate ($z = 1.4$). The former has a relative spread that is mostly below the threshold, while that of the latter is mostly above. Again, it depends on the costs of the higher client rate, whether this increased duration is profitable.

Table 13: The VaR and liquidity constraint bucket weights in percentages for maturities (Mat.) in months, for savings product #1. The parameters used in the client rate model are: $t_n = 10$, $b = \sigma$, $b_r = 0.7\sigma$ and $o = \mu + z \cdot \sigma$, where z is $-1.4, -0.7, 0, 0.7$ and 1.4. Here, μ (0.007) and σ (0.159) are the mean and the standard deviation of the relative spread, respectively, in the historical sample (December 19, 2008 to March 28, 2013). The penultimate row (≤ 11) shows the sum of buckets with a maturity equal to or shorter than 11 months. The average duration (Avg.) is presented in the last row. All buckets are estimated with 1000 simulations, five years (1306 days) into the future.

	VaR					Liquidity constraint				
Mat.	$z = -1.4$	-0.7	0	0.7	1.4	-1.4	-0.7	0	0.7	1.4
0	8.1	8.0	8.0	7.9	6.5	14.5	14.5	14.0	13.6	11.3
1	6.4	6.2	6.2	6.1	4.7	7.0	6.9	6.9	6.2	5.6
2	4.9	4.9	4.9	5.0	3.4	5.1	5.2	5.0	5.0	4.3
3	4.4	4.0	4.0	4.0	3.4	4.1	3.9	3.9	4.1	3.5
4	3.9	4.0	4.0	4.0	3.5	3.6	3.6	3.6	3.6	3.2
5	3.6	3.5	3.5	3.4	2.6	3.4	3.4	3.4	3.2	2.7
6	3.6	3.7	3.7	3.8	3.2	3.2	3.4	2.9	3.0	2.9
7	2.9	3.1	3.1	3.1	2.2	3.3	3.2	3.2	3.4	2.7
8	3.0	2.9	2.9	2.7	2.7	2.4	2.2	2.6	2.3	2.2
9	1.5	1.3	1.2	1.3	1.1	2.3	2.5	2.2	2.3	2.1
10	2.8	2.8	2.8	2.6	1.8	2.6	2.6	2.3	2.4	1.9
11	2.6	2.6	2.6	2.7	1.9	2.4	2.1	2.6	1.9	2.2
12	21.6	21.8	21.5	20.5	15.7	19.0	19.0	19.1	19.3	15.2
24	12.1	12.3	12.2	12.3	8.2	10.8	10.8	10.8	11.2	8.8
36	7.7	7.8	7.5	7.6	6.5	6.4	6.5	6.5	6.3	2.3
48	10.9	11.1	11.9	13.0	32.8	9.9	10.2	10.9	12.2	29.3
≤ 11	47.8	47.0	47.0	46.6	36.8	53.9	53.5	52.7	51.1	44.5
Avg.	15.5	15.7	15.9	16.3	23.5	13.9	14.0	14.4	14.9	20.5

Table 14: The VaR and liquidity constraint bucket weights in percentages for maturities (Mat.) in months, for savings product #2. The parameters used in the client rate model are: $t_n = 10$, $b = \sigma$, $b_r = 0.7\sigma$ and $o = \mu + z \cdot \sigma$, where z is $-1.4, -0.7, 0, 0.7$ and 1.4. Here, μ (-0.219) and σ (0.162) are the mean and the standard deviation of the relative spread, respectively, in the historical sample (December 19, 2008 to March 28, 2013). The penultimate row (≤ 11) shows the sum of buckets with a maturity equal to or shorter than 11 months. The average duration (Avg.) is presented in the last row. All buckets are estimated with 1000 simulations, five years (1306 days) into the future.

	VaR					Liquidity constraint				
Mat.	$z = -1.4$	-0.7	0	0.7	1.4	-1.4	-0.7	0	0.7	1.4
0	2.1	2.1	2.1	2.0	1.9	5.0	5.0	5.0	4.9	4.8
1	2.1	2.1	2.1	2.0	1.8	2.3	2.3	2.3	2.1	2.0
2	1.6	1.6	1.6	1.5	1.5	1.9	1.9	1.8	1.7	1.6
3	1.5	1.5	1.5	1.4	1.2	1.7	1.6	1.5	1.4	1.1
4	1.3	1.3	1.3	1.2	1.1	1.4	1.5	1.5	1.4	1.3
5	1.4	1.4	1.4	1.3	1.3	1.3	1.2	1.3	1.1	0.9
6	1.3	1.3	1.2	1.2	0.9	1.2	1.1	1.2	1.1	0.9
7	1.1	1.1	1.1	0.9	0.9	1.1	1.2	1.1	0.9	0.9
8	1.1	1.1	1.0	1.0	0.9	1.1	1.1	1.0	1.0	0.9
9	0.6	0.6	0.6	0.5	0.4	1.3	1.1	1.1	1.0	0.8
10	0.7	0.7	0.7	0.5	0.4	1.0	1.1	1.0	1.1	0.9
11	1.2	1.2	1.1	0.9	0.7	1.0	1.0	1.0	0.7	0.6
12	10.7	10.5	10.5	9.2	7.9	10.3	10.1	9.6	8.5	7.5
24	8.8	8.5	7.9	7.1	6.1	8.1	7.9	7.7	6.8	5.3
36	7.8	7.6	7.6	6.8	5.7	6.8	6.6	6.8	5.7	4.7
48	56.7	57.2	58.1	62.5	67.3	54.6	55.3	56.1	60.6	65.6
≤ 11	16.0	16.0	15.9	14.4	13.0	20.3	20.1	19.7	18.3	16.8
Avg.	34.1	34.3	34.5	35.9	37.3	32.6	32.8	33.2	34.5	36.0

Lastly, observe that the duration of product #2 is more than twice as long as that of product #1, except for the very attractive client rate scenario. This is due to greater longer maturity buckets and smaller shorter maturity buckets. As has been stated before, the clients of product #2 care less about the client rate. As a consequence, many variables are not included in the models, which makes the simulations less volatile. Even with the continuing historical downward trend in the simulations for the volume of product #2, the smaller uncertainty translates to greater buckets for longer maturities and a larger average duration.

4.7 Evaluation

In this section I evaluate the final model of this research. Firstly, I compare the buckets that would be obtained on the historical sample with buckets from a variant of the model which uses the historical values of explanatory variables. This shows that both generally result in similar buckets, except for one period where the client rate is extremely attractive. Secondly, I compare the model with an alternative and much simpler model, the random walk. This shows that the model does not provide much value for products with an unattractive client rate, because my model and the random walk result in nearly identical simulations and buckets. However, the model is advantageous for products with a more attractive client rate, because it takes into the effects of the client rate, while the random walk does not account for anything.

4.7.1 Historical sample

In this section, I create buckets for two periods in the historical sample with the VaR and the liquidity constraint. This shows that the former is dependent on the sample period, while the latter is not. I also create buckets with a variant of the model of this research and compare these with the historical buckets.

The historical buckets, further referred to as the buckets of model H, are obtained from the actual historical realisation of the deposit volume in the sample period of December 19, 2008 to March 28, 2013 (1115 days). The buckets of H are the optimal buckets for this sample, since they are determined from the actual historical volume. They result in the maximum profit, while always meeting the liquidity demand precisely.

Instead of simulations of the explanatory variables, the variant of the model of this research uses their historical realisation to simulate the transformed deposits and withdrawals. The 1000 simulations of the deposit volume that are created in this way, can be regarded as the optimal performance of the model of this research, since the volume is simulated with the use of the actual historical explanatory variables. The buckets of this model are further referred to as the buckets of model MH.

Table 15 and 16 show the buckets of H and MH for product #1 and #2, respectively. The buckets are estimated with the liquidity constraint and VaR over two samples. The number after the model abbreviation indicates the sample. The first sample runs from December 19, 2008 to November 22, 2011 (763 days) and the second sample runs from April 27, 2010 to March 28, 2013 (763 days). April 27, 2010 is the day on which product #1 reaches its maximum volume. Because both samples are about one month shorter than three years, the greatest available maturity is not 48 months, but only 36 months. Figure 13a and 13b in Appendix D show the 5%, 10%, 50%, 90%, and 95% quantiles of the simulations of MH together with the historical volume of product #1 and #2.

Firstly, the maturity profile of the historical VaR buckets of product #1 is greatly dependent on the estimation sample. Observe the odd maturity profile of the VaR H.1 buckets for product #1 in Table 15. Buckets with a short maturity have a small weight; several buckets with a medium maturity have no weight at all; and the bucket with the longest maturity has an extremely large weight. This extreme maturity profile is caused by the historical deposit volume as can be seen in Figure 1a. The volume declines marginally in the beginning of the sample, then it rises quickly and it does not become lower than the starting volume thereafter. Such an extreme maturity profile can be seen as a strength or weakness of the VaR bucket estimation. A strength, because if the bank foresees such a volume increase, perhaps due to a planned client rate increase, it can maximally profit by investing in bonds with long maturities. A weakness, because this specific maturity profile only holds for the savings that enter the bank at a specific point in time. Deposits that are attracted at another time, have an entirely different maturity profile. For example, observe the VaR H.2 buckets, which are based on the second sample that starts 352 days later. The short and medium maturity buckets have increased; there are no maturities with a zero weight; and the weight of the greatest maturity bucket is a lot smaller. The average duration is also significantly shorter compared to the VaR H.1 buckets, which is due to the great weight shift to maturities shorter than one year. Clearly, the start date of the sample makes a big difference for the historical VaR buckets of product #1. So, the maturity profile must constantly be observed carefully and new savings deposits must be invested accordingly. This makes the management of the VaR buckets labour intensive.

Secondly, Table 15 shows that the liquidity constraint buckets are a lot more consistent than VaR buckets across different samples. The liquidity constraint buckets are more consistent, because the relative volume changes are determined over multiple periods with equally many starting volumes. This is in great contrast to the VaR buckets, which are evaluated against a single starting volume and this makes them greatly dependent on (the starting point of) the sample period. Still, also the liquidity constraint

Table 15: The VaR and liquidity constraint bucket weights in percentages for maturities (Mat.) in months for savings product #1 for model H and MH. The H buckets are estimated from the single actual historical volume. MH represents the model of this research, but with the actual historical explanatory variables, instead of simulations, as input for the deposits and withdrawals model. The penultimate row (≤ 11) shows the sum of buckets with a maturity equal to or shorter than 11 months. The average duration (Avg.) is presented in the last row. The number suffix indicates the sample on which the buckets are estimated; sample 1 runs from March 28, 2008 to November 22, 2011 (763 days) and sample 2 runs from April 27, 2010 to March 28, 2013 (763 days). The buckets of MH are estimated from 1000 simulations.

	VaR				Liquidity constraint			
Mat.	H.1	H.2	MH.1	MH.2	H.1	H.2	MH.1	MH.2
0	0.7	8.0	4.4	8.2	10.4	10.4	12.9	13.9
1	1.3	3.4	5.8	5.0	4.9	4.9	5.7	6.4
2	0.6	1.2	5.7	3.8	3.7	4.1	4.4	4.6
3	2.0	2.6	5.0	4.1	3.7	5.0	4.1	4.3
4	1.0	2.2	4.0	3.1	3.3	4.0	3.9	3.9
5	0.9	2.7	3.3	2.8	2.7	2.8	2.6	3.5
6	0.1	2.8	2.1	3.1	2.9	3.3	2.4	2.9
7	0.0	6.0	0.3	3.5	2.7	2.5	2.6	3.0
8	0.0	4.5	0.0	1.6	3.1	2.5	2.4	2.4
9	0.0	3.3	0.0	1.3	2.1	2.5	2.3	2.1
10	0.0	2.9	0.0	2.6	3.7	2.0	1.9	2.2
11	0.0	2.8	0.0	2.6	3.1	3.1	1.3	2.2
12	0.0	26.3	0.0	22.4	13.0	21.8	12.5	18.5
24	0.0	5.4	5.3	14.9	0.0	5.4	0.0	8.9
36	93.4	25.8	64.3	21.2	40.9	25.8	40.9	21.2
≤ 11	6.6	42.5	30.4	41.5	46.1	47.0	46.6	51.4
Avg.	33.8	16.0	25.2	15.6	18.3	15.1	17.8	13.8

Table 16: The VaR and liquidity constraint bucket weights in percentages for maturities (Mat.) in months for savings product #2 for model H and MH. The H buckets are estimated from the single actual historical volume. MH represents the model of this research, but with the actual historical explanatory variables, instead of simulations, as input for the deposits and withdrawals model. The penultimate row (≤ 11) shows the sum of buckets with a maturity equal to or shorter than 11 months. The average duration (Avg.) is presented in the last row. The number suffix indicates the sample on which the buckets are estimated; sample 1 runs from March 28, 2008 to November 22, 2011 (763 days) and sample 2 runs from April 27, 2010 to March 28, 2013 (763 days). The buckets of MH are estimated from 1000 simulations.

	VaR				Liquidity constraint			
Mat.	H.1	H.2	MH.1	MH.2	H.1	H.2	MH.1	MH.2
0	0.5	0.2	1.9	1.9	4.5	3.9	4.7	4.8
1	2.7	0.2	1.8	1.8	1.9	1.7	2.1	2.1
2	1.6	0.2	1.6	1.5	1.3	1.1	1.5	1.7
3	1.1	0.5	1.3	1.5	1.1	0.8	1.3	1.5
4	1.2	0.4	1.1	1.2	0.5	0.6	1.2	1.3
5	0.3	0.5	1.3	1.2	0.9	0.5	1.0	1.3
6	1.2	0.5	1.1	1.0	0.8	1.3	1.1	1.1
7	1.0	0.3	0.7	1.2	1.2	0.3	0.9	1.1
8	0.7	0.2	0.8	0.4	0.2	0.8	0.9	1.1
9	0.0	0.1	0.7	0.3	0.0	0.7	0.6	1.0
10	0.1	1.6	0.6	0.8	0.2	0.6	0.8	1.0
11	0.0	1.0	0.6	0.8	0.4	0.7	0.5	1.2
12	0.4	11.3	3.9	10.2	0.4	8.2	6.0	8.8
24	5.4	7.5	8.9	9.7	4.9	4.0	3.9	5.8
36	83.6	75.5	73.6	66.4	81.7	74.8	73.5	66.1
≤ 11	10.6	5.7	13.6	13.7	13.0	13.0	16.6	19.2
Avg.	31.8	30.7	29.7	28.0	31.0	29.3	28.7	27.0

buckets are affected by the extreme volume increase in early 2010. The 24-month bucket of H.1 is zero, while it is not for the second sample. The most negative relative volume change in the first sample occurs between the maximum volume at April 27, 2010 and the end of the first sample at November 22, 2011. Since this period is shorter than two years, the 2-year and 3-year liquidity constraint are identical and this gives the 2-year bucket a zero weight. In the second sample, the liquidity constraint for the 3-year period is larger than for the 2-year period, because the period from the top to the bottom at the end of the sample is longer than two years. Hence, the 2-year bucket is non-zero for H.2. The more consistent weights of the liquidity constraint can be used across several time periods, which makes the method more reliable and less labour intensive to manage, compared with the VaR buckets. However, the increased certainty comes at a cost. The average duration of the VaR buckets can occasionally be more than a year longer. The bank should evaluate this risk reward trade-off.

In the historical sample, product #1 experiences an extreme volume increase, but the development of product #2 is more like a gradual decline, see Figure 1c. Therefore, the starting point of the sample has less influence on the VaR buckets and also the liquidity constraint buckets are more consistent than the equivalent buckets of product #1.

Lastly, the MH buckets are created with the use of historical values of the relative spread and other explanatory variables as input for the deposits and withdrawals model. The resulting volume simulations largely show the correct pattern, because the true historical values are used, but due to uncertainty in the deposits and withdrawals model, there is also a certain dispersion around the actual historical value. The quantiles of the volume simulations of product #2 (Figure 13b) contain the historical value nicely, but those of product #1 (Figure 13a) show a less pronounced increase and decrease compared with the real volume. Generally, the buckets of H and MH are close, especially the liquidity constraint buckets and all buckets of product #2. However, the VaR H.1 and MH.1 buckets of product #1 show that even with the actual value of explanatory variables, it can be difficult to obtain the actual buckets. The pattern of buckets with a zero weight is intact, but the average duration differs by more than eight months. These buckets differ the most, because the relation between the volume and relative spread is unable to capture the increase in early 2010 exactly. Also note that most MH buckets are more conservative than the H buckets, but this is not surprising, since the historical buckets are based on one single historical volume. The MH buckets, on the other hand, are based on the 5% quantiles of the VaR and liquidity constraint estimates. While the buckets of H are optimal, the MH buckets could be seen as the optimally achievable buckets from this model, because the correct historical values are used. With that in mind, the extra profit advantage of the VaR buckets in comparison with the liquidity constraint diminishes. The average duration difference between the VaR and liquidity constraint H.1 buckets is about 15.5 months, while the difference between the MH.1 buckets of both methods is 'just' 7.4 months.

Concluding, this section shows that the VaR buckets can be greatly dependent on the estimation sample; deposits that are attracted at another time might have an entirely different maturity profile. The liquidity constraint weights are much less dependent on the sample, because they are determined over multiple periods. On the one hand, the VaR buckets are valuable, because they allow the bank to foresee the effect of a client rate increase on the maturity profile. On the other hand, the always changing maturity profile must constantly be observed carefully and that is labour intensive. Furthermore, the optimally achievable MH buckets show that the VaR buckets are more profitable than the liquidity constraint buckets, but not as much as the H buckets indicate. Although generally the quantiles and buckets of the MH model are close to the historical values, even with the correct historical explanatory variables, it is not always possible to retrieve the exact buckets. This indicates that there is some room for improvement of the model.

4.7.2 Alternative model

In this section I evaluate the buckets of the model of this research, with the buckets that are created with an alternative and much simpler model, the random walk.

The to be evaluated model of this research uses the following parameters for the client rate model: $t_n = 10$, $b = 2\sigma$, $b_r = 1.4\sigma$ and $o = \mu$. Here, μ and σ are the mean and the standard deviation of the relative spread in the historical sample, respectively. The value for b is chosen such that the boundaries of the simulated relative spread are similar to that of the historical relative spread. The buckets of this model will be referred to as the buckets of model M.

An alternative and very simple, but often difficult model to beat, is the random walk model. Such a model for the deposit volume is actually not that bad. The literature mentions to use an AR model and analysis of the historical data shows that the volume of tomorrow is strongly correlated to that of today, resulting in an AR-coefficient extremely close to 1, see Table 1. To prevent negative values in the simulation, the volume is first transformed with the natural logarithm. For product #1, the AR(1) model is estimated with GARCH and for product #2 OLS is used. The model for product #2 additionally includes the dummy for the outlier and models for both product #1 and #2 account for accrued interest. The buckets of the random walk model will be referred to as the buckets of model RW.

Table 17 and 18 contain the buckets of model M and RW for product #1 and

#2, respectively. The buckets are estimated with the 5% quantile of the VaR and the liquidity constraint estimates of 1000 simulations. Each simulation runs five years (1306 days) into the future, and from this 5-year period, two partly overlapping subsamples of about three years are created. The first subsample runs from March 28, 2013 to March 30, 2016 (785 days) and the second runs from March 30, 2015 to March 30, 2018 (785 days). Because both subsamples have a length of approximately three years, the greatest available maturity is 36 months instead of 48 months. Figure 13c to 13f in Appendix D depict the 5%, 10%, 50%, 90% and 95% quantiles of the simulations of model M and RW for product #1 and #2.

Figure 13c and 13d show that the quantiles of the product #1 volume simulations of model M and RW are quite different. The quantiles of model M show little dispersion and a declining volume for the first 2 to 2.5 years. After that period, there is a greater dispersion for increasing volume simulations. This can be explained by the relative spread, which is an important variable for model M for product #1. The relative spread is positively related to the deposits and the threshold ensures an even greater effect for a large positive relative spread. The relative spread is low at the start of all simulations, which causes a declining volume. Due to the passive nature of the client rate model, it takes about 2.5 years before the relative spread is high in several simulations. This can be seen from the increasing 90% and 95% quantiles. Since the relative spread is incorporated in the GARCH model with a threshold, a high relative spread also causes a greater dispersion. The quantiles of model RW show a slowly declining volume with a greater dispersion compared with model M. Especially the rough and serrated 95% quantile shows that several simulations predict extremely large deposit volumes. Due to the small number of parameters that model RW incorporates, its residuals are larger than those of model M, which causes a greater dispersion. The slowly declining volume can be explained by the slightly negative residuals. Although model M has a better foundation with logical variables, I find the quantiles of model RW more plausible. The great certainty with which the declining volume of product #1 is predicted by model M is questionable. In reality the bank does not only adjust the client rate if the relative spread is outside the boundaries. With one or several quick client rate increases, we can expect a much higher volume than is predicted by the 90% and 95% quantiles. So, the 90% and 95% quantile should be higher. This can be accomplished by extending the client rate model with random client rate changes.

Unlike the different upper quantiles, the lower quantiles differ not that much and this can also be seen in the similar VaR buckets. Further comparison of the M and RW buckets, reveals that the liquidity constraint buckets of product #1 differ the most. The overnight bucket (0 months) of the random walk model is about 8 to 10 percent points

Table 17: The VaR and liquidity constraint bucket weights in percentages for maturities (Mat.) in months for savings product #1 for model M and RW. M represents the model created in this research with $t_n = 10$, $b = 2\sigma$, $b_r = 1.4\sigma$ and $o = \mu$ as parameters for the client rate model. Here, μ (0.007) and σ (0.159) are the mean and the standard deviation of the relative spread, respectively, in the historical sample (December 19, 2008 to March 28, 2013). RW represents the random walk model. The penultimate row (≤ 11) shows the sum of buckets with a maturity equal to or shorter than 11 months. The average duration (Avg.) is presented in the last row. The number suffix indicates the sample on which the buckets are estimated; sample 1 runs from March 28, 2013 to March 30, 2016 (785 days) and sample 2 runs from March 30, 2015 to March 30, 2018 (785 days). All buckets are estimated with 1000 simulations.

	VaR				Liquidity constraint			
Mat.	M.1	M.2	RW.1	RW.2	M.1	M.2	RW.1	RW.2
0	8.0	7.4	6.2	8.0	13.2	13.2	21.4	23.0
1	6.2	5.5	6.1	6.4	6.4	6.6	7.2	7.8
2	4.9	4.7	5.9	5.2	4.9	4.6	4.6	4.1
3	4.0	4.2	4.2	4.0	3.9	4.2	3.1	2.6
4	4.0	3.5	4.2	3.0	3.8	3.7	2.8	2.5
5	3.5	3.4	3.6	2.5	3.3	2.9	2.4	2.5
6	3.7	3.3	3.2	3.6	3.3	3.2	2.7	2.4
7	3.1	3.5	2.8	2.6	3.0	2.7	2.3	1.8
8	2.9	3.1	2.4	3.2	2.6	2.7	2.0	2.3
9	1.2	1.5	2.0	1.7	2.6	2.4	2.2	2.0
10	2.8	2.3	2.1	2.8	2.0	2.4	1.9	2.3
11	2.6	2.3	2.5	1.7	2.3	2.4	1.2	2.1
12	21.5	22.6	18.8	19.0	18.9	19.5	15.9	14.5
24	12.2	13.0	11.6	10.4	10.8	10.2	7.0	6.4
36	19.3	19.7	24.5	25.9	19.0	19.3	23.2	23.8
≤ 11	47.0	44.8	45.1	44.7	51.2	50.9	53.9	55.4
Avg.	14.4	14.8	15.8	15.9	13.6	13.7	13.4	13.4

Table 18: The VaR and liquidity constraint bucket weights in percentages for maturities (Mat.) in months for savings product #2 for model M and RW. M represents the model created in this research with $t_n = 10$, $b = 2\sigma$, $b_r = 1.4\sigma$ and $o = \mu$ as parameters for the client rate model. Here, μ (-0.219) and σ (0.162) are the mean and the standard deviation of the relative spread, respectively, in the historical sample (December 19, 2008 to March 28, 2013). RW represents the random walk model. The penultimate row (≤ 11) shows the sum of buckets with a maturity equal to or shorter than 11 months. The average duration (Avg.) is presented in the last row. The number suffix indicates the sample on which the buckets are estimated; sample 1 runs from March 28, 2013 to March 30, 2016 (785 days) and sample 2 runs from March 30, 2015 to March 30, 2018 (785 days). All buckets are estimated with 1000 simulations.

	VaR				Liquidity constraint			
Mat.	M.1	M.2	RW.1	RW.2	M.1	M.2	RW.1	RW.2
0	2.1	1.9	2.4	2.1	4.8	4.9	4.7	4.7
1	2.1	1.9	2.2	2.0	2.1	2.1	2.0	2.1
2	1.6	1.7	1.8	1.8	1.7	1.6	1.8	1.7
3	1.5	1.7	1.4	1.6	1.6	1.6	1.8	1.7
4	1.3	1.4	1.0	1.4	1.4	1.5	1.4	1.5
5	1.4	1.3	1.3	1.3	1.1	1.4	1.3	1.4
6	1.3	1.4	1.0	1.2	1.1	1.2	1.0	1.2
7	1.1	1.1	1.1	1.1	1.2	1.1	1.1	1.1
8	1.1	0.9	1.3	1.0	1.0	1.1	1.0	1.0
9	0.6	0.5	0.5	0.6	1.0	1.1	1.0	0.9
10	0.7	0.6	0.7	0.5	0.9	0.8	0.9	0.9
11	1.2	0.9	1.0	1.0	0.9	1.0	1.2	1.0
12	10.5	10.5	10.5	10.3	9.5	9.7	9.5	9.4
24	8.5	9.7	8.4	8.8	7.1	7.0	6.6	6.4
36	64.9	64.5	65.4	65.4	64.5	64.1	64.8	64.9
≤ 11	16.0	15.3	15.7	15.5	18.9	19.2	19.1	19.2
Avg.	27.4	27.5	27.5	27.5	26.8	26.7	26.8	26.8

larger than the respective bucket of model M. Due to the small number of regressors, the residuals in the random walk model are larger. This causes greater fluctuations in the volume simulations, and subsequently a large overnight bucket. The extra weight for the overnight bucket is compensated by the buckets with a longer maturity and the average durations differ not much.

The buckets of model M and RW for product #2 differ marginally and also the average duration is almost identical. This can be declared by the fact that both models are much alike. Contradictory to product #1, product #2 is not very dependent of the relative spread and its model does not contain thresholds. The model largely consists of a constant and dummies. Also Figure 13e and 13f show that the simulations of model M and RW are very much alike.

In Section 4.7.1 we have seen that the VaR buckets can be strongly dependent of the sample on which they are estimated. This particularly holds for product #1. However, the buckets in Table 17 and 18 do not differ much across the samples. That is because these simulations share no specific client rate scenario; each simulation is independent. Consequently, the simulations have no simultaneous volume increase and all maturities have a non-zero weight. This shows that also the VaR buckets can be more or less consistent, if the simulations share no specific scenario.

Given the differences between model M and RW for product #1 and #2, I conclude that the additional value of my model depends on the nature of the product. For products that are marginally dependent of the relative spread, like product #2, the model has no clear additional value. For these products, the simulations and buckets of my model and the random walk are almost identical. For products that are more attractive and do depend on the relative spread, my model is advantageous. My model takes into account this attractiveness, while the random walk does not account for anything. Despite this fact, the buckets of my model and the random walk in Table 17 generally differ little. However, this is just due to the lack of a clear client rate scenario. The boundaries in this client rate model are set such that they are similar to the historical sample. However, these boundaries are very wide and the regime can therefore not be regarded as attractive or unattractive. When we compare the random walk buckets with a clearly attractive client rate, $z = 1.4$ in Table 13 in Section 4.6.2, the differences are much larger.

5 Conclusion

Liquidity management is an important aspect of the risk management at many companies. Even successful companies can experience the drastic consequences of inadequate liquidity management in the short run. A frequently used technique to cope with liquidity management is duration matching, which aims to offset the magnitude and timing of incoming and outgoing cash flows. However, duration matching is not straightforward for savings deposits, because these have no fixed maturity. The goal of this research is therefore to find a way in which the maturity of savings deposits can be correctly determined.

This research aims to answer that question by specifying models for the deposits and withdrawals in order to create realistic simulations for the deposit volume. The models include logical and important variables, such as the relative spread and the nationwide deposit volume, and important events, such as client rate changes. Furthermore, a simplified version of the client rate adjustment process is incorporated to observe and regulate the effect of different client rate regimes. Afterwards, VaR and the liquidity constraint of Bardenhewer (2007) are used to determine the maturity from the deposit volume simulations.

Comparison of the VaR and liquidity constraint buckets reveals several differences. Firstly, the liquidity constraint buckets are stricter, because multiple time periods are evaluated, instead of a single one for the VaR buckets. This means that the cumulative weight of liquidity constraint buckets up to a certain maturity is at least as high as that of the VaR buckets. Consequently, the average duration of the VaR buckets is higher, and so is the profit, but also the risk. Secondly, the liquidity constraint buckets are more consistent across samples. The VaR buckets are only consistent across samples if the simulations are independent and share no specific scenario. If they do share a scenario, the VaR buckets are not very consistent across samples. However, this property can be valuable. The VaR buckets allow the bank to foresee the effect of a client rate scenario on the maturity profile, which can be a lot more profitable than the average liquidity constraint profile. On the other hand, due to the changing maturity profiles, VaR is more labour intensive. Concluding, the liquidity constraint is stricter and more consistent and yields less profit and risk. The bank should evaluate whether the extra profit of the VaR buckets, is worth the extra time and risk.

The final model of this research is evaluated by comparing it with an alternative and much simpler model, the random walk. Given the differences between the final model and the random walk model for product #1 and #2, I conclude that the additional value of my model depends on the nature of the product. For products that are marginally dependent of the relative spread, like product #2, my model has no clear additional value. For these products, the simulations and buckets are almost identical to those of the random walk. For products that are more attractive and do depend on the relative spread, my model is advantageous. My model takes into account the attractiveness of the product, while the random walk does not account for anything. Despite this fact, the buckets of my model and the random walk appear similar. However, this is just due to the lack of a clear client rate scenario. The boundaries in the client rate model are set such that they are similar to the historical sample. However, these boundaries are very wide and the regime can therefore not be regarded as attractive or unattractive. The differences are much larger, if the random walk buckets are compared with a clearly attractive client rate, which confirms the additional value of the model.

However, the model is not perfect and there is certainly room for improvement. Firstly, a comparison of the historical buckets and the buckets of the model with historical explanatory variables shows that the buckets are generally similar, but not always. Especially the increase in early 2010 was not captured completely. A closer examination of the relation between the deposits and relative spread might solve this. Secondly, only nine interest rate decreases are present in the data set and not a single interest rate increase. More data should be examined to get a clear view of the effect of client rate changes. That might also reveal whether the magnitude of the change and the level of the relative spread at the time of change are important. Additionally, this model makes no assumptions on the reaction of other banks to a client rate change. It could be argued that a change of one bank triggers changes at other banks. Lastly, the quantiles of the volume simulations are questionable. The incorporation of random client rate changes can make the client rate model less passive and they might make the simulations more realistic.

Further research might differentiate between several client groups. In this research, every client is treated equally. However, it could be that wealthy and less wealthy persons act differently. In addition to that, wealthy clients have a much greater effect on the deposit volume due to their larger balance. In fact, data analysis shows that some large deposit volume changes are due to single wealthy clients. Wealthy clients are therefore more risky.

Lastly, the bank should evaluate the profitability of a few decisions. The model of this research predicts that the average duration of a very attractive client rate regime is significantly longer. The bank should evaluate whether the increased profit of a higher average duration outweighs the increased cost of a higher client rate. Similarly, the bank should evaluate whether it should use VaR or liquidity constraint buckets. The former method results in a larger average duration, while the latter entails less risk.

References

Bardenhewer, M. (2007). Modelling non-maturing products. In L. Matz & P. Neu

(Eds.), *Liquidity risk measurement and management* (pp. 220–256). John Wiley and Sons.

Cox, J., Ingersoll, J., & Ross, S. (1985). A theory of the term structure of interest rates. *Econometrica*, *53*(2), 385–407.

De Pooter, M. (2007). Examining the Nelson-Siegel class of term structure models. *Tinbergen Institute Discussion Papers. Tinbergen Institute*.

DeWachter, H., Lyrio, M., & Maes, K. (2006). A multi-factor model for the valuation and risk management of demand deposits. *National Bank of Belgium*, *Working Paper No. 83*.

Diebold, F., & Li, C. (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics*, *130*(2), 337–364.

DNB. (2012). Uitgangspunten voor het Internal Liquidity Adequacy Assessment Process (ILAAP). *De Nederlandsche Bank*.

Ellis, D., & Jordan, J. (2001). The evaluation of credit union non-maturity deposits. *National Credit Union Administration, Working Paper*.

Frauendorfer, K., & Schürle, M. (2003). Management of non-maturing deposits by multistage stochastic programming. *European Journal of Operational Research*, *151*(3), 602–616.

Frauendorfer, K., & Schürle, M. (2007). Dynamic modelling and optimization of non-maturing accounts. In L. Matz & P. Neu (Eds.), *Liquidity risk measurement and management* (pp. 327–359). John Wiley and Sons.

Hamilton, J. (1994). State-space models. In R. Engle & D. McFadden (Eds.), *Handbook of econometrics* (Vol. IV, pp. 3039–3080). Elsevier Science.

Heath, D., Jarrow, R., & Morton, A. (1992). Bond pricing and the term structure of interest rates. *Econometrica*, *60*(1), 77–105.

Hutchison, D., & Pennacchi, G. (1996). Measuring rents and interest rate risk in imperfect financial markets. *The Journal of Financial and Quantitative Analysis*, *31*(3), 399–417.

Jarrow, R., & Van Deventer, D. (1998). The arbitrage-free valuation and hedging of demand deposits and credit card loans. *Journal of Banking and Finance*, *22*(3), 249–272.

Kalkbrener, M., & Willing, J. (2004). Risk management of non-maturing liabilities. *Journal of Banking and Finance*, *28*(7), 1547–1568.

Kim, C.-J., & Nelson, C. (1999). *State-space models with regime switching.* MIT Press Books.

Maes, K., & Timmermans, T. (2005). Measuring the interest rate risk of Belgian regulated savings deposits. *Financial Stability Review, National Bank of Belgium*, *3*(1), 137–151.

Nelson, C., & Siegel, A. (1987). Parsimonious modeling of yield curves. *The Journal of Business*, *60*(4), 473–489.

Nyström, K. (2008). On deposit volumes and the valuation of non-maturing liabilities. *Journal of Economic Dynamics and Control*, *32*(3), 709–756.

O'Brien, J. (2000). Estimating the value and interest rate risk of interest-bearing transactions deposits. *Board of Governors of the Federal Reserve*.

OeNB. (2008). Guidelines on managing interest rate in the banking book. *Oesterreichische Nationalbank*.

Paraschiv, F., & Schürle, M. (2010). Modeling client rate and volumes of non-maturing accounts. *Institute for Operations Research and Computational Finance, University of St. Gallen*, *Working Paper*.

Politis, D., & Romano, J. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, *89*(428), 1303–1313.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, *5*(2), 177–188.

Wielhouwer, J. (2003). On the steady state of the replicating portfolio. *OR Spectrum*, *25*(3), 329–343.

Wilson, T. (1994). Optimal values. *Balance Sheet*, *3*(3), 13–20.

A Replicating portfolio

In this section I use the replicating portfolio to mimic the client rate offered by Allianz with the 1, 2, 3, 6 and 9-month Euribor rates and the 1, 2, 3, 4 and 5-year Euro Swap rate. The observation sample of the client rates runs from December 19, 2008 to March 28, 2013 (1115 observations). The observation sample of the replicating yields runs from January 1, 2002 to March 28, 2013 (2933 observations). The latter sample is much larger, which is necessary to construct long term moving averages as suggested in Bardenhewer (2007). First I apply the general static replicating portfolio and thereafter the extensions of Bardenhewer (2007).

A.1 Introduction

The replicating portfolio is constructed from several interest paying instruments in such a way that its yield mimics the client rate, except for a certain margin,

$$m_t = \sum_i r_{i,t} \cdot w_i - r_{cl,t}. \tag{36}$$

Here, the yield of the portfolio is calculated as the weighted average of the interest paying instruments. The weights for the individual interest paying instruments are determined with an optimisation criterion subject to one or more constraints. Possible objectives are the minimisation of the standard deviation of the margin,

$$\min_w \quad \sqrt{\mathrm{Var}[m_t]}, \tag{37}$$

or maximisation of the Sharpe ratio of the margin,

$$\max_w \quad \frac{\mathrm{E}[m_t]}{\sqrt{\mathrm{Var}[m_t]}}. \tag{38}$$

The Sharpe ratio is defined as the mean of the margin divided by the standard deviation of the margin. The sum of the weights must be equal to one,

$$\sum_i w_i = 1, \tag{39}$$

and often the weights are also required to be greater than zero individually,

$$w_i \geq 0 \quad \forall i. \tag{40}$$

This also implies that weights are equal to or smaller than one. The strategy is executed by investing the entire deposit volume in the bonds of the replicating portfolio according the weights.

With the static replicating portfolio approach, the weights are determined once over a historical period. The weights are not updated and remain the same for a long period

Table 19: Matlab code for static replicating portfolio

```
% rb    bond rate               [T x N] matrix
% ra    allianz client rate     [T x 1] vector
% w     weight                  [N x 1] vector
[T,N] = size(rb);
eq = @(w) -mean(rb*w-ra) / std(rb*w-ra);
w = fmincon(eq,ones(N,1)/N,[],[],ones(1,N),1,zeros(N,1),ones(N,1));
```

of time. Over time, the value of the replicating bonds and the deposit volume do change. However, neither of them has an effect on the weights of the replicating portfolio. The first is being accounted for by trading in the respective instruments such that the weights remain constant. The second is being accounted for by investing the additional funds or by selling the bonds according to the weights. When a bond expires, the released funds are reinvested at the maturity of the expiring bond.

To illustrate the basic static replicating approach, I construct a replicating portfolio from the 1, 2, 3, 6 and 9-month Euribor rates and the 1, 2, 3, 4 and 5-year Euro Swap rate. Although OeNB (2008) states that "*this method can be easily implemented, without any special software requirements*", I find that when the number of bonds increases, Excel has a hard time in finding a solution. I use the Matlab code in Table 19, because it does not give any problems and it converges quickly.

The optimisation of the static replicating portfolio results in a portfolio that is solely constructed of the longest maturity, the five-year bond. The yield of the replicating portfolio and the client rate are depicted in Figure 8. The Figure shows that the margin is far from stable, which is due to the fact that the portfolio is constructed of just one bond. A volatile margin is not attractive for the bank, as this reflects more risk. Additionally, the mean margin is -13 basis points, which makes it even more unattractive for the bank, because on average this yields a loss. Furthermore, this portfolio implies that cash flows of the deposits are best mimicked with the cash flows of a five-year bond and that is not realistic. All in all, the results of the static replicating portfolio are far from satisfactory.

A.2 Bardenhewer

So far, the static replicating portfolio approach is not quite appealing. Bardenhewer (2007) offers some extensions that might improve the model. He suggests to use a trend, moving averages, and a liquidity constraint.

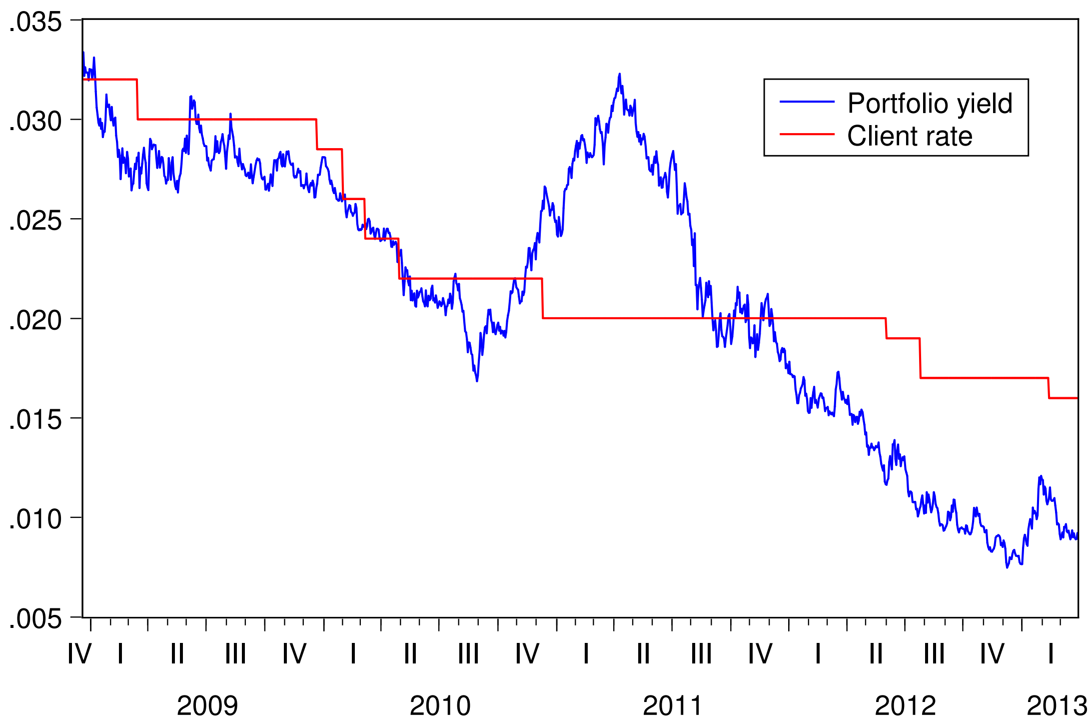


Figure 8: Time series of client rate and static replicating portfolio yield

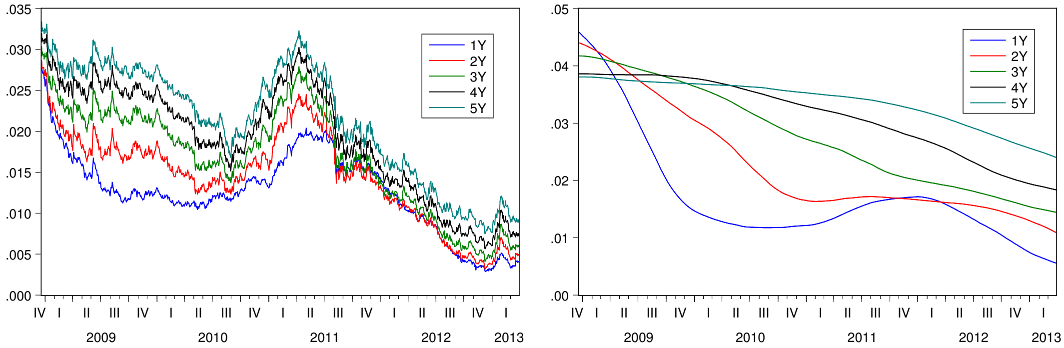
A.2.1 Trend

Bardenhewer (2007) finds that the volume should be decomposed in a trend and an unexpected part that is not explained by the trend. Only the trend component should be replicated with the portfolio and the unexpected part should be fully invested in the shortest maturity bond available. The shortest maturity bond functions as a buffer. For example, if the trend explains 90%, the remaining 10% is invested in the shortest maturity and the other weights are multiplied by 0.9. However, my volume data does not show a clear trend, nor is it about constant or stationary over time. Therefore, I decide to not incorporate the trend suggestion.

A.2.2 Moving average

Another addition of Bardenhewer (2007) is to implement moving averages of the replicating yields instead of the yields at a certain point in time. The moving average should be taken over the same period as the maturity of that bond, e.g. the one-month moving average of a bond with a maturity of one month, and a two-month moving average of a bond with a maturity of two months.

So why does it help to take moving averages? Figure 9a shows that in the 'normal' case, the slope of the yield curve does not fluctuate much. Consider an even less fluc-



(a) **(b)**

Figure 9: Time series of 'normal' yields in (a) and time series of moving average yields in (b).

tuating yield curve over time, e.g. a yield curve with always the same upward slope. The yields can still change over time, but only a parallel shift of the entire yield curve can occur. If we try to make a replicating portfolio with these yields, the highest yield will always get a weight of one. For, a combination of yields always results in a lower margin, with exactly the same volatility of the margin. Something similar happens with the real 'normal' yield curve. Alternatively, if we use moving averages, the slope of the yield curve is not constant any more, see Figure 9b. So, a portfolio of bonds might result in a better Sharpe ratio of the margin. In reality this approach can be implemented by renewing all bonds of every maturity at the same time. Every day bonds of each maturity should expire and be renewed. This implies that one should have at least 21 one-month bonds, 42 two-month bonds and so on, assuming that a month consists of 21 business days. In the literature, mostly monthly data is used. In that case only one one-month bond is needed, two two-month bonds and so on.

With the implementation of moving averages 27% is invested in two-year bonds and 73% is invested in five-year bonds. The yield of the replicating portfolio and the client rate is depicted in Figure 10. The margin is a lot more stable and moreover, it is clearly positive. On average the margin is plus 80 basis points. However, only a few bonds get a non-zero weight and the maturities of the bonds are still very high. Too high actually, because it is unlikely that savings stick with the bank for at least two years.

A.2.3 Liquidity constraint and market mix

Bardenhewer (2007) also finds that a lot of weight is assigned to long maturities and that might cause insufficient liquidity. He suggests to use a liquidity constraint that makes sure that there is always enough money to satisfy the demands of the customers. At least in the historical estimation period. To determine the weights with the liquidity constraint, first calculate the maximum relative historical decrease of the volume over

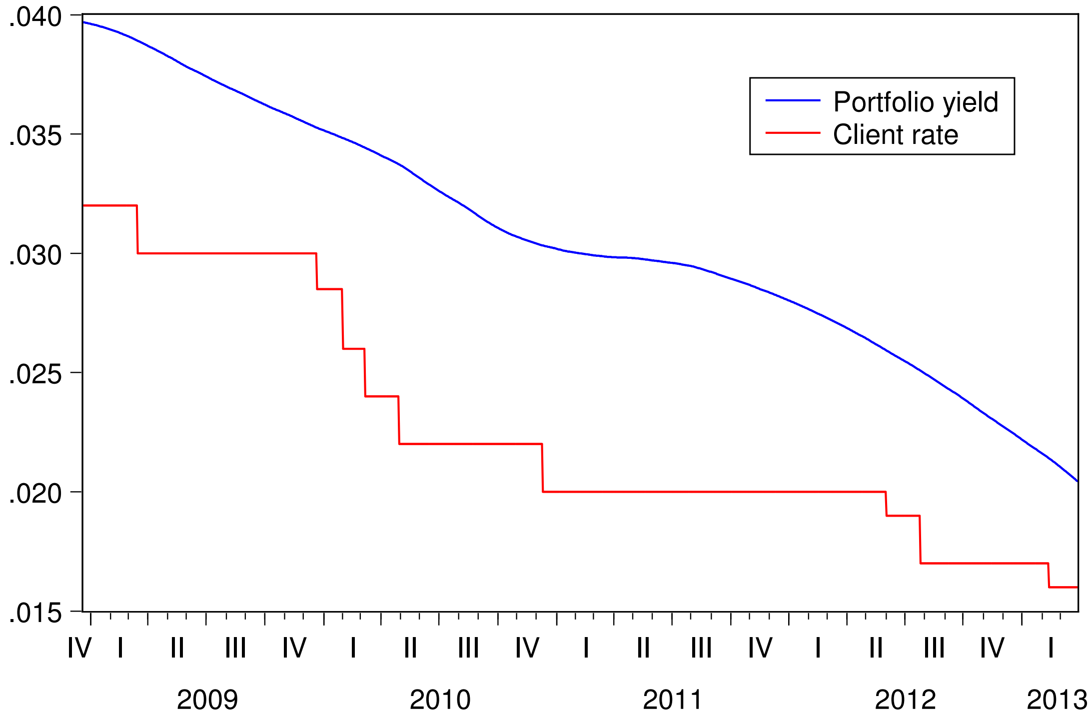


Figure 10: Time series of client rate and static replicating portfolio yield with moving average rates

each available maturity period. Note that the relative historical decrease over a long period is always greater than or equal to that of a shorter subsample period. Then, to determine the individual weights, take the increments of the decreases. For example, assume that the maximum historical relative decrease is 10% and 30% over a period of one month and two months respectively. The weights for the one-month and two-month bond are then 10% and 20%. Lastly, set the weight of the longest maturity such that all weights sum to 100%. The liquidity constraint weights are thus only influenced by the volume and not by the rate. This in contrast to the replicating portfolio approach, where the weights are solely determined by the rate.

To combine the weights determined by the volume and by the rate, Bardenhewer (2007) determines the 'market mix' and states: "*The market mix weight is chosen for each bucket such that the maximum cumulated weights are met.*" In this case the market mix weights are entirely determined by the liquidity constraint. The weights are depicted in panel A of Table 20 and it clearly shows the extreme differences between the rate and volume driven weights. The liquidity constraint weights are more equal and make a lot more sense, because also short maturities get a non-zero weight. Moreover, note that the market mix weights are entirely determined by the liquidity constrained. Should a bank trust the replicating portfolio approach blindly, it would experience severe liquidity problems.

A.3 Subsamples

Lastly, I examine the assumption of constant weights for the static replicating portfolio approach. The original observation sample is divided into two samples of respectively 557 and 558 observations. I use the moving average approach with the maximum Sharpe ratio as objective, together with the liquidity constraint to determine the market mix weights. This results in the weights depicted in panel B and C of Table 20. Observe that the weights of the replicating portfolio approach driven solely by the rates are not constant at all. The first subsample gives large weights to the three-year and four-year bonds, whereas on the whole sample it gives large weights to the one-year and five-year bonds. In the second subsample, the five-year bond even gets a weight of 100%. As stated earlier, this is not realistic. On the other hand, the liquidity constraint gives quite similar weights in both subsamples and the total sample.

A.4 Conclusion

The static replicating portfolio method puts a too large weight on long maturities and too few maturities get a non-zero weight. The implementation of moving averages might improve the method a little, but it does not solve all problems. The calculated weights result in a volatile and occasionally negative margin. Furthermore, the static replicating portfolio assumes that the weights are the same over the whole estimation period, but this is clearly not the case, see panel B and C of Table 20.

Personally I think that the replicating portfolio lacks a connection with the volume of the savings accounts and that it focuses too much on the deposit rate. Bardenhewer (2007) tries to account for the volume with the incorporation of a trend, but that is not suitable for my data set. Even with the trend addition, the connection with the volume remains weak.

The addition of the liquidity constraint seems to work very well, as the weights are more realistic and also more or less constant over time. The market mix weights in Table 20 are almost entirely determined by the constraint. This puts extra doubt in the value of the replicating portfolio and more confidence in the liquidity constraint. The liquidity constraint might as well be used on its own, without the replicating portfolio technique.

As a final remark, I should note that the current economical situation is quite different from the past. The current yield curve is extremely low, very close to zero. The literature assumes that the margin is positive, but currently that is not the case.

Table 20: Portfolio weights determined with moving average rates (1) and the liquidity constraint (2). Line (3) and (4) show the cumulated weights of both techniques. Line (5) and (6) show the (cumulated) weights of the market mix. All weights are rounded to integer percentage values. Panel A is estimated over the entire sample, panel B and C on the first and second half respectively.

Panel A: December 19, 2008 - March 28, 2013 (1115 obs.)

	1m	2m	3m	6m	9m	1y	2y	3y	4y	5y
(1) MA							27			73
(2) Constraint	10	4	4	11	8	7	22	5		28
(3) MA cum.							27	27	27	100
(4) Constraint cum.	10	14	18	30	38	45	67	72	72	100
(5) Market mix cum.	10	14	18	30	38	45	67	72	72	100
(6) Market mix	10	4	4	11	8	7	22	5		28

Panel B: December 19, 2008 - February 7, 2011 (557 obs.)

	1m	2m	3m	6m	9m	1y	2y	3y	4y	5y
(1) MA								72	28	
(2) Constraint	10	4	4	9	9	9	22	5		28
(3) MA cum.								72	100	100
(4) Constraint cum.	10	14	18	27	36	45	67	72	72	100
(5) Market mix cum.	10	14	18	27	36	45	67	72	100	100
(6) Market mix	10	4	4	9	9	9	22	5	28	

Panel C: February 8, 2011 - March 28, 2013 (558 obs.)

	1m	2m	3m	6m	9m	1y	2y	3y	4y	5y
(1) MA										100
(2) Constraint	9	5	4	11	8	7	12	1		43
(3) MA cum.										100
(4) Constraint cum.	9	14	18	30	38	45	56	57	57	100
(5) Market mix cum.	9	14	18	30	38	45	56	57	57	100
(6) Market mix	9	5	4	11	8	7	12	1		43

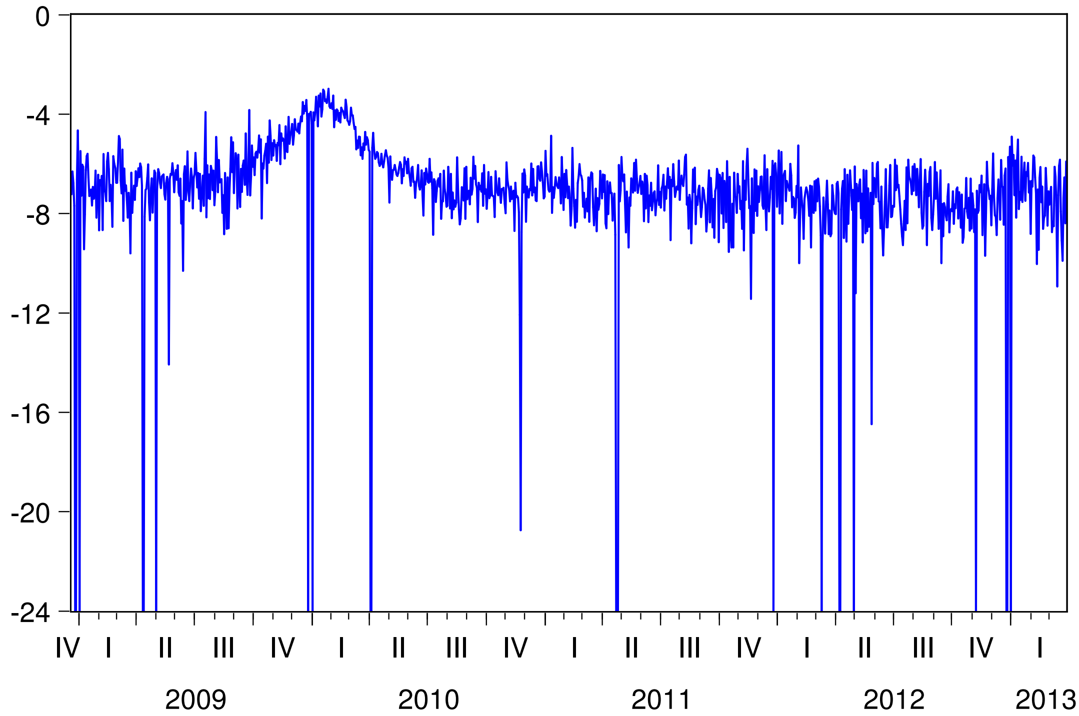


Figure 11: A time series of the relative deposits after being transformed with the natural logarithm. The -24 values are truncated $-\infty$ values.

B Alternative transformation for relative deposits

Section 3.2 uses (2) to transform the log-normally distributed relative deposits to a more normally distributed series, see Figure 2b. At first glance a more obvious approach might be to transform with the natural logarithm instead, so $D_{tr} = \ln(D_{rel})$. However, the natural logarithm yields a less intuitive result for the domain (0 to 0.05) of the relative deposits, see Figure 11. Large positive deposits would become small and negative, and small positive deposits would become large and negative. Hence, it does not conserve the shape as well as (2). Moreover, due to national holidays such as Christmas and Easter, some days have no deposits, for which $D_{rel} = 0$. As the relative deposits converge to zero, the natural log converges to the impractical $-\infty$, whereas (2) converges to zero, which can more easily be captured with a dummy variable.

C Effect of client rate adjustment

This section illustrates the effect of a client rate decline and a client rate increase with a simplified numerical example.

For this illustration I assume that a bank has 5000 accounts. These accounts have a balance of €100 and every day €0.25 is deposited on and withdrawn from each account.[6] A normal individual is not likely to deposit and withdraw the same amount on a given day, but each account might as well consist of multiple accounts and be used by many people, in which case it is valid. In this example, consumers are supposed to be either satisfied with the client rate or not. So, they cannot be extremely satisfied or only marginally. As a consequence, satisfied clients make daily deposits and withdrawals, and unsatisfied people are no client at all.

Now, we inspect a time period of 30 days, in which a client rate decrease occurs on day 10 and an increase on day 20. Due to the decrease on day 10, the entire balance of 48 accounts is withdrawn on the same day and these accounts also do not make their daily €0.25 deposits and withdrawals any more. From day 11 to day 17, respectively 29, 17, 10, 6, 3, 2 and 1 accounts follow suit. They empty their accounts and cease their daily deposits and withdrawals. The effect of the client rate decrease on the transformed deposits and withdrawals is illustrated in Figure 12. This shows the large spike in transformed withdrawals on day 10, caused by the discontinuation of the accounts. However, the effect of the ceased daily deposits on the transformed deposits is so marginal that it is not visible in the figure.

Suppose that the bank changes its mind and increases the client rate to the former higher level on day 20; this has the opposite effect. From day 20 to day 27, respectively 48, 29, 17, 10, 6, 3, 2 and 1 accounts redeposit their balance and from the following day they also make the daily transactions again. This has a clear effect on the deposits, but not on the withdrawals.

Concluding, with the assumptions from above, a client rate decrease has a significant effect on the withdrawals, but not on the deposits. Vice versa, a client rate increase has a significant effect on the deposits, but not on the withdrawals. Following this reasoning, the exponential decline variable should only be included for a decrease in the withdrawals model and for an increase in the deposits model.

[6] The values in this example, 5000 accounts, a balance of €100, and daily transactions of €0.25, are chosen such that the transformed deposits and withdrawals more or less correspond with the estimated values in Figure 5. The accounts that leave after a client rate change, respectively 48, 29, 17, 10, 6, 3, 2 and 1, follow an exponential decline pattern. Only the half life time of the exponential decline is much faster than estimated from the real data.

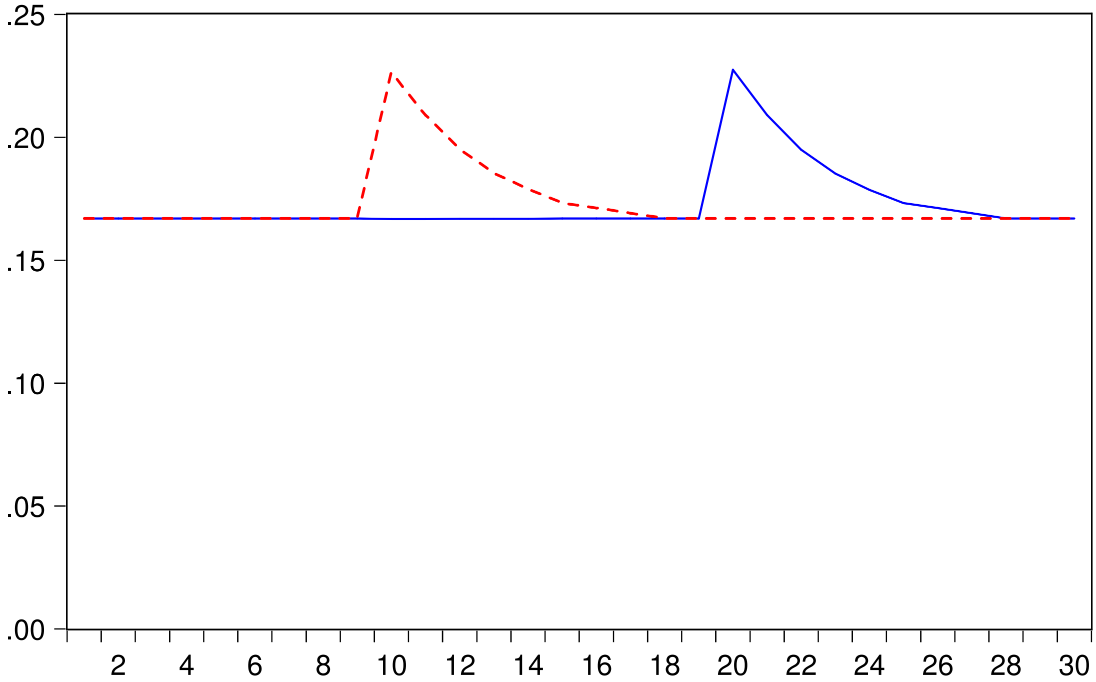


Figure 12: The transformed deposits (blue) and the transformed withdrawals (red) when a client rate decrease occurs on day 10 and a client rate increase on day 20.

D Quantiles of volume simulation

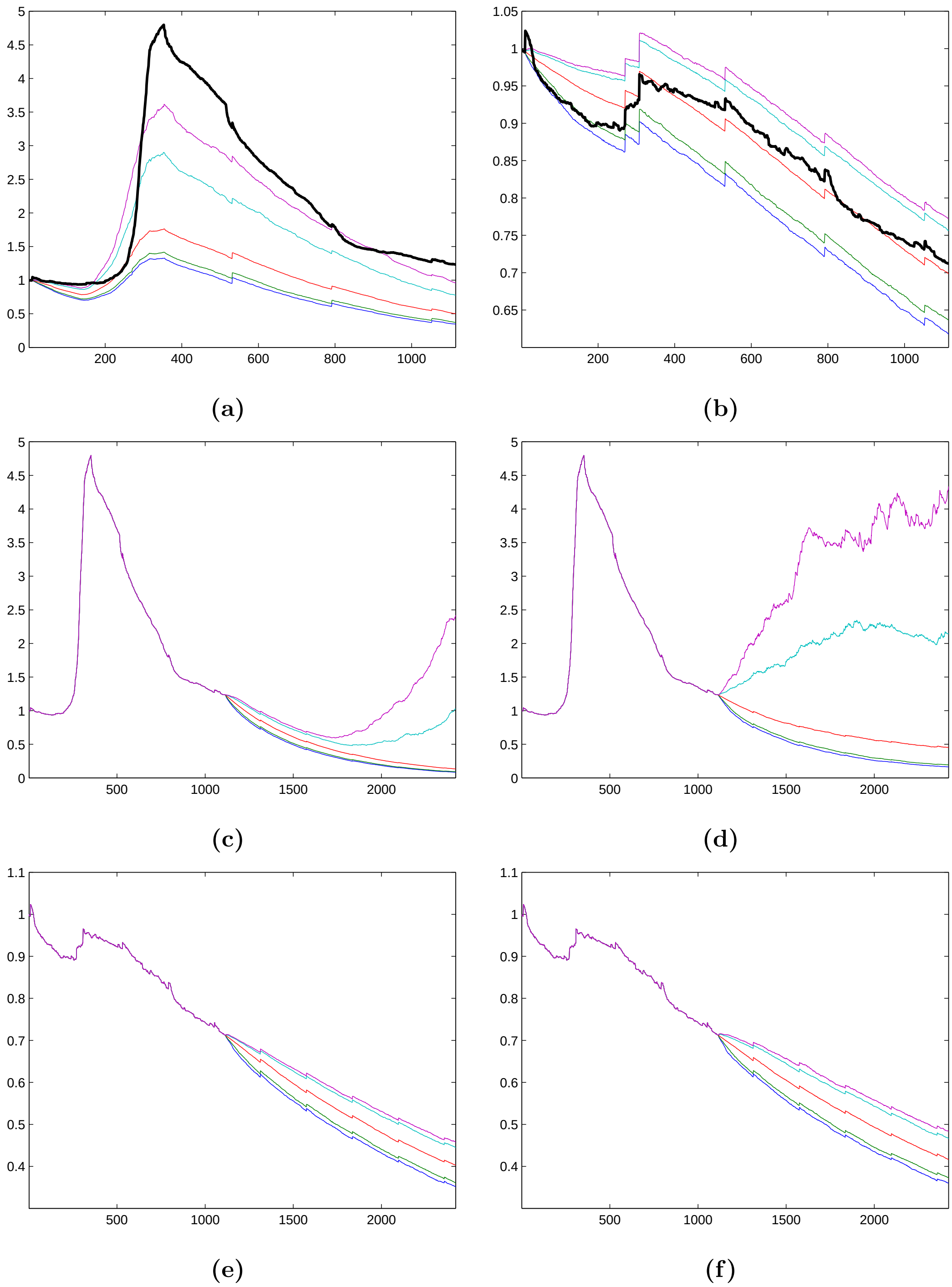


Figure 13: Figure (a) and (b) show the quantiles of the volume simulations of MH together with the actual historical volume in black, for product #1 and #2 respectively. Figure (c) and (d) show the quantiles of the volume simulations of M, for product #1 and #2 respectively, and figure (e) and (f) depict the quantiles of model RW. The printed quantiles are the 5%, 10%, 50%, 90% and 95% quantiles.