FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

MASTER THESIS

Bc. Tomáš Rusý

Stochastic Programming Problems in Asset–Liability Management

Department of Probability and Mathematical Statistics

Supervisor of the master thesis:	doc. RNDr. Ing. Miloš Kopa, Ph.D.
Study programme:	Mathematics
Study branch:	Probability, Mathematical Statistics and Econometrics

Prague 2017

I declare that I carried out this master thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In Prague, May 12, 2017 signature of the author

Title: Stochastic Programming Problems in Asset–Liability Management

Author: Bc. Tomáš Rusý

Department: Department of Probability and Mathematical Statistics

Supervisor: doc. RNDr. Ing. Miloš Kopa, Ph.D., Department of Probability and Mathematical Statistics

Abstract: The main objective of this thesis is to build a multi–stage stochastic program within an asset–liability management problem of a leasing company. At the beginning, the business model of such a company is introduced and the stochastic programming formulation is derived. Thereafter, three various risk constraints, namely the chance constraint, the Value–at–Risk constraint and the conditional Value–at–Risk constraint along with the second–order stochastic dominance constraint are applied to the model to control for riskiness of the optimal strategy. Their properties and their effects on the optimal decisions are thoroughly investigated, while various risk limits are considered. In order to obtain solutions of the problems, random elements in the model formulation had to be approximated by scenarios. The Hull – White model calibrated by a newly proposed method based on maximum likelihood estimation has been used to generate scenarios of future interest rates. In the end, the performances of the optimal solutions of the problems for unconsidered and unfavourable crisis scenarios were inspected. The used methodology of such a stress test has not yet been implemented in stochastic programming problems within an asset–liability management.

Keywords: asset–liability management, multi–stage stochastic programming, risk constraints, stochastic dominance, stress test

At this place, I would like to thank a number of people who significantly contributed to the development of my master thesis.

First, my acknowledgements go to my supervisor doc. RNDr. Ing. Miloš Kopa, Ph.D., who had answers to all of my questions and gave me a number of priceless pieces of advice which were essential to keep me on track with the thesis. I am also indebted to him for providing me with a full GAMS licence, which enabled me to solve the model and to reach desired results.

Second, a big thanks goes to the asset–liability management unit of Česká Spořitelna, a.s., especially to Ing. Michal Walos, who came up with the idea to focus on a business model of a leasing company, and to Ing. Martina Hejdová, who patiently explained me everything regarding the way how such a company works.

I would also like to acknowledge Dr. Sebastiano Vitali, Ph.D., for his advice on running GAMS from different software, which allowed me to perform the analysis in much larger extent than it would be otherwise possible.

Last, but not least I would like to thank my beautiful Eliška, whose love and presence have given me the all–needed motivation and perseverance to learn and improve which have been indispensable during my time at Charles University.

Thank you.

Contents

Introduction **2**

1 Introduction to a Multi–Stage Stochastic Programming **4**
1.1 Motivation Example 4
1.2 Multi–Stage Stochastic Linear Program 7

2 Asset–Liability Model of a Leasing Company **10**
2.1 Business Model of a Leasing Company 10
2.1.1 Model Variables 10
2.1.2 Benchmark Strategy 13
2.2 The Model as a Multi–Stage Stochastic Program 14
2.3 Management of Interest Rate Risk 17
2.3.1 Chance Constraint 17
2.3.2 Value–at–Risk Constraint 19
2.3.3 Conditional Value–at–Risk Constraint 20
2.3.4 Second–Order Stochastic Dominance Constraint 22

3 Scenario Generation in the ALM of a Leasing Company **25**
3.1 The Hull–White Model 25
3.1.1 Introduction of the Model 25
3.1.2 Change of Measure 27
3.1.3 Estimation of the Parameters 28
3.2 Demand for Loans of the Leasing Company 38

4 Model Analysis **41**
4.1 Model Specification and Parameter Settings 41
4.2 Results and Sensitivity Analysis 44
4.2.1 No–Risk Constraint Problem 44
4.2.2 Chance Constraint 48
4.2.3 Value–at–Risk Constraint 49
4.2.4 Conditional Value–at–Risk Constraint 51
4.2.5 Second–Order Stochastic Dominance Constraint 52

5 Stress Testing of Optimal Solutions **56**
5.1 Methodology of Stress Testing 56
5.2 Stress Test Analysis 58

Conclusion **62**

Bibliography **64**

List of Figures **67**

List of Tables **69**

Introduction

Stochastic programming is a quickly developing methodology, which finds its application in many different fields. Especially the area of financial planning and control has seen plenty of implementations of stochastic programming models. The first such an application arose in 1986, when Kusy and Ziemba published their famous paper dealing with a bank asset–liability model (Kusy and Ziemba, 1986). Even though they did not consider any risk constraints in their program, the way they formulated the dynamics of cash flows and government regulations showed how flexible this technique can be. Yet, at that time the computational power did not allow to solve such problems with a reasonably dense scenario tree.

That changed with the advent of the 21st century. The arrival of new technologies aided by the availability of software implied that researchers were able to calculate more complex and more realistic problems. This caused a boom of stochastic programming models, which quickly emerged in the literature. Many of these applications were done in a cooperation with commercial sponsors. We should mention at least the most famous contribution of Carino et al. (1994), who presented a very first asset–liability model for an insurance company. The same area was subject of study of Hoyland (1998), while Dert (1995) wrote a dissertation on the use of stochastic programming within an asset–liability model of a pension fund. A case study on a Czech pension fund was worked out by Dupačová and Polívka (2009), which dates as the first application of stochastic programming within asset–liability management in the Czech environment.

In this thesis, we will focus on formulating a stochastic program of an asset–liability model of a leasing company. The business of such a company stems from closing loans with clients, who need money to purchase some products, and with a bank, so the company has money to lend. The management has the right to decide how much and for how long period of time they borrow money. Every closed loan is associated with some costs while it also contributes to the structure of the portfolio of the company. We will develop a stochastic program such that its optimal solution will propose the best borrowing strategy which the management could employ. This will be determined by the value of the portfolio formed by the leasing company's loans at the investment horizon. Moreover, risk constraints will be introduced to the model in order to control the portfolio value in the worst scenarios. Namely, we will discuss the effect of a chance constraint, a Value–at–Risk constraint, a conditional Value–at–Risk constraint and of a second–order stochastic dominance constraint on the portfolio value of the solution at the investment horizon. Last, but not least, we introduce a stress tests which investigates the optimal strategy returns' sensibility to unconsidered (and unfavourable) scenarios. Such a test helps managers to determine how their portfolio of loans would perform in a crisis scenario. With respect to that information, they can adjust the scenario tree or reformulate the risk constraints if the results of the test are not satisfactory.

The structure of the thesis is formed in the following way. First, in Chapter 1, we describe the problem of a private investor in order to introduce the concept of a multi–stage stochastic programming, which is defined afterwards. In the second chapter, we present a business model of a leasing company subsequently rewritten in the language of stochastic programming. Moreover, we specify the risk and the second–order stochastic dominance constraints which will be employed in the asset–liability model and discuss their interpretation and implementation. The third chapter is devoted to a scenario generation procedure of random elements of the problem. We present the Hull – White model (Hull and White, 1990) and describe the calibration procedure we developed for this application. We also create a model which is used for generating scenarios for demand for loans of the leasing company. We utilise all the knowledge we gained in the first three chapters in Chapter 4, where the results of the model analysis are presented. We consider a range of values of the risk constraints' parameters and investigate how they affect the optimal value of the problem and the consequent decisions implied to the management. In the final chapter, a stress test is performed to answer a question of what would be the optimal value of a predetermined strategy in the case of an unexpected development of interest rates.

1. Introduction to a Multi–Stage Stochastic Programming

Stochastic programming has been widely used as a powerful tool in solving optimisation problems with uncertainties. In this chapter, we will focus on the introduction of this method and the illustration of its possible use. First, we will show a simple and a popular example of a multi–stage problem of a private investor which motivates more complicated and more realistic applications of stochastic programming. The definition of a multi–stage stochastic linear programming will follow. Thereafter, under the assumption that the stochastic distribution is discrete with finite number of atoms, we will show how one can derive an equivalent large–scale linear program. This result is often utilised when solving multi–stage stochastic programs, as usually one approximates the continuous distribution of random elements in the problem by a discrete one. We tried to keep this introductory part rather short, so if the reader is interested in a more detailed and a more complex introduction, we refer him to an excellent book of Birge and Louveaux (1997), which served as our major source of information about stochastic programming. Final note in this chapter will relate to the scenario reduction techniques, which can help us to reduce the complexity of a scenario tree. These can be of use when the tree becomes too bushy, which makes the equivalent deterministic problem hardly solvable.

1.1 Motivation Example

At the very beginning of this thesis, we will talk about the idea behind a (multi–stage) stochastic programming and we will mention areas where it was successfully applied on real problems. We will illustrate this concept on a well–known problem of a private investor who seeks to reach a target wealth given initial capital at some time horizon.

In general, stochastic programming models build on deterministic optimisation problems by incorporating uncertainty, which is inevitably part of real world situations. In financial applications, this uncertainty is often caused by random returns of stocks or other assets; in our asset–liability problem, the uncertainty will be represented by unknown value of future interest rates and by a random demand for the company's products. In other fields, it could be the weather, which can considerably influence yields of crops and therefore affect farmers' revenues, or the system's reliability implied by its construction, which might be of interest for engineers. The goal of people facing such problems is usually to find an initial setting which performs the "best on average" in the uncertain environment.

The term "best on average" is of a vague meaning, but in general one can interpret it as having the maximal/minimal value of the expectation of some objective function under the distribution which describes the uncertainty. That is exactly the setting under which stochastic programs can be implemented to find the optimal decision. Often, especially in multi–stage stochastic programs, we are not able to calculate the expected value of our objective function under the distribution which describes the randomness (which is generally unknown). Usually adopted idea in finding a solution of such a problem is to create possible scenarios of the uncertain events and to solve the equivalent large–scale deterministic program. These scenarios can be constructed for example from historical observations of the random event. We will show how one can obtain the deterministic program later in this chapter.

Multi–stage stochastic programs extend the general idea of stochastic programming in such a way that we can make decisions during the time as we observe realizations of random events. For example, as we will see in the following problem of a private investor, he can adjust investment strategies after some time when he observes returns on his investment up to that time. In contrast, a farmer can buy some different fertilizers to compensate for lack of rain during spring.

Let us now describe the well–known problem of a private investor, presented for example in Birge and Louveaux (1997), Section 1.2, to illustrate the concept of a multi–stage stochastic programming. Let us assume that a father thinks about the education of his son. He should start attending a college in $T = 9$ years. The father wants to save money for the tuition fees $G = 250000$ CZK, while he currently has a budget of $B = 200000$ CZK, which he can afford to invest. However, as he is a busy man, he knows he will only have time to adjust his investment strategy after every three years. Moreover, let us for simplicity assume, that he can invest only into a term deposit yielding a constant return

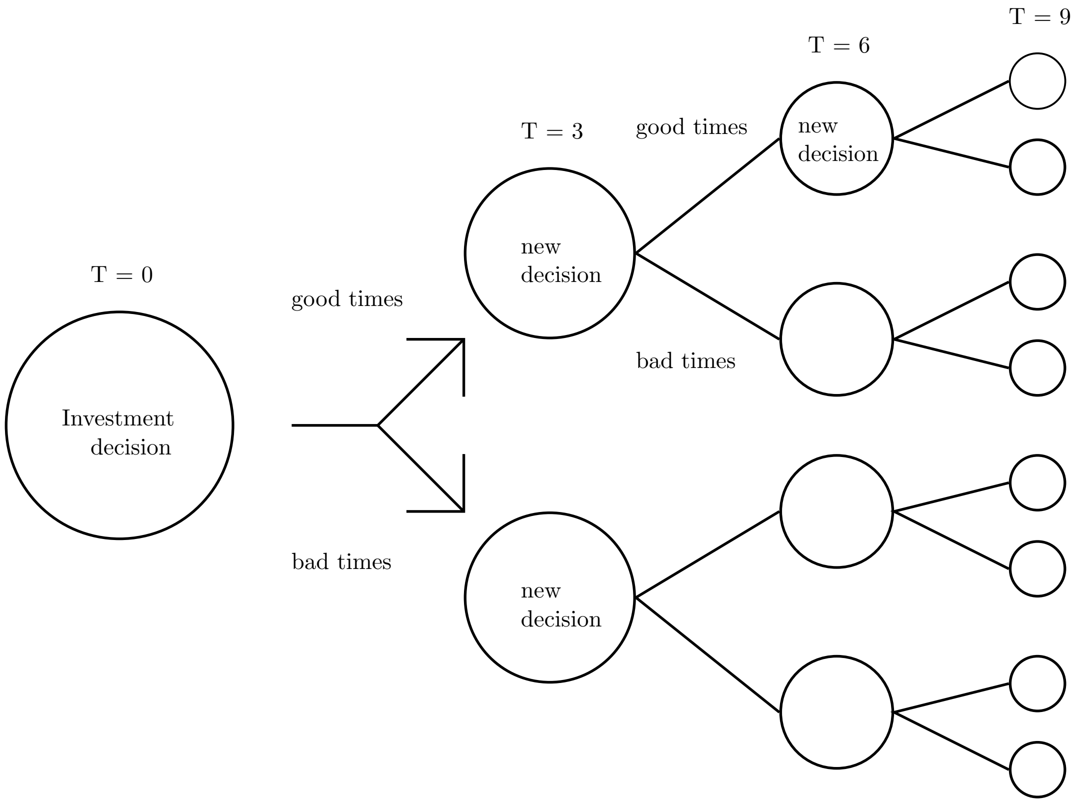

Figure 1.1: A scenario tree of the private investor problem.

of 1.9% a year (= 5.809% over three years) or into a fund whose return varies over time. If times are good, the fund exhibits return of 3.5% per year (= 10.8718% over three years), on the other hand, in bad times its return is only 0.5% per year (= 1.5075% over three years). We also assume that bad and good times happen with equal probability and are independent of the previous states.

The corresponding scenario tree of such a formulated problem is sketched in Figure 1.1. One can see that at every decision time, the investor can choose how much money he invests into the term deposit and how much money into the fund. After three years, he will find himself in a node in the next stage representing the realized state of the economy. There, he finds out the value of his portfolio and he is allowed to readjust his strategy in order to react to the most recent development. We will denote $S_k, k \in \{0,3,6,9\}$ the set of all different nodes s_k at time k and $a(s_k)$ the ancestor of the node $s_k, k \in \{3,6,9\}$. Variable $x_{s_k}(1)$ will represent the amount invested in the term deposit and $x_{s_k}(2)$ analogically the amount invested into the fund in the node $s_k \in S_k,\ \ k \in \{0,3,6\}$. We also denote $\rho_{s_k}(i)$ the return of the term deposit/fund in the node $s_k, k \in \{3,6,9\}, i \in \{1,2\}$.

With such a notation, the problem can be formulated as follows:

$$\max \sum_{s_9 \in S_9} \left(q y^+_{s_9} - r y^-_{s_9} \right),$$

$$\text{s.t.} \ \sum_{i=1}^{2} x_{s_0}(i) = B, \qquad s_0 \in S_0, \tag{1.1}$$

$$\sum_{i=1}^{2} x_{s_k}(i) = \sum_{i=1}^{2} \rho_{s_k}(i) x_{a(s_k)}(i), \qquad s_k \in S_k, k = 3,6, \tag{1.2}$$

$$G + y^+_{s_9} - y^-_{s_9} = \sum_{i=1}^{2} \rho_{s_9}(i) x_{a(s_9)}(i), \qquad s_9 \in S_9, \tag{1.3}$$

$$y^+_{s_9}, y^-_{s_9}, x_{s_k}(i) \geq 0, \qquad s_k \in S_k, \ \ k = 0,3,6, \ \ i = 1,2.$$

The variable $y^+_{s_9}$ represents the surplus over the target G which the investor has in the node $s_9 \in S_9$. Analogically, $y^-_{s_9}$ stands for the shortage. Parameter q summarizes the additional utility the investor has from a unit surplus, while r expresses the utility lost by the cost incurred by borrowing money in order to pay the tuition fees. In our analysis, we set $q = 1$ and $r = 4$, so we penalized four times more missing the target than exceeding it. Equation (1.1) describes the initial budget constraint of the investor, while (1.2) corresponds to the budget constraints he faces in the course of the time. It requires the investor to invests full amount obtained from the investment made in the previous period. The final constraints (1.3) evaluate the surplus/shortage on G in the final stage of the private investor problem.

The optimal solution of the problem in the first period was to invest a majority of funds into the fund hoping for good times. If that was the case, it proposed in the next period to adopt a more conservative strategy — to invest more into the term deposit. Under such a strategy, the investor could reach target G if at least one good time happened. In the case of both bad times states in the final two stages, the investor still managed to finish just below the target, mostly thanks to the conservative strategy. On the other hand, if the first state was a bad time, the investor's only chance to reach the target was to invest all in the fund both periods and hope for good times. Thanks to the small amount invested in the first stage into the term deposit, the optimal strategy has reached the target with two good times coming after the first bad time too.

Summing this all together, the optimal objective value of the problem was -45137. In comparison, if the investor replaced the random returns by their expectations, he would end up by investing only into the fund, which would imply the value of the objective function to be -50531. What is even more striking is that under such a strategy, he would reach the goal only in one scenario — with a probability 0.125. On the other hand, by following the optimal strategy suggested by the multi–stage stochastic program, he would have reached the target in four scenarios — with a probability 0.5. The ability of stochastic programming to meet these so called probabilistic constraints is possibly even more important. One can imagine that the value of fulfilling a predetermined goal can be of a great importance for a company, both in a way that it avoids fines or earns bonuses. Stochastic programming can suggest the optimal investing strategy with respect to reaching such a goal as well.

1.2 Multi–Stage Stochastic Linear Program

Let us now move to a general definition of a multi–stage stochastic linear program, which will be inspired by Birge and Louveaux (1997), Section 3.5. First, we will introduce the stochastic process which generates randomness in the program. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\xi_t, t \in T$ be a collection of real–valued random vectors on that probability space with a marginal probability distribution denoted as Ξ_t. The set T is some ordered index set usually with a meaning of time. Assume that the decisions of the multi–stage stochastic program take place at times $t_1, \ldots, t_n \in T$. At time $t_k, k \in 1, \ldots, n$, we observe a realization of the stochastic process $\xi_{t_k}(\omega_k), \omega_k \in \Omega$. Let us denote $\mathcal{F}_k$ the σ-field generated by the process up until the time t_k. The decision, which will take place at time t_k will be denoted by x_k. These decisions must be non–anticipative, hence they must not depend on future realizations of ω. In other words, x_k are $\mathcal{F}_k$ measurable. Without loss of generality, we can assume that $\mathcal{F}_1 = \{\emptyset, \Omega\}$, which in other words means that ω_1 is known and ξ_1 and x_1 are deterministic. For more simple notation, let us denote $\bar{\omega}_k = (\omega_1, \ldots, \omega_k)$ the observed history of the process up to time t_k and similarly the history of decisions $\bar{x}_k = (x_1, \ldots, x_k)$.

For easier understanding the multi–stage stochastic linear program we will formulate the definition in the so called nested form. This reads as follows:

$$\begin{aligned} \min \quad & c_1^T x_1 + \mathbb{E}_{\Xi_{t_2}}\Big[\varphi_1(x_1, \omega_2)\Big], \\ \text{s.t.} \quad & A_1 x_1 = b_1, \\ & x_1 \geq 0, \end{aligned} \tag{1.4}$$

where $\varphi_k(\cdot,\cdot), k = 1,\ldots,n-1$ are given recursively as

$$\begin{aligned} \varphi_k(\bar{x}_k,\bar{\omega}_{k+1}) = \min \quad & c_{k+1}^T(\bar{\omega}_{k+1})x_{k+1} + \mathbb{E}_{\Xi_{t_{k+2}}|\bar{\omega}_{k+1}}\Big[\varphi_{k+1}(\bar{x}_{k+1},\bar{\omega}_{k+2})\Big], \\ \text{s.t.} \quad & B_{k+1}(\bar{\omega}_{k+1})\bar{x}_k + A_{k+1}(\bar{\omega}_{k+1})x_{k+1} = b_{k+1}(\bar{\omega}_{k+1}), \\ & x_{k+1} \geq 0, \end{aligned} \tag{1.5}$$

where $\varphi_n(\bar{x}_n) = 0$. All relationships in (1.5) are meant to hold almost surely with respect to the corresponding probability distribution Ξ_{t_k}. We also assume that all expectations exist and all minima are attained. Vector $b_1 \in \mathbb{R}^{m_1}$ together with matrix $A_1 \in \mathbb{R}^{m_1\times n_1}$ form known initial constraints on x_1. Vector $c_1 \in \mathbb{R}^{n_1}$ is also known and defines the effect of the initial decision x_1 on the value of the objective function. Symbol $\Xi_{t_{k+2}}|\bar{\omega}_{k+1}$ denotes the conditional probability distribution of the random vector $\xi_{t_{k+2}}$ given the observed value of $\bar{\omega}_{k+1}$. We can realize that $\xi_{t_{k+1}}(\bar{\omega}_{k+1})^T = (c_{k+1}(\bar{\omega}_{k+1})^T, b_{k+1}(\bar{\omega}_{k+1})^T, \text{vec}(A_{k+1}(\bar{\omega}_{k+1}))^T, \text{vec}(B_{k+1}(\bar{\omega}_{k+1}))^T)$ is an N_k-dimensional random vector, living on a probability space $(\Omega, \mathcal{F}_{k+1}, \mathbb{P})$. This vector expresses all the stochastic components of the problem which arise in the stage $k+1$. Symbol $\text{vec}(\cdot)$ is the vectorization operator.

The core in understanding the multi–stage formulation is interpreting the function $\varphi_k(\bar{x}_k,\bar{\omega}_{k+1})$. This function specifies the optimisation problem we will face in the next period. Hence at time t_k we minimize together the contribution of the current decision to the objective function plus the (expected) minimal value we can achieve in the next period with such a decision. We should mention why we call the above–described multi–stage program linear. It stems from the fact that all the constraints are only linear functions of the decisions variables, while the same holds for the objective function. We will utilise this fact when we will assume that Ξ_{t_k} is a discrete probability distribution with a finite number of atoms for all k. Then such a stochastic program can be rewritten into a large–scale linear program.

Let us suppose that ξ_{t_k} has such a distribution. Then there is a collection of scenarios $\{\omega^1,\ldots,\omega^S\}$ such that $\mathbb{P}\Big(\omega \in \{\omega^1,\ldots,\omega^S\}\Big) = 1$. For $s = 1,\ldots,S$, the probability of such a scenario $\omega^s = (\omega_2^s,\ldots,\omega_n^s)$ is $p^s > 0$. Each scenario generates coefficients $\xi_{t_2}^s,\ldots,\xi_{t_n}^s$. Moreover, scenarios can be organized into the form of a scenario tree, where for every decision time, we consider together scenarios which have the same history up to that time. Formally, for time $t_k, k = 2,\ldots,n$, we define S_k as a set of indices so that it holds $\{\bar{\omega}_{s_k}|s_k \in S_k\} = \{\bar{\omega}_k^s|s = 1,\ldots,S\}$, while we also have that $\forall s_k^1, s_k^2 \in S_k, s_k^1 \neq s_k^2 : \bar{\omega}_{s_k^1} \neq \bar{\omega}_{s_k^2}$. The set S_k represents set of all different nodes of the scenario tree at time t_k. From the tree structure, it is clear that every node $\bar{\omega}_{s_{k+1}}$ has got a unique ancestor $\bar{\omega}_{s_k}$. This node will be denoted as $s_k = a(s_{k+1})$. On the other hand, a node $\bar{\omega}_{s_k}$ can have a multiple number of descendants. Finally, let us denote $p_{s_k} > 0$ the probability of the node $\bar{\omega}_{s_k}$ at time t_k. It must hold $\sum_{s_k\in S_k} p_{s_k} = 1$. Let us also denote $A_{s_k}, b_{s_k}, B_{s_k}$ and c_{s_k} the corresponding realizations of the stochastic components of the problem. Now we have all the required ingredients to reformulate the multi–stage stochastic linear program as in equations (1.4) and (1.5) as a linear program.

First, we need to notice that the expected values in the objective functions will become only weighted sums of linear functions of the decision variables. The afore–mentioned non–anticipativity constraints require that for all scenarios with the same history until time t_k we must make the same decisions up to that time. This corresponds to having the same decision for all scenarios sharing a node s_k in S_k. Hence by using the tree structure representation of scenarios, we include the non–anticipativity constraints implicitly (they do not require any constraints in the model formulation). Moreover, we can see that the constraints in (1.5) are at time t_k the same for all scenarios having a common node of the scenario tree. Because we have also only one decision vector for this node, we can rewrite constraints (1.5) together for all scenarios sharing the same node at once. Using all these findings, we can express the multi–stage stochastic linear program as in (1.4) and (1.5) as

$$\min \quad c_1^T x_1 + \sum_{s_2 \in S_2} p_{s_2} c_{s_2}^T x_{s_2} + \cdots + \sum_{s_n \in S_n} p_{s_n} c_{s_n}^T x_{s_n}, \tag{1.6}$$

$$\begin{array}{llllllll}
\text{s.t.} & A_1 x_1 & & & & & = b_1, & \\
& B_{s_2} x_1 & + & A_{s_2} x_{s_2} & & & = b_{s_2}, & s_2 \in S_2, \\
& & & B_{s_3} x_{a(s_3)} & + & A_{s_3} x_{s_3} & = b_{s_3}, & s_3 \in S_3, \\
& & & & \ddots & \ddots & \vdots & \\
& & & & B_{s_n} x_{a(s_n)} & + \quad A_{s_n} x_{s_n} & = b_{s_n}, & s_n \in S_n,
\end{array}$$

$$x_1 \geq 0, \;\; x_{s_k} \geq 0, \;\; s_k \in S_k, \;\; k = 2, \ldots, n.$$

The above derived result is often used for continuous distributions of the stochastic process. The usual methodology is to generate scenarios from the continuous distribution and then to use this discrete distribution as an approximate to the continuous distribution. There are various ways how to generate scenarios. These could be created based on historical observations, a suitable calibrated mathematical model, or from an expert opinion. It is often done with the use of a combination of the three approaches. Clearly, the quality of scenarios influence the quality of the solution, which makes the scenario generation procedure massively important. The ultimate goal is to obtain the best possible representation of the continuous distribution. One could possibly enhance the accuracy by increasing the number of scenarios. This however also increases the computational difficulty of the program, so one needs to find a balance between the number of scenarios and the required computational time. Various techniques have been developed in order to reduce a large scenario tree while keeping a high relative accuracy of the solution. To be more specific, a scenario reduction process based on finding a smaller tree with a given cardinality as close as possible to the original tree was proposed by Dupačová et al. (2003). These methods could be also based on a forward selection method, see e.g. Heitsch and Römisch (2003).

In our multi–stage model we will employ some risk constraints which lead to a large–scale mixed–integer linear program. The generalization of the formulation to a mixed–integer program is straightforward as one only adds restrictions on the set of possible values of parameters into the model. We will not present a precise formulation of it. Instead, we will point out in the asset–liability model formulation where binary variables appear.

2. Asset–Liability Model of a Leasing Company

This chapter will be devoted to a description of a business model of a leasing company. We will present the problem which analysts tackle and we will formulate it in the language of stochastic programming. Thereafter, we will introduce four concepts of how to control the interest rate risk which the company faces and we will show how one can incorporate them into the model formulation. Moreover, we will discuss the interpretation of the optimal solutions when such constraints are employed.

2.1 Business Model of a Leasing Company

A leasing company is an enterprise, whose scope of business is to lend money to clients to purchase some products. On the other hand, the leasing company itself borrows money from a bank in order to have money to lend. This will represent the business model of the company, whose aim will be to make money on better interest it gets from the bank. For simplicity, we will not consider defaults of clients on their loans. That is because of the collateral, which is associated with leasing loans. In the case of a default of the client the collateral is sold to repay the loan. Hence the company does not loose any (substantial amount of) money. Generalizing our model to consider such losses would not be difficult as we would just subtract the part of a mark–up on the interest rate corresponding to the risk profile of a client as an insurance.

We will treat our problem as a discrete time. We will assume that all cash flows will take place at times $t_0 < t_1 < \cdots < t_n < \cdots$, where $n \in \mathbb{N}$. Time t_0 can be thought as present and time t_n as the investment horizon. That will be the time at which we will be aiming to maximize our profit. All loans of the leasing company which were closed before time t_n and did not mature by the time t_n will be thought to be sold for their market value at the investment horizon. Moreover, no new contracts will be closed at time t_n or later. The most obvious choice for times t_i would be to set them as an equidistant sequence with monthly or yearly period. Note that these times will be in fact the decision times, where the board will have the opportunity to set the amount of money they borrow from the bank. This should be also taken into account when considering what time structure within the model should be employed. Still, we want to stress that we do not require the sequence to be equidistant.

2.1.1 Model Variables

First, we will introduce the concept of interest rates which expresses the cost of borrowing money. We will denote y_t^τ the annualized, continuously compounded time t yield–to–maturity with time–to–maturity τ. The curve

$$y_t^\tau, \qquad \tau > 0$$

represents the yield curve at time t. We should note that for $t = t_0$, the curve $y_{t_0}^{\tau}, \tau > 0$ is the currently observed market yield curve and it is therefore given. On the other hand, for $t = t_i, i \geq 1$, the curve $y_{t_i}^{\tau}, \tau > 0$ is (given information at time t_0) random. For that reason, it will be in latter chapters denoted by $y_{t_i}^{\tau}(\omega), \tau > 0$ to stress its randomness. The notation of random variables in our problem will be discussed in more detail in Section 2.2, where we will present the formulation of this problem as a multi–stage stochastic program.

The leasing company will borrow from the bank at a rate

$$s_t^{\tau} = y_t^{\tau} + s(\tau).$$

Quantity $s(\tau)$ can be interpreted as a spread between the market yield–to–maturity and the offered rate for a loan from the bank to the leasing company. We assume $s(\tau)$ to be given and one can expect it to be increasing in τ for example because of an increasing probability of default of the leasing company in time. Analogically, we will define a rate at which clients can borrow from the leasing company. This will be denoted by

$$r_t^{\tau} = y_t^{\tau} + s(\tau) + m(\tau) = s_t^{\tau} + m(\tau).$$

We will refer to $m(\tau)$ as to a mark–up, which the company charges. We assume it is the same for all clients and fixed, so it does not change in time.

Let us move to the cash–flows between the leasing company and clients. We will assume that clients at time t_i demand d_{t_i,t_j} of loans till time $t_j, j > i, i < n$ and that all the demanded loans will be granted by the leasing company. Demand is random (will depend on the random element ω), non–negative and it might depend on the actual level of interest rates. The loan d_{t_i,t_j} will be repaid at all times t_k such that $t_i < t_k \leq t_j$ with the same intensity λ. In other words, clients will pay at time $t_k, i < k \leq j$ the amount $\lambda \cdot (t_k - t_{k-1})$. The value of λ can be derived from the relationship

$$d_{t_i,t_j} = \sum_{k=i+1}^{j} \lambda(t_k - t_{k-1}) \exp\{-r_{t_i}^{t_k - t_i}(t_k - t_i)\}.$$

The amount of money repaid from a loan d_{t_i,t_j} at time t_k will be denoted as $d_{t_i,t_j}(t_k)$. It can be expressed as

$$d_{t_i,t_j}(t_k) = \frac{(t_k - t_{k-1}) d_{t_i,t_j}}{\sum_{l=i+1}^{j}(t_l - t_{l-1}) \exp\{-r_{t_i}^{t_l - t_i}(t_l - t_i)\}}, \qquad k = i+1, \ldots, j. \quad (2.1)$$

In the case that the times are equidistant. The formula (2.1) simplifies to

$$d_{t_i,t_j}(t_k) = \frac{d_{t_i,t_j}}{\sum_{l=i+1}^{j} \exp\{-r_{t_i}^{t_l - t_i}(t_l - t_i)\}}, \qquad k = i+1, \ldots, j. \quad (2.2)$$

One can notice in (2.2), that the right hand side does not depend on k, which corresponds to the fact that the amount of instalments is the same at all times and equals λ times the length of a period between the decision stages. From this formula, one can derive the amount of money repaid back by clients at time t_k. This quantity will be denoted by R_{t_k} and is given by

$$R_{t_k} = \sum_{t_i < t_k \le t_j} d_{t_i,t_j}(t_k). \tag{2.3}$$

Note that in the sum above, index k stays fixed and we sum over the indices i and j. That means we sum over the loans closed before time t_k which have not yet matured. Similarly, we will denote D_{t_k} the amount of money which the leasing company has lent to clients at time t_k.

$$D_{t_k} = \sum_{t_k < t_j} d_{t_k,t_j}. \tag{2.4}$$

Next, we will define the corresponding quantities for cash–flows with the bank. The main difference will be that the amount of money the leasing company borrows from the bank is not random. It is in fact a decision variable in our problem which will be subject to optimisation. Let us denote the amount of money the leasing company wants to borrow from time t_i to time t_j as $x_{t_i,t_j}, i < n, i < j$. The amount of loans repaid back to the bank from a loan x_{t_i,t_j} at time t_k will be denoted by $x_{t_i,t_j}(t_k)$ and it can be determined as

$$x_{t_i,t_j}(t_k) = \frac{(t_k - t_{k-1})x_{t_i,t_j}}{\sum_{l=i+1}^{j}(t_l - t_{l-1})\exp\{-s_{t_i}^{t_l-t_i}(t_l - t_i)\}}, \qquad k = i+1, \dots, j. \tag{2.5}$$

Summing up (2.5) over all loans running at time t_k, one obtains the total amount of money repaid back to the bank — denoted by Q_{t_k}.

$$Q_{t_k} = \sum_{t_i < t_k \le t_j} x_{t_i,t_j}(t_k). \tag{2.6}$$

Finally, the total amount borrowed at time t_k is

$$X_{t_k} = \sum_{t_k < t_j} x_{t_k,t_j}. \tag{2.7}$$

Now, denote B_{t_k} the amount of money the leasing company has on its account immediately after time t_k and E_{t_k} the company's running costs from time t_k to t_{k+1}. Then the following condition has to be met

$$B_{t_k} = B_{t_{k-1}} \exp\{y_{t_{k-1}}^{t_k - t_{k-1}}(t_k - t_{k-1})\} - E_{t_{k-1}} + R_{t_k} - D_{t_k} + X_{t_k} - Q_{t_k}. \tag{2.8}$$

The relationship in (2.8) describes the movements on current account of the company. It needs to pay the instalments to the bank and lend money to clients, while it receives some money back from clients and loans from the bank. Moreover, it has to pay its costs (wages, rents). Also the risk free interest from the money left on the bank account in the last period is added.

What will be needed in the model are variables describing the actual state of business. In other words, we will need to find a relationship for the value of assets and for the value of liabilities at every time. Let us denote $P(t_k, t_l)$ time t_k price of a zero–coupon bond paying one at time $t_l, l > k$. This can be interpreted as the discount factor from time t_l to time t_k. The value of assets of the leasing company at time t_k will be denoted by A_{t_k} and it meets the following relation

$$A_{t_k} = \sum_{t_i \le t_k < t_l \le t_j} P(t_k, t_l) d_{t_i,t_j}(t_l). \tag{2.9}$$

Similarly, the value of the company's liabilities at time t_k will be denoted by L_{t_k}. It holds that

$$L_{t_k} = \sum_{t_i \le t_k < t_l \le t_j} P(t_k, t_l) x_{t_i,t_j}(t_l). \tag{2.10}$$

Finally, we will define V_{t_k} as the value of the portfolio at time t_k. This will consist of the value of assets, liabilities and of the amount of money on the company's current account. Formally,

$$V_{t_k} = A_{t_k} - L_{t_k} + B_{t_k}. \tag{2.11}$$

This yields the most basic formulation of the optimisation problem as

$$\begin{aligned} \max_{x \in \mathcal{X}} \quad & \mathbb{E} V_{t_n}, \\ \text{s.t.} \quad & B_{t_k} \ge 0, \ \text{a.s.} \qquad \forall k = 1, \ldots, n, \\ & x_{t_i,t_j} \ge 0, \qquad \forall i < j, \quad 0 \le i < n. \end{aligned}$$

One can notice that given the definition of V_{t_n} as in (2.11) and the relationships (2.5), (2.6), (2.7), (2.8) and (2.10), the above stated optimisation problem is a concisely described multi–stage stochastic linear program. We will state a precise and a more detailed definition of such an optimisation problem later in this chapter.

2.1.2 Benchmark Strategy

The optimal solution of the stochastic program proposes a way how to structure a portfolio of loans so it leads to the maximum profit, while it meets some predetermined conditions. Thereafter, we would like the manager of the company to act according to it. However, it is obvious that before changing the strategy, the manager would like to see how does the new optimal strategy compare to the one he is currently practising — to a benchmark strategy. Such a comparison is important at least from the two following reasons. First, we can decide if the additional income we get is worth changing the routines ingrained in the company, and second, we can identify places where we would like the optimal strategy to do better and so we can adjust the model formulation. Currently, the business is done in the following way. Consider a client who comes to the leasing company demanding a loan of d_{t_i,t_j}. The leasing company closes a deal with the client and it immediately closes a mirror deal with the bank too. By the term mirror deal, we mean a loan with the same notional and the same maturity — the company borrows d_{t_i,t_j}. By such a behaviour, the leasing company closes its position and gains profit, which is generated by the difference between the two rates. Let us define the benchmark strategy as

$$x^0_{t_i,t_j} = d_{t_i,t_j}. \tag{2.12}$$

This implies that

$$\begin{aligned}
x^0_{t_i,t_j}(t_k) &= \frac{(t_k - t_{k-1})d_{t_i,t_j}}{\sum_{l=i+1}^{j}(t_l - t_{l-1})\exp\{-s_{t_i}^{t_l - t_i}(t_l - t_i)\}},\\
Q^0_{t_k} &= \sum_{t_i < t_k \le t_j} x^0_{t_i,t_j}(t_k),\\
X^0_{t_k} &= \sum_{t_k < t_j} d_{t_k,t_j} = D_{t_k},\\
B^0_{t_k} &= B^0_{t_{k-1}} \exp\{y_{t_{k-1}}^{t_k - t_{k-1}}(t_k - t_{k-1})\} - E_{t_{k-1}} + R_{t_k} - Q^0_{t_k}.
\end{aligned}$$

One can also express the difference $R_{t_k} - Q^0_{t_k}$ to see the earnings at time t_k. The liabilities of the benchmark strategy are

$$L^0_{t_k} = \sum_{t_i \le t_k < t_l \le t_j} P(t_k, t_l) x^0_{t_i,t_j}(t_l).$$

Finally, the value of the benchmark portfolio can be expressed as

$$V^0_{t_k} = A_{t_k} - L^0_{t_k} + B^0_{t_k}.$$

The formulae are analogical to what we have seen in the previous section, because we only plugged in for x_{t_i,t_j} the values of d_{t_i,t_j} as given in (2.12). We can see, that if we employ such a strategy, there is no way to manipulate with the leasing company's income. Even though it cannot loose anything on interest rate movements, it is still not assured to earn profit as it might not be able to gain enough to cover for its costs. Therefore, there is no guarantee that the benchmark strategy meets the (survival) condition

$$B^0_{t_k} \ge 0, \qquad \forall k = 1, \dots, n.$$

So let us assume that our leasing company is healthy enough that by employing such a strategy it does not end up without any money on its account.

The main idea why it could be beneficial for the company to behave differently and not to follow the benchmark strategy is that the short term rates are usually lower than the long term rates. Hence the company instead of closing a mirror deal with maturity ten years would borrow money only for shorter periods multiple times. This could generate additional profit. However, that would imply the company to open its interest rate risk position which could cause significant losses (or gains) in the case of an unexpected interest rate movement. For this reason, an introduction of certain measures of risk is required. Imposing risk limits in the optimisation problem would restrict the set of feasible strategies. Then we will be able to control the probability of extreme and unfavourable outcomes.

2.2 The Model as a Multi–Stage Stochastic Program

In the following section, we will reformulate our asset–liability problem of the leasing company as a multi–stage stochastic program. The reformulation will be provided in terms of both multi–stage stochastic programming and linear programming, where the distribution of random variables will be approximated by scenarios.

First, let us recall the structure of our problem. We have times $t_0 < t_1 < \cdots < t_n < \cdots$ at which cash–flows of the leasing company are exchanged and where t_n is the investment horizon at which we want to maximize the value of the portfolio. However, we are now at time t_0 and the value itself, along with many other factors which influence it, is random. This randomness will be described by a random vector $\omega = (\omega_0, \omega_1, \dots, \omega_n)$, its history up to time $k \le n$ will be denoted $\bar{\omega}_k = (\omega_0, \dots, \omega_k)$. We consider ω_0 to be given — non–random, as it represents the information at current time. Variables and parameters of the model were described in Section 2.1, but the notation in this section will be extended. Random components of the problem which depend on realizations of the random vector up to time t_k will be denoted by the established notation and by a term $\bar{\omega}_k$. This will specify the information when we observe the value of the random element. The multi–stage stochastic program without any risk constraints then reads as follows:

$$
\begin{aligned}
\max_{x_{t_i,t_j}} \quad & \mathbb{E} V_{t_n}(\bar{\omega}_n), && (2.13)\\
\text{s.t.} \quad d_{t_i,t_j}(t_k,\bar{\omega}_i) &= \frac{(t_k - t_{k-1}) d_{t_i,t_j}(\bar{\omega}_i)}{\sum_{l=i+1}^{j}(t_l - t_{l-1}) \exp\{-r_{t_i}^{t_l - t_i}(\bar{\omega}_i)(t_l - t_i)\}}, \qquad i < k \le j,\\
x_{t_i,t_j}(t_k,\bar{\omega}_i) &= \frac{(t_k - t_{k-1}) x_{t_i,t_j}(\bar{\omega}_i)}{\sum_{l=i+1}^{j}(t_l - t_{l-1}) \exp\{-s_{t_i}^{t_l - t_i}(\bar{\omega}_i)(t_l - t_i)\}}, \qquad i < k \le j,\\
R_{t_k}(\bar{\omega}_{k-1}) &= \sum_{t_i < t_k \le t_j} d_{t_i,t_j}(t_k,\bar{\omega}_i), \qquad 0 < k \le n,\\
Q_{t_k}(\bar{\omega}_{k-1}) &= \sum_{t_i < t_k \le t_j} x_{t_i,t_j}(t_k,\bar{\omega}_i), \qquad 0 < k \le n,\\
D_{t_k}(\bar{\omega}_k) &= \sum_{t_k < t_j} d_{t_k,t_j}(\bar{\omega}_k), \quad X_{t_k}(\bar{\omega}_k) = \sum_{t_k < t_j} x_{t_k,t_j}(\bar{\omega}_k), \quad 0 \le k < n,\\
B_{t_0}(\omega_0) &= X_{t_0}(\omega_0) - D_{t_0}(\omega_0),\\
B_{t_k}(\bar{\omega}_k) &= \frac{B_{t_{k-1}}(\bar{\omega}_{k-1})}{P(t_{k-1}, t_k, \bar{\omega}_{k-1})} - E_{t_{k-1}} + X_{t_k}(\bar{\omega}_k) - Q_{t_k}(\bar{\omega}_{k-1})\\
&\quad + R_{t_k}(\bar{\omega}_{k-1}) - D_{t_k}(\bar{\omega}_k), \qquad 1 \le k < n,\\
B_{t_n}(\bar{\omega}_{n-1}) &= \frac{B_{t_{n-1}}(\bar{\omega}_{n-1})}{P(t_{n-1}, t_n, \bar{\omega}_{n-1})} - E_{t_{n-1}} + R_{t_n}(\bar{\omega}_{n-1}) - Q_{t_n}(\bar{\omega}_{n-1}),\\
A_{t_n}(\bar{\omega}_n) &= \sum_{t_i < t_n < t_l \le t_j} P(t_n, t_l, \bar{\omega}_n) d_{t_i,t_j}(t_l, \bar{\omega}_i),\\
L_{t_n}(\bar{\omega}_n) &= \sum_{t_i < t_n < t_l \le t_j} P(t_n, t_l, \bar{\omega}_n) x_{t_i,t_j}(t_l, \bar{\omega}_i),\\
V_{t_n}(\bar{\omega}_n) &= A_{t_n}(\bar{\omega}_n) - L_{t_n}(\bar{\omega}_n) + B_{t_n}(\bar{\omega}_{n-1}),\\
B_{t_k}(\bar{\omega}_k) &\ge 0, \quad 0 \le k \le n, \qquad x_{t_i,t_j}(\bar{\omega}_i) \ge 0, \quad 0 \le i < j, \; i < n,
\end{aligned}
$$

where all equalities and inequalities are meant to hold almost surely with respect to the probabilistic distribution of ω. We can see, that there is basically only one restriction in the model as we require the leasing company to have non–negative amount on its current account at every time. Other relationships describe the dynamics of cash–flows of the company. The exact solution of such a problem is unreachable, since even if we knew the distribution of the random vector ω, determining the expected value of the objective function as a deterministic function of our decisions seems impossible. For that reason, we will restrict ourselves to approximating the stochastic distribution by discrete scenarios, which will be organized in the form of a scenario tree. The process of generating scenarios will be subject of discussion in the next chapter. Here, we will specify the structure of the tree. At every time $t_k, k = 0, 1, \ldots, n$ we will consider nodes $\bar{\omega}_k^s, s \in S_k$ with equal probability $1/|S_k|$. We require $|S_0| = 1$, as the only element of S_0 is the current state and that is non–random. Every node at time t_k will have for all times $t_0 \leq t_i < t_k$ a unique ancestor denoted by $a_i(\bar{\omega}_k^s)$. The non–anticipativity constraint requires a single decision at every node, which we will shortly denote $x_{t_k,t_j}^s = x_{t_k,t_j}(\bar{\omega}_k^s), s \in S_k$. Employing such a tree structure of our scenarios, we can use the relationship between multi–stage stochastic programs with discrete probability distribution with finite number of atoms and linear programs which was presented in Section 1.2. This implies that the multi–stage asset–liability model of the leasing company can be written in terms of linear programming as

$$\max_{x_{t_i,t_j}^s} \quad \frac{1}{|S_n|} \sum_{s \in S_n} V_{t_n}(\bar{\omega}_n^s) \tag{2.14}$$

$$\begin{aligned}
\text{s.t. } d_{t_i,t_j}(t_k, \bar{\omega}_i^s) &= \frac{(t_k - t_{k-1}) d_{t_i,t_j}(\bar{\omega}_i^s)}{\sum_{l=i+1}^{j} (t_l - t_{l-1}) \exp\{-r_{t_i}^{t_l - t_i}(\bar{\omega}_i^s)(t_l - t_i)\}}, \quad i < k \leq j, s \in S_i, \\
x_{t_i,t_j}(t_k, \bar{\omega}_i^s) &= \frac{(t_k - t_{k-1}) x_{t_i,t_j}^s}{\sum_{l=i+1}^{j} (t_l - t_{l-1}) \exp\{-s_{t_i}^{t_l - t_i}(\bar{\omega}_i^s)(t_l - t_i)\}}, \quad i < k \leq j, s \in S_i, \\
R_{t_k}(\bar{\omega}_{k-1}^s) &= \sum_{t_i < t_k \leq t_j} d_{t_i,t_j}(t_k, a_i(\bar{\omega}_{k-1}^s)), \qquad 0 < k \leq n, \ \ s \in S_{k-1}, \\
Q_{t_k}(\bar{\omega}_{k-1}^s) &= \sum_{t_i < t_k \leq t_j} x_{t_i,t_j}(t_k, a_i(\bar{\omega}_{k-1}^s)), \qquad 0 < k \leq n, \ \ s \in S_{k-1}, \\
D_{t_k}(\bar{\omega}_k^s) &= \sum_{t_k < t_j} d_{t_k,t_j}(\bar{\omega}_k^s), \ \ X_{t_k}(\bar{\omega}_k^s) = \sum_{t_k < t_j} x_{t_k,t_j}^s, \ \ 0 \leq k < n, \ \ s \in S_k, \\
B_{t_0}(\omega_0^s) &= X_{t_0}(\omega_0^s) - D_{t_0}(\omega_0^s), \quad s \in S_0, \\
B_{t_k}(\bar{\omega}_k^s) &= \frac{B_{t_{k-1}}(a_{k-1}(\bar{\omega}_k^s))}{P(t_{k-1}, t_k, a_{k-1}(\bar{\omega}_k^s))} - E_{t_{k-1}} + X_{t_k}(\bar{\omega}_k^s) - Q_{t_k}(a_{k-1}(\bar{\omega}_k^s)) \\
&\quad + R_{t_k}(a_{k-1}(\bar{\omega}_k^s)) - D_{t_k}(\bar{\omega}_k^s), \quad 1 \leq k < n, s \in S_k, \\
B_{t_n}(\bar{\omega}_{n-1}^s) &= \frac{B_{t_{n-1}}(\bar{\omega}_{n-1}^s)}{P(t_{n-1}, t_n, \bar{\omega}_{n-1}^s)} - E_{t_{n-1}} + R_{t_n}(\bar{\omega}_{n-1}^s) - Q_{t_n}(\bar{\omega}_{n-1}^s), s \in S_{n-1}, \\
A_{t_n}(\bar{\omega}_n^s) &= \sum_{t_i < t_n < t_l \leq t_j} P(t_n, t_l, \bar{\omega}_n^s) d_{t_i,t_j}(t_l, a_i(\bar{\omega}_n^s)), \quad s \in S_n, \\
L_{t_n}(\bar{\omega}_n^s) &= \sum_{t_i < t_n < t_l \leq t_j} P(t_n, t_l, \bar{\omega}_n^s) x_{t_i,t_j}(t_l, a_i(\bar{\omega}_n^s)), \quad s \in S_n, \\
V_{t_n}(\bar{\omega}_n^s) &= A_{t_n}(\bar{\omega}_n^s) - L_{t_n}(\bar{\omega}_n^s) + B_{t_n}(a_{n-1}(\bar{\omega}_n^s)), \quad s \in S_n, \\
B_{t_k}(\bar{\omega}_k^s) &\geq 0, \quad 0 \leq k < n, s \in S_k, \quad x_{t_i,t_j}^s \geq 0, \quad s \in S_i, \ \ 0 \leq i \leq j, i < n.
\end{aligned}$$

We have seen the formulation of the optimisation problem which seeks to maximize expected revenue. However, one would hardly found a manager who would be willing to follow the optimal strategy suggested by such a program. That is due to a human nature and the phenomena called risk aversion. We prefer to prevent losses to gaining additional revenue. This motivates the introduction of certain measures of risk which will force the optimal strategy to avoid too risky investment. These will be described along with their interpretation and implementation to the model formulation in the next section.

2.3 Management of Interest Rate Risk

We stated that by employing the benchmark strategy, the leasing company closes its position in terms of interest rate risk. That does not mean that the leasing company is insensitive to interest rate changes, as the value of the portfolio depends on interest rate. The correct interpretation is that the income generated by its assets covers all instalments implied by its liabilities. Hence the company does not need to close any new loans with the bank in order to cover its current liabilities. If we do not employ the benchmark strategy, it could happen that we committed to repay a big loan to the bank next period while our earnings are not going to be sufficient. Hence we rely on the fact that we will borrow money in the next period and in the case of an interest rate increase, the loan might turn out to be very expensive. In contrast, such a situation cannot happen within the benchmark strategy, where by definition our earnings will always be greater than our instalments at every time period. For that reason, one needs to introduce risk constraints, which will set limits on riskiness of a strategy. The ones we will introduce can be divided into two groups, the first one will compare the optimal strategy to the benchmark while constraints in the second group will control the return for the worst scenarios.

2.3.1 Chance Constraint

The first constraint we will introduce will be the so called chance or probabilistic constraint. This constraint was first formulated by Telser (1955) and used in asset–liability models for example by Dert (1995) or Klein Haneveld et al. (2010). Roughly speaking, this constraint ensures that our strategy will be better than the benchmark strategy with some probability. It has the following form:

$$\mathbb{P}\Big(V_{t_n} - V^0_{t_n} \geq 0\Big) \geq 1 - \alpha, \tag{2.15}$$

where $\alpha \in \langle 0, 1\rangle$ denotes a given probability and $\mathbb{P}$ represents the probability implied by the probability distribution Ξ of random vectors $\xi_{t_k}, k = 0, 1, \ldots, n$. The constraint (2.15) says that the final value of our portfolio must not be lower than the final value of a portfolio when adopting the benchmark strategy with a probability greater than or equal to $1 - \alpha$. So combining this constraint with our optimisation problem, we will be aiming to find a strategy with the maximal expected value of the corresponding portfolio given that the strategy is allowed to loose to the benchmark with a probability α at maximum.

This constraint is very easily interpreted, which is its greatest advantage. When a manager is going to decide, whether he adopts the new strategy, one of his first questions will be how much better it is to the current one. This constraint gives him an immediate answer. Its disadvantage might be that if Ξ is discrete with a finite number of atoms it leads to a mixed integer linear program so one will need to restrict the amount of scenarios, as every scenario will lead to a new integer variable.

Next, we will derive the mixed–integer linear programming reformulation of (2.14) which is implied by the condition (2.15). First, we will need to determine what will be the value of a portfolio following the benchmark strategy (2.12) at time n. By setting the decisions as in (2.12), from the formulation in (2.14), we get the following relationships:

$$
\begin{aligned}
x^0_{t_i,t_j}(t_k,\bar{\omega}^s_i) &= \frac{(t_k - t_{k-1}) d_{t_i,t_j}(\bar{\omega}^s_i)}{\sum_{l=i+1}^{j}(t_l - t_{l-1})\exp\{-s^{t_l - t_i}_{t_i}(\bar{\omega}^s_i)(t_l - t_i)\}}, \quad i < k \le j, \quad s \in S_i, \\
Q^0_{t_k}(\bar{\omega}^s_{k-1}) &= \sum_{t_i < t_k \le t_j} x^0_{t_i,t_j}(t_k, a_i(\bar{\omega}^s_{k-1})), \qquad 0 < k \le n, \quad s \in S_{k-1}, \\
X^0_{t_k}(\bar{\omega}^s_k) &= D_{t_k}(\bar{\omega}^s_k), \quad 0 \le k < n, \quad s \in S_k, \qquad B^0_{t_0} = 0, \\
B^0_{t_k}(\bar{\omega}^s_{k-1}) &= \frac{B^0_{t_{k-1}}(\bar{\omega}^s_{k-1})}{P(t_{k-1}, t_k, \bar{\omega}^s_{k-1})} - E_{t_{k-1}} - Q^0_{t_k}(\bar{\omega}^s_{k-1}) + R_{t_k}(\bar{\omega}^s_{k-1}), \qquad (2.16) \\
&\qquad 1 \le k \le n, \quad s \in S_{k-1}, \\
L^0_{t_n}(\bar{\omega}^s_n) &= \sum_{t_i < t_n < t_l \le t_j} P(t_n, t_l, \bar{\omega}^s_n) x^0_{t_i,t_j}(t_l, a_i(\bar{\omega}^s_n)), \quad s \in S_n, \\
V^0_{t_n}(\bar{\omega}^s_n) &= A_{t_n}(\bar{\omega}^s_n) - L^0_{t_n}(\bar{\omega}^s_n) + B^0_{t_n}(a_{n-1}(\bar{\omega}^s_n)), \quad s \in S_n.
\end{aligned}
$$

The equations as specified above describe the dynamics of liabilities under the benchmark strategy. Hence in the parameter $V^0_{t_n}(\bar{\omega}^s_n)$, we should obtain time t_n value of a portfolio of the benchmark strategy for a given scenario of the random vector ω. We should also bear in mind that we should check that the benchmark strategy is "feasible", meaning that it does not run out to debt. That could happen if the earnings are lower than the costs of running the company.

Given that we now know what would be the final value of a portfolio under the benchmark strategy, it is not difficult to employ the condition (2.15) in the linear program (2.14). Considering our problem as a discrete with a finite number of atoms, the fact that we want the probability of not loosing to the benchmark to be greater than $1-\alpha$ corresponds to loosing to the benchmark in $100 \cdot \alpha\%$ of scenarios at maximum. In this case, α is usually set to be in the region $(0.001, 0.1)$. Let us therefore introduce binary variables $z^s, s \in S_n$, where $z^s = 1$ indicates that our strategy lost to the benchmark. This leads to a condition

$$
\begin{aligned}
V^0_{t_n}(\bar{\omega}^s_n) - V_{t_n}(\bar{\omega}^s_n) \le M \cdot z^s, \quad z^s \in \{0,1\}, \quad s \in S_n, \\
\sum_{s \in S_n} z^s \le \alpha \cdot |S_n|, \qquad (2.17)
\end{aligned}
$$

where M is some arbitrary large number (greater than any potential loss). From the constraint (2.17), one can clearly see that if the benchmark does better than the strategy in a scenario s, it forces z^s to jump to one. The second inequality controls the number of such jumps. It is straightforward to deduce that the number of binary variables will be equal to the number of scenarios in the program.

2.3.2 Value–at–Risk Constraint

The next constraint we will be discussing is a well–known Value–at–Risk (VaR). Consider a profit function V, then for a given confidence level $\alpha \in (0,1)$, $\mathrm{VaR}(V)$ measures what is the worst outcome which we experience among the best $100 \cdot \alpha\%$ cases. Or in other words, it says that earnings smaller than $\mathrm{VaR}(V)$ happen only with a low probability $1-\alpha$. These are the implications of a usual definition of VaR as described below. That is more frequently formulated for a loss function.

Definition 1 (J.P. Morgan Risk Metrics, 1995)**.** *Let Y be a random loss function with a cumulative distribution function F_Y and $\alpha \in (0,1)$. Then α*–Value–at–Risk of Y $\mathrm{VaR}_\alpha(Y)$ *is the $\alpha-$quantile of the random variable Y, i.e.*

$$\mathrm{VaR}_\alpha(Y) = F_Y^{-1}(\alpha),$$

where F_Y^{-1} is the quantile function of Y.

Note that in the chance constraint, α is considered close to zero while in VaR, it is thought to be close to one. That is due to the fact that within VaR we consider a loss as our constrained function. Let us define

$$Y_{t_k} = -V_{t_k}, \qquad k = 1, \dots, n.$$

Then Y_{t_n} is the realized loss of the strategy we adopt. One can note that we expect this loss to be negative (as we want to gain a positive profit). The Value–at–Risk constraint is then formulated in the following way:

$$\mathrm{VaR}_\alpha(Y_{t_n}) \le u_\alpha, \qquad \alpha \in (0,1), \;\; u_\alpha \in \mathbb{R}. \tag{2.18}$$

This can be interpreted in such a way that we choose only from strategies which meet that their Value–at–Risk at the confidence level α is smaller than some predetermined amount u_α. The motivation behind such a constraint stems from the usual methodology which is practised in large companies, where the risk department sets limits on Value–at–Risk and the asset–liability management unit is obliged to meet it. Hence we will not be optimising the portfolio with respect to Value–at–Risk, but rather we will be selecting from strategies which meet a predetermined value of VaR. Equivalently, we can rewrite (2.18) as

$$\mathbb{P}\Big(Y_{t_n} \le u_\alpha\Big) \ge \alpha.$$

The above presented result allows us to employ the condition (2.18) in the linear program (2.14). Let us again assume that the distribution Ξ is discrete with a finite number of atoms and recall the notation with scenarios as in the previous case. There we introduced binary variables $z^s, s \in S_n$, each corresponding to one scenario. The meaning of z^s will be also very similar, as now, the value of $z^s = 1$ would indicate exceeding the VaR limit in the scenario s. One can formulate the VaR constraint (2.18) in the following way:

$$-V_{t_n}(\bar{\omega}_n^s) - u_\alpha \leq M \cdot z^s, \quad z^s \in \{0,1\}, \quad s \in S_n,$$
$$\sum_{s \in S_n} z^s \leq (1-\alpha) \cdot |S_n|. \tag{2.19}$$

One could also think of a different loss function Y_{t_k}. For example instead of looking at a loss as a negative profit, we could see the loss as the difference between the value of the benchmark portfolio and the value of the optimal strategy. That would mean to define the loss function as:

$$Y'_{t_k} = -\Big(V_{t_k} - V^0_{t_k}\Big), \qquad k = 1, \dots, n.$$

It is straightforward to see that for $u_\alpha = 0$, the Value–at–Risk constraint (2.18) for such a defined loss function becomes the chance constraint (2.15). This formulation therefore generalises the chance constraint, as it allows us to specify by how much we want the new strategy to be better than the benchmark and at which confidence level. Implementing this constraint to the linear program under a discrete probability distribution for such a loss function is analogical to the chance constraint so we will not be discussing it here.

We see the fact that Value–at–Risk is widely used in the financial industry as the greatest advantage of this constraint because managers are usually quite familiar with it. It has been implemented in various asset–liability stochastic programs, especially in models where the VaR constraint is required by regulations. For example we can mention a recent work by de Oliveira et al. (2017). Comparing (2.17) and (2.19), one can see that the VaR constraint will be computationally similar to the chance constraint thanks to the same number of integer variables. Implementation of such a constraint therefore leads to an integer program where every scenario generates one integer variable. This induces a restriction on the total number of scenarios.

2.3.3 Conditional Value–at–Risk Constraint

The third constraint we will be dealing with is based on the conditional Value–at–Risk (CVaR), which is sometimes called the mean excess loss or the expected shortfall. It differs from VaR in such a way that instead of measuring the worst possible outcome between the best $100 \cdot \alpha\%$ cases, it measures the expected value from the worst $100 \cdot (1-\alpha)\%$ cases. In other words, VaR is not interested in what happens among the worst cases, while CVaR is. Pflug (1999) has described other advantages of CVaR with respect to VaR constraint. Most notably he mentions that CVaR is a coherent risk measure (see Artzner et al. (1999)), so it meets the four properties which are desirable for a measure of risk. On the other hand VaR in general fails to meet the sub–additivity condition. The optimisation properties of CVaR are exploited in Rockafellar and Uryasev (2000), who find CVaR far superior to VaR mostly thanks to its convexity and linearity. They note that optimising a portfolio with respect to a minimum CVaR often leads to a linear programming task. The definition of CVaR is also usually formulated for a loss function.

Definition 2 (Rockafellar and Uryasev, 2002)**.** *Let Y be a random loss function and $\alpha \in (0,1)$. Then we define α–*conditional Value–at–Risk of Y $\mathrm{CVaR}_\alpha(Y)$ *as:*

$$\mathrm{CVaR}_\alpha(Y) = \inf_{a\in\mathbb{R}} \left\{ a \ + \frac{1}{1-\alpha}\mathbb{E}[Y-a]^+ \right\},$$

where $[\cdot]^+$ denotes a positive part of a real number.

Moreover, Rockafellar and Uryasev (2000) showed that if Y has a continuous distribution function, then it holds that

$$\mathrm{CVaR}_\alpha(Y) = \mathbb{E}[Y|Y > \mathrm{VaR}_\alpha(Y)],$$

which justifies the other names such as mean excess loss.

Regarding our asset–liability problem, the loss function can be for CVaR defined in the same way as for VaR. The constraint has the following form:

$$\mathrm{CVaR}_\alpha(Y_{t_n}) = \inf_{a\in\mathbb{R}} \left\{ a \ + \frac{1}{1-\alpha}\mathbb{E}[Y_{t_n}-a]^+ \right\} \le v_\alpha, \qquad (2.20)$$

where $v_\alpha \in \mathbb{R}$ is similarly as in the VaR constraint a given number and $\alpha \in (0,1)$. One can see how this condition will be rewritten in the scenario formulation of the multi–stage stochastic program. For every scenario s, a new variable z^s, bounded by two constraints $z^s \ge Y^s_{t_n} - a$ and $z^s \ge 0$, will be introduced. These conditions will force that $z^s \ge [Y^s_{t_n} - a]^+$. Hence a weighted sum of these new variables $z^s, s \in S_n$ will restrict from above the expected value. If we find at least one a such that

$$a \ + \frac{1}{1-\alpha}\mathbb{E}[Y_{t_n}-a]^+ \le a \ + \frac{1}{1-\alpha}\frac{1}{|S_n|}\sum_{s\in S_n} z^s \le v_\alpha,$$

is satisfied, then we will know that the infimum in (2.20) is smaller then the limiting value v_α and the condition (2.20) will be satisfied. Therefore we will introduce the variable a (which in fact for continuous distributions represents VaR of the portfolio) to the program. Summing it all up, the implementation of the CVaR constraint to the linear program (2.14) preserves its linearity. Hence the corresponding problem will be also a linear programming task. The conditional Value–at–Risk constraint can be formally implemented in the following way. Let $z^s, s \in S_n$ be variables, $z^s \in \mathbb{R}$ and $a \in \mathbb{R}$. The scenario formulation of the constraint (2.20) has the form of:

$$\begin{aligned} z^s \ge -V_{t_n}(\bar{\omega}^s_n) - a, \quad z^s \ge 0, \quad s \in S_n, \\ a + \frac{1}{1-\alpha}\frac{1}{|S_n|}\sum_{s\in S_n} z^s \le v_\alpha, \quad a \ \in \mathbb{R}. \end{aligned} \qquad (2.21)$$

This result confirms the statements we have made during the description of this risk measure that it leads to a linear program which means it should be computationally much more efficient to optimise. Consequently, when employing such a condition, we could afford to generate more scenarios and hence to obtain a better approximation of the original distribution.

2.3.4 Second–Order Stochastic Dominance Constraint

The final constraint we will employ in our asset–liability management problem will be a *second–order stochastic dominance* (SSD) constraint. The stochastic dominance is a modern tool for comparing possible returns of two portfolios of investors. The general idea behind the definition of SSD is to create a partial ordering of returns of portfolios so the first portfolio is ranked better than the second portfolio if no investor with a non–decreasing and concave utility function would prefer to invest into the second portfolio. Usually, the formal definition of SSD is given through so called integrated cumulative probability functions.

Definition 3 (Hadar and Russell, 1969)**.** *Let X be a random variable and let F_X be its cumulative distribution function. Let us define the* integrated probability cumulative distribution function of X *as*

$$F_X^{(2)}(x) = \int_{-\infty}^{x} F_X(u)du.$$

This allows us to state the following definition.

Definition 4 (Hadar and Russell, 1969)**.** *Let V and B be random variables and let $F_V^{(2)}$ and $F_B^{(2)}$ be their integrated probability cumulative distribution functions. Then we say that V dominates B by a* second–order stochastic dominance *($V \succeq_{SSD} B$) if and only if*

$$F_V^{(2)}(y) \leq F_B^{(2)}(y), \qquad \forall y \in \mathbb{R}.$$

The interpretation which we have stated above has been shown in Hadar and Russell (1969), who give the following equivalence:

$$V \succeq_{SSD} B \quad \Leftrightarrow \quad \forall u \in U_2 : \mathbb{E}u(V) \geq \mathbb{E}u(B),$$

where U_2 is a set of all non–decreasing and concave functions, which represent all possible utility functions of non–satiated risk–averse investors. More on the relationship between (first, second and third order) stochastic dominance and utility functions can be found in Levy (1992).

The condition which we want to impose in our asset–liability problem becomes now quite obvious. We would like our new strategy, which generates (random) return V_{t_n} to dominate the benchmark by the second–order stochastic dominance. Moreover, we will generalise a bit the equivalent constraint, which will be formulated together with a parameter b measuring by how much the benchmark is dominated. Such a condition can be formally written as

$$V_{t_n} \succeq_{SSD} V_{t_n}^0 + b, \tag{2.22}$$

and we will refer to it as to the *second–order stochastic b–dominance constraint.* Strategy which meets the constraint (2.22) for $b = 0$ has a very strong interpretation in the sense that every risk–averse manager would prefer (or at worst would be indifferent between) the new strategy to the benchmark strategy. This statement seems very powerful. As the ordering implied by the concept of the second–order stochastic dominance is only partial, it might easily happen that the only portfolio which meets the condition (2.22) with $b = 0$ is the benchmark portfolio itself. For this reason, we might consider also $b < 0$, as this can quantify how far is the new strategy from dominating the benchmark strategy.

There is also a connection between the second–order stochastic dominance and the conditional Value–at–Risk. This was exploited by Kopa and Chovanec (2008) who used results of Ogryczak and Ruszczynski (2002) to show that it holds:

$$V \succeq_{SSD} B \quad \Leftrightarrow \quad \mathrm{CVaR}_\alpha(-V) \leq \mathrm{CVaR}_\alpha(-B) \ \forall \alpha \in (0,1).$$

This can be interpreted in such a way that if we find a strategy V which dominates the benchmark B, then the strategy V will not have a mean of the worst $100 \cdot \alpha\%$ cases smaller than the benchmark B for all $\alpha \in (0,1)$. We are also still yet to formulate the constraint (2.22) in the optimisation model and using the consistence with the CVaR measure could be of use if we assumed a finite number of scenarios. However, the approach we will adopt will be different.

We will make use of the fact that we will approximate the stochastic distribution Ξ by scenarios, which are assumed to have equal probability. Under such settings, one can apply results of Hardy et al. (1934) on majorization, which were later used in stochastic ordering. From there Kuosmanen (2004) has proved the following theorem.

Theorem 1. *Let V and B be random variables with m possible outcomes denoted by $V^1, \ldots, V^m$ and $B^1, \ldots, B^m$, where every outcome has a probability $1/m$. Then $V \succeq_{SSD} B$ if and only if*

$$\begin{aligned} \exists W = \{w_{ij}\}_{i,j=1}^m \in \mathbb{R}^{m \times m} : \quad & w_{ij} \geq 0, \;\; W\mathbf{1} = \mathbf{1}, \mathbf{1}^T W = \mathbf{1}^T, \\ & V^i \geq \sum_{j=1}^m w_{ij} B^j, \quad i = 1, \ldots, m, \end{aligned}$$

where $\mathbf{1}$ is an $m-$dimensional vector of ones.

Proof. For proof see Kuosmanen (2004), Theorem 3. □

Matrix W in the theorem above is called a doubly stochastic matrix, its column sums and row sums are equal to one and its elements are non–negative. One can derive that they are also not greater than one. Theorem 1 has a straightforward application to our asset–liability model. The distribution of the value of a strategy is in the scenario version of the asset–liability program a discrete random variable with every scenario forming an atom of the distribution with the same probability as other atoms. The condition (2.22) is equivalent to the following:

$$\begin{aligned} V_{t_n}(\bar{\omega}_n^{s_i}) - b \geq \sum_{j=1}^{|S_n|} w_{ij} V_{t_n}^0(\bar{\omega}_n^{s_j}), \quad s_i \in S_n, \\ w_{ij} \geq 0, \;\; \sum_{i=1}^{|S_n|} w_{ij} = 1, \;\; \sum_{j=1}^{|S_n|} w_{ij} = 1. \end{aligned} \tag{2.23}$$

The second–order stochastic b–dominance constraint can be therefore rewritten as a linear program with the number of new variables as a square of the number of scenarios. This places some upper bound on the number of scenarios we can use, but given that the corresponding problem is a linear programming task, a reasonable number can still be considered.

3. Scenario Generation in the ALM of a Leasing Company

In this chapter, we will describe the process of generating scenarios for the multi–stage stochastic program of the leasing company. We will introduce the Hull – White model of Hull and White (1990), which will be used for generating scenarios of future interest rates. We will describe the calibration procedure of the model and also the generating process itself. In the second part, we will develop a model which will describe the randomness of demand for the leasing company's products and how it depends on current interest rates. That model will be developed particularly for our analysis and it will assume that the decision stages are equidistant with a one year period.

3.1 The Hull–White Model

The Hull – White model of Hull and White (1990) is one of the most used models for pricing interest rate derivatives. It uses exogenous information in the form of the observed market yield curve to match the current term structure of interest rates. Thanks to this feature, predictions of yields based on this model are close to market expectations, which increases the model's credibility. This model is generally known under the risk neutral measure which itself includes the market price of risk, while the value of the parameters of the model is usually obtained by fitting to the observed market prices of interest rate derivatives. However, there are not enough trades in the Czech market so we have to adopt a different calibration procedure. That will be inspired by Chen and Scott (1993) who estimated the Cox–Ingersoll–Ross model of Cox et al. (1985) by the maximum likelihood method. We will develop an extended version of this method, which will be described in detail. Yet first, we will need to derive the properties of the Hull – White model under the real world measure to justify the estimation of the model's parameters on historical rates.

3.1.1 Introduction of the Model

Let us first describe the definition of the Hull – White under the risk neutral measure — denoted by $\mathbb{Q}$. This part will be taken from other literature, mainly from a well–known book on interest rate modelling of Brigo and Mercurio (2001). Thereafter, we will use the results of Vasicek (1977) and Harrison and Pliska (1981) to derive the corresponding quantities of the model under the real world measure. The model is defined by the following dynamics:

$$dr_t = \Big(\bar{\theta}(s,t) - \bar{\alpha} r_t\Big)dt + \bar{\sigma} dW^Q(t),$$

where $W^Q(t)$ is the standard $\mathbb{Q}$–Brownian motion and $s \leq t$ defines the time when we observe the yield curve. Parameter $\bar{\alpha}$ stands for the mean reversion factor and $\bar{\sigma}$ summarizes the volatility of the short rate r_t. Finally, $\bar{\theta}(s,t)$ is set such that the observed market prices at time s are fitted perfectly, i.e.

$$\bar{\theta}(s,t) = \frac{\partial f^M(s,u)}{\partial u}\bigg|_{u=t} + \bar{\alpha} f^M(s,t) + \frac{\bar{\sigma}^2}{2\bar{\alpha}}\left(1 - e^{-2\bar{\alpha}(t-s)}\right),$$

where $f^M(s,t)$ is the market instantaneous forward rate from time s at time t. This is defined as

$$f^M(s,t) = -\frac{\partial \log P^M(s,u)}{\partial u}\bigg|_{u=t}.$$

$P^M(s,t)$ is the time s market observed price of a zero coupon bond paying one at time t. If we denote

$$\bar{\gamma}(s,t) = f^M(s,t) + \frac{\bar{\sigma}^2}{2\bar{\alpha}^2}\left(1 - e^{-\bar{\alpha}(t-s)}\right)^2,$$

then as stated in Brigo and Mercurio (2001), the distribution of r_t conditioned on the knowledge of r_u, $t \geq u \geq s$ is normal with the conditional expected value $\bar{\mu}_{t|u,s}$ and variance $\bar{\sigma}^2_{t|u}$,

$$\begin{aligned}
\bar{\mu}_{t|u,s} &= e^{-\bar{\alpha}(t-u)} r_u + \bar{\gamma}(s,t) - \bar{\gamma}(s,u) e^{-\bar{\alpha}(t-u)}, \\
\bar{\sigma}^2_{t|u} &= \frac{\bar{\sigma}^2}{2\bar{\alpha}}\left(1 - e^{-2\bar{\alpha}(t-u)}\right).
\end{aligned}$$

Especially for the case $u = s$, we can express the conditional expected value $\bar{\mu}_{t|s}$ and the variance $\bar{\sigma}^2_{t|s}$,

$$\begin{aligned}
\bar{\mu}_{t|s} &= f^M(s,t) + \frac{\bar{\sigma}^2}{2\bar{\alpha}^2}\left(1 - e^{-\bar{\alpha}(t-s)}\right)^2, \\
\bar{\sigma}^2_{t|s} &= \frac{\bar{\sigma}^2}{2\bar{\alpha}}\left(1 - e^{-2\bar{\alpha}(t-s)}\right).
\end{aligned}$$

This implies the transition density $f_r(r_t|r_u, s)$ of r_t conditioned on the value of r_u and the information at time s to be given by

$$f_r(r_t|r_u, s) = \frac{1}{\sqrt{2\pi\bar{\sigma}^2_{t|u}}} \exp\left\{-\frac{\left(r_t - \bar{\mu}_{t|u,s}\right)^2}{2\bar{\sigma}^2_{t|u}}\right\}. \tag{3.1}$$

Let us denote $P(r_t, s, t, T)$ the time t model implied zero coupon bond price maturing at time T with market information observed at time s, $s \leq t \leq T$. This can be determined from the term structure equation derived by Vasicek (1977) and it has an explicit solution which can be written in a form

$$P(r_t, s, t, T) = \exp\{\bar{A}(s,t,T) - \bar{B}(t,T) r_t\}, \tag{3.2}$$

where functions $\bar{A}(s,t,T)$ and $\bar{B}(t,T)$ have the following form:

$$\bar{B}(t,T) = \frac{1 - e^{-\bar{\alpha}(T-t)}}{\bar{\alpha}}, \tag{3.3}$$

$$\bar{A}(s,t,T) = \log\left(\frac{P^M(s,T)}{P^M(s,t)}\right) + \bar{B}(t,T) f^M(s,t) - \frac{\bar{\sigma}^2 \bar{B}(t,T)^2}{4\bar{\alpha}}\left(1 - e^{-2\bar{\alpha}(t-s)}\right). \tag{3.4}$$

The density (3.1) determines the distribution of the short rate r_t (under the risk neutral measure). Given that the annualized continuously compounded time t yield–to–maturity $t+\tau$ can be calculated as

$$y_t(s,\tau) = -\frac{1}{\tau}\log P(r_t, s, t, t+\tau), \tag{3.5}$$

then from (3.2), we can see that $y_t(s,\tau)$ is only a linear function of r_t. And because any linear transformation of a random variable with normal distribution also has normal distribution, it means that $y_t(s,\tau)$ has normal distribution. We will omit the index s in the notation as it should be clear from the context at which time the market observed yield curve was used to define the model.

3.1.2 Change of Measure

The equivalent parametrisation under the real world measure, denoted as $\mathbb{P}$ will be derived in the following lines. The change of measure will be achieved by the Girsanov's theorem, whose results were translated by Harrison and Pliska (1981) into the world of trading in a real market environment. In reality, a drift of a cumulative market price of risk is subtracted from the real world Brownian motion to obtain the relationship between the risk neutral measure and the real world measure. This yields

$$W^Q(t) = W^P(t) - \int_0^t \lambda(u, r_u)du,$$

where W^P is a Brownian motion under the real world measure. We will for simplicity assume that the market price of risk is constant — so that it does not depend on the value of the short rate and that it does not change in time. Hence we set

$$\lambda(t, r_t) = \lambda.$$

Plugging this into the risk neutral dynamics of the model, we obtain

$$dr_t = \Big(\theta(s,t) - \alpha r_t\Big)dt + \sigma dW^P(t), \tag{3.6}$$

where $\sigma = \bar{\sigma}, \alpha = \bar{\alpha}$ and

$$\theta(s,t) = \frac{\partial f^M(s,u)}{\partial u}\bigg|_{u=t} + \alpha f^M(s,t) - \sigma\lambda + \frac{\sigma^2}{2\alpha}\left(1 - e^{-2\alpha(t-s)}\right) = \bar{\theta}(s,t) - \sigma\lambda.$$

One can easily see, that this dynamics has very much the same shape as the risk neutral formulation with only slight difference in the parameter definition. Under such a formulation of the market price of risk the mean reversion factor and the volatility of the short rate are the same. The final thing we need to derive is the bond price under this measure and the distribution of the short rate.

The term structure equation for calculating the bond prices of Vasicek (1977) is the same under both measures (for every market price of risk). That implies that also its solution will be the same and therefore the bond prices under the real world measure are also given by the relationships in (3.3) and (3.4). We only need to express the risk neutral measure parameters by their real world counterparts. But this requires no work as both the mean reversion factor and the volatility of the short rate are the same under the two measures. Next, we will need to solve the stochastic differential equation (3.6). If we employ Ito's lemma to find dynamics of a product $e^{\alpha t} r_t$ and integrate it up, we obtain

$$r_t = e^{-\alpha(t-u)} r_u + \int_u^t e^{-\alpha(t-v)} \theta(s,v) dv + \sigma \int_u^t e^{-\alpha(t-v)} dW^P(v).$$

Let us denote $\mu_{t|u,s}$ the conditional expected value of the short rate given times $t \geq u \geq s$ in the real world measure specification and $\sigma^2_{t|u}$ its standard deviation. We can note, that

$$\mu_{t|u,s} = e^{-\alpha(t-u)} r_u + \int_u^t e^{-\alpha(t-v)} \theta(s,v) dv,$$
$$\sigma^2_{t|u} = \sigma^2 \int_u^t e^{-2\alpha(t-v)} dv.$$

Using that $\theta(s,t) = \bar{\theta}(s,t) - \sigma\lambda$, we obtain

$$\mu_{t|u,s} = e^{-\alpha(t-u)} r_u + \int_u^t e^{-\alpha(t-v)} \bar{\theta}(s,v) dv - \int_u^t e^{-\alpha(t-v)} \sigma\lambda dv.$$

Again, using the connection between the real world and the risk neutral measure, the first two parts together form exactly the equation for the mean in the risk neutral world. So we can use the results for this quantity as given in Brigo and Mercurio (2001). The second integral in the equation is easy to calculate. We can express

$$\mu_{t|u,s} = \bar{\mu}_{t|u,s} - \frac{\sigma\lambda}{\alpha}\left(1 - e^{-\alpha(t-u)}\right). \tag{3.7}$$

Finally, the formula for the variance of the short rate looks:

$$\sigma^2_{t|u} = \bar{\sigma}^2_{t|u} = \frac{\sigma^2}{2\alpha}\Big(1 - e^{-2\alpha(t-u)}\Big). \tag{3.8}$$

Equations (3.7) and (3.8) together with the fact that the distribution of the short rate is normal specifies its distribution. Using this result together with what we discussed in this section, with equations (3.3) and (3.4) specifying the bond price formula and with the relationship between yields–to–maturity and bond prices, as in (3.5), we can determine the distribution of yield–to–maturity under the real world parametrisation of the Hull – White model.

3.1.3 Estimation of the Parameters

We will use the maximum likelihood method for the estimation of the parameters of the Hull – White model. This method was first suggested by Chen and Scott (1993) who analysed the Cox–Ingersoll–Ross model of Cox et al. (1985). We will describe the construction of the likelihood given panel data of observed yields in a more general way. Thereafter we will use it to find the estimates of the Hull – White model parameters for the Czech interest rate curve based on historical data.

Construction of the Log–Likelihood Function

Let $y_t(\tau)$ be a yield at time t with time–to–maturity $\tau > 0$. From (3.2) and (3.5), we can derive the following relationship for the model implied yield–to–maturity:

$$y_t(\tau) = a(s,t,\tau) + b(t,\tau)r_t, \tag{3.9}$$

where

$$a(s,t,\tau) = -\frac{\bar{A}(s,t,t+\tau)}{\tau} \quad \text{and} \quad b(t,\tau) = \frac{\bar{B}(t,t+\tau)}{\tau}.$$

Using relationships (3.3) and (3.4), we can express:

$$b(t,\tau) = b(\tau) = \frac{1-e^{-\alpha\tau}}{\alpha\tau}, \tag{3.10}$$

$$a(s,t,\tau) = F^M(s,t,t+\tau) - b(\tau)f^M(s,t) + \tau\frac{\sigma^2 b(\tau)^2}{4\alpha}\left(1-e^{-2\alpha(t-s)}\right), \tag{3.11}$$

where the function $F^M(s,t,T)$ represents the forward rate at time s from time t with maturity T. This can be be calculated from the relationship

$$F^M(s,t,T) = -\frac{1}{T-t}\log\left(\frac{P^M(s,T)}{P^M(s,t)}\right).$$

In the latter parts, we will denote $a(s,t,\tau)$ just by $a(\tau)$ as the value of the time indices s and t should be clear from the context. We need to remember that $a(\tau)$ and $b(\tau)$ depend also on the model parameters. Let us assume that at time t we observe n yields $y_t(\tau_1),\ldots,y_t(\tau_n)$ with times–to–maturity $\tau_1,\ldots,\tau_n$. We will need a vector equivalent to (3.9). Let us denote the following quantities:

$$\begin{aligned} y_t &= \left(y_t(\tau_1),\ldots,y_t(\tau_n)\right)^T, \\ A &= \left(a(\tau_1),\ldots,a(\tau_n)\right)^T, \\ B &= \left(b(\tau_1),\ldots,b(\tau_n)\right)^T. \end{aligned}$$

Using this, we have

$$y_t = A + Br_t. \tag{3.12}$$

Equation (3.12) summarizes how the short rate influences all yields–to–maturity which we observe, so we can interpret it as the model implied yield curve. It is uniquely determined by the model dynamics, its parameters, times–to–maturity $\tau_1,\ldots,\tau_n$, and by time t.

Our goal is to find a density of the observed yields, while we know the density of r_t. From (3.9) it follows, that given one observed yield–to–maturity $y_t(\tau)$, we can reconstruct the value of the short rate only by inverting the relationship. However, if we do it for different time–to–maturity τ, we are more than likely to obtain a different value of the short rate for the same time t, which does not really make sense. This theoretical relationship in reality does not hold exactly. For that reason, we will introduce a concept that the observed yields are measured with pricing errors. These will summarize fluctuations of real yields around the model implied yield curve. One can imagine that if only one yield is observed, the short rate is determined only by the one inverse relationship. With two yields, there is a room for one error – one yield measured precisely and one with an error. By a straightforward extension of this principle, we need to have $n-1$ errors to be able to fit n observed yields. So let ε_t be a random vector which represents this error. We will assume that it is an $n-1$ dimensional normal, with mean 0_{n-1} and a regular variance–covariance matrix Ω. Moreover, we will also assume that this random vector is independent of the short rate. In the mathematical notation, we write

$$\varepsilon_t \sim \mathcal{N}_{n-1}\Big(0, \Omega\Big).$$

Next, we define C as an $n \times n-1$ matrix which distributes these $n-1$ errors into n components of yields. Combining this with what we saw in (3.12), we get

$$z_t = A + Br_t + C\varepsilon_t = y_t + C\varepsilon_t. \tag{3.13}$$

Let us compare this equation to (3.12). We have added the pricing errors $C\varepsilon_t$ to make the relationship realistic. As we have tried to describe before, in practice it is impossible to find such r_t so we would match the (future) observed yield curve for any value of the parameters in the model dynamics. As we said, y_t can be interpreted as the theoretical (model implied) yield–to–maturity, while z_t represents the observed yield–to–maturity. Our aim will be to find a suitable model for the random variable z_t, which we observe. Equation (3.13) can be rewritten as follows:

$$z_t = A + \begin{pmatrix} B & C \end{pmatrix} \begin{pmatrix} r_t \\ \varepsilon_t \end{pmatrix}, \tag{3.14}$$

where $\begin{pmatrix} B & C \end{pmatrix}$ is an $n \times n$ matrix which defines the transformation between yields and the short rate/errors. Let us denote it $T = \begin{pmatrix} B & C \end{pmatrix}$.

We are particularly interested in the conditional density of z_t at time t given time s. We will make use of the relationship in (3.14) and use a theorem which describes the relationship of densities of transformed random variables. We have $z_t = g(r_t, \varepsilon_t)$, where g is a linear transformation, so it is monotone and also the determinant of the transformation is given by the determinant of the matrix T. Let us denote the conditional density of z_t as $f_z(z_t|z_s)$ and the density of errors as $f_\varepsilon(\varepsilon_t)$. Then, with the use of the independence of the short rate and errors, one can express

$$f_z(z_t|z_s) = \frac{1}{|\det T|} f_r(\hat{r}_t|r_s) f_\varepsilon(\hat{\varepsilon}_t), \tag{3.15}$$

where

$$\begin{pmatrix} \hat{r}_t \\ \hat{\varepsilon}_t \end{pmatrix} = T^{-1}(z_t - A). \tag{3.16}$$

We have derived the conditional density of the observed yield curve implied by the model. Let us now assume that our observations are measured at m periods, at times $t_1 < \cdots < t_m$. In other words, we observe n–dimensional time series $\{z_{t_j}\}_{j=1}^m$. Denote $\theta \in \Theta$ the vector of the parameters of the model. It consists of the parameters of the one–factor short–rate model and of the parameters of the error distribution. Then the likelihood of the data can be expressed as

$$\begin{aligned} L(\theta; z_{t_1}, \dots, z_{t_m}) &= f_{z_{t_1},\dots,z_{t_m}}(z_{t_1}, \dots, z_{t_m}), \\ &= f_z(z_{t_m}|z_{t_{m-1}}) \cdots f_z(z_{t_2}|z_{t_1}) f_z(z_{t_1}). \end{aligned}$$

In general, one does not consider the density $f_z(z_{t_1})$ so we will ignore it and work only with the conditional likelihood function. We do not expect omitting the density of the first observation to affect the analysis, as it is a usual practise to take the first observation in the series as given. The conditional likelihood can be written as

$$L(\theta; z_{t_1}, \dots, z_{t_m}|z_{t_1}) = \prod_{j=2}^{m} f_z(z_{t_j}|z_{t_{j-1}}) = \prod_{j=2}^{m} \frac{1}{|\det T|} f_r(\hat{r}_{t_j}|r_{t_{j-1}}) f_\varepsilon(\hat{\varepsilon}_{t_j}).$$

From the above equation we can derive the form of the conditional log–likelihood function.

$$\ell(\theta) = -(m-1)|\det T| + \sum_{j=2}^{m} \log f_r(\hat{r}_{t_j}|r_{t_{j-1}}) + \sum_{j=2}^{m} \log f_\varepsilon(\hat{\varepsilon}_{t_j}), \tag{3.17}$$

where $\hat{r}_{t_j}$ and $\hat{\varepsilon}_{t_j}$ are given as in (3.16). The above equation shows us how to construct the likelihood for the observed yields–to–maturity. To complete the log–likelihood, we must specify the matrix of the transformation T, we need to know its inverse and its determinant. This means to specify the matrix C as values of A, B which appear in (3.16) along with the density f_r are determined from the short–rate dynamics. The correlation matrix of errors f_ε can be also specified by ourselves as one of the assumptions of the model.

In our data analysis, we have employed the following parametrization:

$$C = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ \frac{1}{\sqrt{n-1}} & \frac{1}{\sqrt{n-1}} & \cdots & \frac{1}{\sqrt{n-1}} \end{pmatrix}, \quad \Omega = \sigma_e^2 \begin{pmatrix} 1 & \rho & \rho & \cdots & \rho \\ \rho & 1 & \rho & \cdots & \rho \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \rho & \cdots & \rho & 1 & \rho \\ \rho & \rho & \cdots & \rho & 1 \end{pmatrix}, \tag{3.18}$$

where $0 \le \rho < 1$. The choice was done based on our empirical experience which we gained from analysing the dataset (to be introduced shortly). Basically, our goal was to choose C and Ω so we could assume that the fitted errors $\hat{\varepsilon}_{t_j}$ were generated from such a density. To be able to use C as in (3.18), we need to be able to calculate the inverse of the matrix T and its determinant. The matrix T itself can be expressed as

$$T = \begin{pmatrix} b_1 & 1 & 0 & \cdots & 0 \\ b_2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ b_{n-1} & 0 & 0 & \cdots & 1 \\ b_n & \frac{1}{\sqrt{n-1}} & \frac{1}{\sqrt{n-1}} & \cdots & \frac{1}{\sqrt{n-1}} \end{pmatrix},$$

where $b_1, \dots, b_n$ are the elements of B as in (3.12). Let us denote

$$a = b_n - \frac{1}{\sqrt{n-1}} \sum_{i=1}^{n-1} b_i. \tag{3.19}$$

Then, after applying some linear algebra calculations, one can express the inverse matrix as

$$T^{-1} = \frac{1}{a} \begin{pmatrix} -\frac{1}{\sqrt{n-1}} & -\frac{1}{\sqrt{n-1}} & \cdots & -\frac{1}{\sqrt{n-1}} & 1 \\ a + \frac{b_1}{\sqrt{n-1}} & \frac{b_1}{\sqrt{n-1}} & \cdots & \frac{b_1}{\sqrt{n-1}} & -b_1 \\ \frac{b_2}{\sqrt{n-1}} & a + \frac{b_2}{\sqrt{n-1}} & \cdots & \frac{b_2}{\sqrt{n-1}} & -b_2 \\ \vdots & & \ddots & & \vdots \\ \frac{b_{n-1}}{\sqrt{n-1}} & \frac{b_{n-1}}{\sqrt{n-1}} & \cdots & a + \frac{b_{n-1}}{\sqrt{n-1}} & -b_{n-1} \end{pmatrix}. \tag{3.20}$$

Next, we need to calculate the determinant of T. By applying the Laplace expansion to the first column and then to the last row, we get:

$$\begin{aligned} \det T &= \sum_{i=1}^{n-1} (-1)^{i+1} b_i \cdot \frac{1}{\sqrt{n-1}} (-1)^{n-1+i} \cdot 1 + (-1)^{n+1} b_n \cdot 1, \\ &= \sum_{i=1}^{n-1} (-1)^{n+2i} b_i \frac{1}{\sqrt{n-1}} + (-1)^{n+1} b_n, \\ &= (-1)^n \left(\frac{1}{\sqrt{n-1}} \sum_{i=1}^{n-1} b_i - b_n \right). \end{aligned}$$

Taking the absolute value of the determinant yields

$$|\det T| = \left| \frac{1}{\sqrt{n-1}} \sum_{i=1}^{n-1} b_i - b_n \right| = |a|. \tag{3.21}$$

Looking at the inverse matrix and at the determinant, we can see that the necessary condition for this parametrisation to work is that $a \neq 0$. Then T is regular and it enables us to use this parametrisation for the definition of the model. The fact that $a \neq 0$ does not appear to be a limiting restriction. The next step in expressing the likelihood in (3.17) is to derive the density of errors. For that reason, we will need to calculate the inverse and the determinant of the matrix Ω.

The calculation of the determinant is rather complicated, but we can adopt the following trick. We know, that it is going to be a polynomial in ρ of the degree $n-1$, so it can only have $n-1$ real roots at maximum. Moreover, if for some ρ_1, the kernel of a matrix has dimension k, then the polynomial $(\rho - \rho_1)^k$ divides the determinant. One can see that for $\rho_1 = 1$ the kernel has got dimension $n-2$. The final parameter, for which Ω is singular is for $\rho_1 = (-1)/(n-2)$. Using this, and the knowledge that the absolute coefficient in the polynomial is $\sigma_e^{2(n-1)}$ – a product of diagonal elements, we can express the determinant as

$$\det \Omega = \sigma_e^{2(n-1)} (1-\rho)^{n-2} (1 + (n-2)\rho). \tag{3.22}$$

Expressing the inverse matrix is tricky too. In this case, we will assume that it contains only two different elements. An off–diagonal element c/σ_e^2 and an on–diagonal element d/σ_e^2. If we multiply such a defined matrix with the matrix Ω

and set it equal to the identity matrix, we get the two following equations:

$$\begin{aligned} d + (n-2)\rho c &= 1, \\ c + \rho d + (n-3)\rho c &= 0. \end{aligned}$$

Solving this, one obtains

$$\begin{aligned} c &= -\frac{\rho}{1-\rho} \cdot \frac{1}{1+(n-2)\rho}, \\ d &= \frac{1}{1-\rho} - \frac{\rho}{1-\rho} \cdot \frac{1}{1+(n-2)\rho} = \frac{1}{1-\rho} + c. \end{aligned}$$

In matrix notation, we get

$$\Omega^{-1} = \frac{1}{\sigma_e^2} \begin{pmatrix} d & & c \\ & \ddots & \\ c & & d \end{pmatrix}. \tag{3.23}$$

Combining these results, we can express the density of $\varepsilon_t \sim \mathcal{N}_{n-1}(0, \Omega)$ as

$$\begin{aligned} \log f_\varepsilon(\varepsilon_t) = &-\frac{n-1}{2} \log 2\pi - \frac{n-2}{2} \log(1-\rho) - \frac{1}{2} \log(1+(n-2)\rho) \\ &- \frac{n-1}{2} \log \sigma_e^2 - \frac{1}{2\sigma_e^2} \sum_{i=1}^{n-1} d\varepsilon_{t,i}^2 - \frac{1}{2\sigma_e^2} \sum_{i,j=1, i \neq j}^{n-1} c\varepsilon_{t,i}\varepsilon_{t,j}, \\ \log f_\varepsilon(\varepsilon_t) \propto &-\frac{n-2}{2} \log(1-\rho) - \frac{1}{2} \log(1+(n-2)\rho) - \frac{n-1}{2} \log \sigma_e^2 \\ &- \frac{1}{2\sigma_e^2} \frac{1}{1-\rho} \sum_{i=1}^{n-1} \varepsilon_{t,i}^2 + \frac{1}{2\sigma_e^2} \frac{\rho}{1-\rho} \cdot \frac{1}{1+(n-2)\rho} \sum_{i,j=1}^{n-1} \varepsilon_{t,i}\varepsilon_{t,j}. \end{aligned}$$

Assume now that we observe *iid* random variables $\hat{\varepsilon}_{t_2}, \ldots, \hat{\varepsilon}_{t_m} \sim \mathcal{N}_{n-1}(0, \Omega)$. If we denote $N = (m-1)(n-1)$ and

$$\hat{E}^2 = \sum_{k=2}^{m} \sum_{i=1}^{n-1} \hat{\varepsilon}_{t_k,i}^2, \qquad \hat{E}^1 = \sum_{k=2}^{m} \sum_{i,j=1}^{n-1} \hat{\varepsilon}_{t_k,i}\hat{\varepsilon}_{t_k,j}, \tag{3.24}$$

then the log–likelihood function of such a time series has the form of

$$\begin{aligned} \ell_2(\theta_2) = &\sum_{k=2}^{m} \log f_\varepsilon(\varepsilon_{t_k}) \\ \propto &-(m-1)\frac{n-2}{2} \log(1-\rho) - (m-1)\frac{1}{2} \log(1+(n-2)\rho) \\ &-(m-1)\frac{n-1}{2} \log \sigma_e^2 - \frac{1}{2\sigma_e^2} \frac{1}{1-\rho} \hat{E}^2 + \frac{1}{2\sigma_e^2} \frac{\rho}{1-\rho} \cdot \frac{1}{1+(n-2)\rho} \hat{E}^1, \end{aligned}$$

where $\theta_2 = (\rho, \sigma_e^2)^T$ are the parameters of the error distribution. Analogically, we will denote $\theta_1 = (\alpha, \sigma, \lambda)^T$ the Hull – White model parameters. Here, we would like the reader to realise the connection between $\ell_2(\theta_2)$ and the third summand in (3.17). Next, we will try to maximize the log–likelihood $\ell_2(\theta_2)$ over the parameter space for σ_e^2 and ρ. We will derive the score equations and then we will set them equal to zero and try to get explicit expressions for the maximum likelihood (ML) estimates. First, we start with the parameter σ_e^2.

$$\frac{\partial \ell_2(\theta_2)}{\partial \sigma_e^2} = -\frac{N}{2\sigma_e^2} + \frac{\hat{E}^2}{2\sigma_e^4}\frac{1}{1-\rho} - \frac{\hat{E}^1}{2\sigma_e^4}\frac{\rho}{1-\rho} \cdot \frac{1}{1+(n-2)\rho}.$$

This gives us an estimate

$$\hat{\sigma}_e^2 = \frac{1}{N}\frac{1}{1-\rho}\left(\hat{E}^2 - \frac{\rho}{1+(n-2)\rho}\hat{E}^1\right). \tag{3.25}$$

Equation (3.25) tells us, what is the ML for any ρ. To be able to use this result, we need to find an estimate for ρ itself.

$$\begin{aligned}
\frac{\partial \ell_2(\theta_2)}{\partial \rho} &= \frac{(m-1)(n-2)}{2(1-\rho)} - \frac{(m-1)(n-2)}{2(1+(n-2)\rho)} - \frac{1}{2\sigma_e^2}\frac{\hat{E}^2}{(1-\rho)^2} \\
&\quad + \frac{1}{2\sigma_e^2}\frac{\hat{E}^1}{(1-\rho)^2} \cdot \frac{1-\rho+\rho}{1+(n-2)\rho} - \frac{1}{2\sigma_e^2}\frac{\hat{E}^1}{1-\rho} \cdot \frac{(n-2)\rho}{(1+(n-2)\rho)^2}, \\
&= \frac{(m-1)(n-2)}{2(1-\rho)} - \frac{(m-1)(n-2)}{2(1+(n-2)\rho)} \\
&\quad - \frac{(m-1)(n-1)}{1-\rho}\frac{1}{2\sigma_e^2}\frac{1}{N}\frac{1}{1-\rho}\left(\hat{E}^2 - \frac{\rho}{1+(n-2)\rho}\hat{E}^1\right) \\
&\quad + \frac{1}{2\sigma_e^2}\frac{\hat{E}^1}{1-\rho} \cdot \frac{1}{1+(n-2)\rho} - \frac{1}{2\sigma_e^2}\frac{\hat{E}^1}{1-\rho} \cdot \frac{(n-2)\rho}{(1+(n-2)\rho)^2}.
\end{aligned}$$

Plugging into the last equation the ML estimate $\hat{\sigma}_e^2$ as in (3.25), and setting the score equation to 0, we obtain:

$$\begin{aligned}
0 &= \frac{(m-1)(n-2)}{2(1-\rho)} - \frac{(m-1)(n-2)}{2(1+(n-2)\rho)} - \frac{(m-1)(n-1)}{2(1-\rho)} \\
&\quad + \frac{1}{2\sigma_e^2}\frac{\hat{E}^1}{1-\rho} \cdot \frac{1}{1+(n-2)\rho} - \frac{1}{2\sigma_e^2}\frac{\hat{E}^1}{1-\rho} \cdot \frac{(n-2)\rho}{(1+(n-2)\rho)^2}, \\
0 &= -\frac{m-1}{1-\rho} - \frac{(m-1)(n-2)}{(1+(n-2)\rho)} + \frac{1}{\hat{\sigma}_e^2}\frac{\hat{E}^1}{1-\rho} \cdot \frac{1}{1+(n-2)\rho}\left(1 - \frac{(n-2)\rho}{1+(n-2)\rho}\right), \\
0 &= -\frac{m-1}{1-\rho} - \frac{(m-1)(n-2)}{(1+(n-2)\rho)} + \frac{(m-1)(n-1)}{\left(\hat{E}^2 - \frac{\rho}{1+(n-2)\rho}\hat{E}^1\right)} \cdot \frac{\hat{E}^1}{(1+(n-2)\rho)^2}, \\
0 &= -\frac{n-1}{(1-\rho)} + \frac{n-1}{\left(\hat{E}^2 - \frac{\rho}{1+(n-2)\rho}\hat{E}^1\right)} \cdot \frac{\hat{E}^1}{1+(n-2)\rho}.
\end{aligned}$$

Continuing the algebraic manipulations, we get

$$\frac{1+(n-2)\rho}{(1-\rho)} = \frac{\hat{E}^1}{\left(\hat{E}^2 - \frac{\rho}{1+(n-2)\rho}\hat{E}^1\right)},$$

$$(1 + (n-2)\rho)\frac{\hat{E}^2}{\hat{E}^1} - \rho = 1 - \rho.$$

Finally, we obtain

$$\hat{\rho} = \frac{1}{n-2}\frac{\hat{E}^1 - \hat{E}^2}{\hat{E}^2}. \tag{3.26}$$

To finish our job, plugging our result from (3.26) into (3.25) yields:

$$\hat{\sigma}_e^2 = \frac{1}{N}\hat{E}^2. \tag{3.27}$$

That is an extremely nice result, which one would hardly expect given the complexity of the likelihood function. Moreover, both formulas look very reasonable and have strong interpretation. Plugging formulas (3.26) and (3.27) into the likelihood function $\ell_2(\theta_2)$ yields (after a bit of algebra)

$$\max_{\theta_2} \ell_2(\theta_2) \propto -\frac{(m-1)(n-2)}{2}\log\left((n-1)\hat{E}^2 - \hat{E}^1\right) - \frac{m-1}{2}\log \hat{E}^1. \tag{3.28}$$

Consequently, instead of the log–likelihood function for the vector parameter θ as in (3.17), we can consider only the profile log–likelihood function for parameters θ_1. This should increase our chances in getting valid maximum likelihood estimates as we have basically eliminated two parameters from a need of an uncertain numerical optimisation. The profile log–likelihood has the following form:

$$\begin{aligned} \ell(\alpha, \sigma, \lambda) \propto &- (m-1)\log|a\ | - \frac{m-1}{2}\log\left(2\pi\sigma_{dt}^2\right) - \frac{1}{2\sigma_{dt}^2}\sum_{j=2}^{m}\left(\hat{r}_{t_j} - \mu_{t_j|t_{j-1}}\right)^2 \\ &- \frac{(m-1)(n-2)}{2}\log\left((n-1)\hat{E}^2 - \hat{E}^1\right) - \frac{m-1}{2}\log \hat{E}^1, \end{aligned} \tag{3.29}$$

where we have implicitly assumed that we have equidistant observations. This allowed us to write σ_{dt}^2 as a common variance of the short rate for all observations. Summing this up, when one employs the parametrization of the model (3.13) as we did in (3.18), he can express the profile log–likelihood of the parameters of the Hull – White model as stated in (3.29). This can be then optimised given historical observations of rates with the use of the relationships (3.16) and (3.24).

Data Analysis

Our aim is to find a model which will represent the interest rate in the Czech market, hence we will fit the model to a yield curve derived from the price of swaps on PRIBOR (Prague Interbank Offered Rate). From these prices, one can construct the zero–coupon yield curve by a method called bootstrapping (see Hull (2008)). In the following, we will call these rates swap rates. For estimation, we have used monthly data from January 2009 up to November 2016 — together 95 observations. The beginning of the data set was decided to be after the financial crisis during which the market view on risk factors affecting the value of interest rates changed rapidly. The data we we will use for estimation are plotted in Figure 3.1.

To optimise the profile log–likelihood (3.29), we have used software *R* and the function *nmkb* in the package *dfoptim* of Varadhan et al. (2011). It performs a kind of a heuristic search around the parameter space to find an extreme. However, as the point has met the *Karush–Kuhn-Tucker* conditions, it confirms that we found a local maximum. Moreover, we have plotted a number of 2D cuts across the parameter space which all suggested that we have reached the global maximum too. The optimal values of the parameters were the following:

$$\hat{\alpha} = 0.036963242, \qquad \hat{\sigma} = 0.005958489, \qquad \hat{\lambda} = 0.901977764. \tag{3.30}$$

To illustrate the implied rates by the calibrated Hull – White model, we will provide two figures. These will be connected to the asset–liability model analysis in the sense that the figures will be based on the tree structure which will be used in the stochastic program. The stages of the program will be $0, 1, \ldots, 6$, years, so we will need to obtain scenarios for these times. We will start from one node representing the current time. This node will have eight descendants. This number will be halved in the next stage and from year two, all nodes will have only two descendants. The decision for such a structure was to keep the number of scenarios relatively small while we wanted to have better description of the distribution of the first two stages than if we kept the number of descendants from every node constant.

The generation process of the yield curve scenarios will not be random, as we will choose them so they represent the "known" — model implied distribution as well as possible. The process will be as follows. Assume that we are to generate

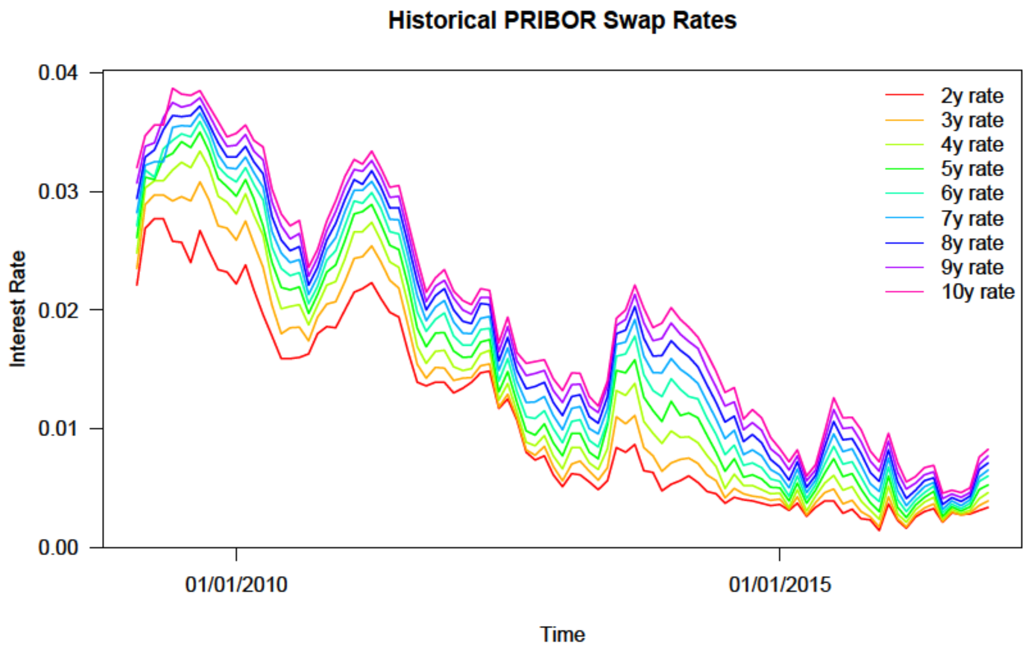

Figure 3.1: Monthly observations of the yield curve derived from the swap rates on PRIBOR from January 2009 to November 2016.

n scenarios from a given univariate distribution. Then we split the interval $[0,1]$ into n equidistant intervals $[0,1/n],\dots,[1-1/n,1]$ and from these middle points will be chosen. Hence we select a sequence

$$\left\{\frac{1}{2n}+\frac{k}{n},\qquad k=0,\dots,n-1\right\}.$$

The short rate scenarios will be created as the quantiles of the short rate distribution corresponding to the probabilities in the sequence. Long rates and zero–coupon bond prices will be derived from the short rate based on (3.9) and (3.2).

The two afore–mentioned figures are the following. First, in Figure 3.2, we show node values of a yield–to–maturity with time–to–maturity one year. One might be surprised with the number of scenarios with negative yield, but we think that given which yields are currently observed around in Europe, the probability of negative yields in the Czech market are not unrealistic. For example, in the final stage, little bit less than 15% of yields suggest negative return.

The second figure depicts scenario's yield curves for all 8 year–one nodes. Their shape is shown in Figure 3.3. One can notice that the rates increase with maturity for every scenario, which is what economists would expect. One can compare this figure to Figure 3.2, as the colors of the scenarios are the same in both figures. These should mainly illustrate rationality of the scenarios as they reflect the current market environment quite well. This is important to realise before the analysis as it forms the foundation for the optimisation procedure on which strong and interpretable results can be build.

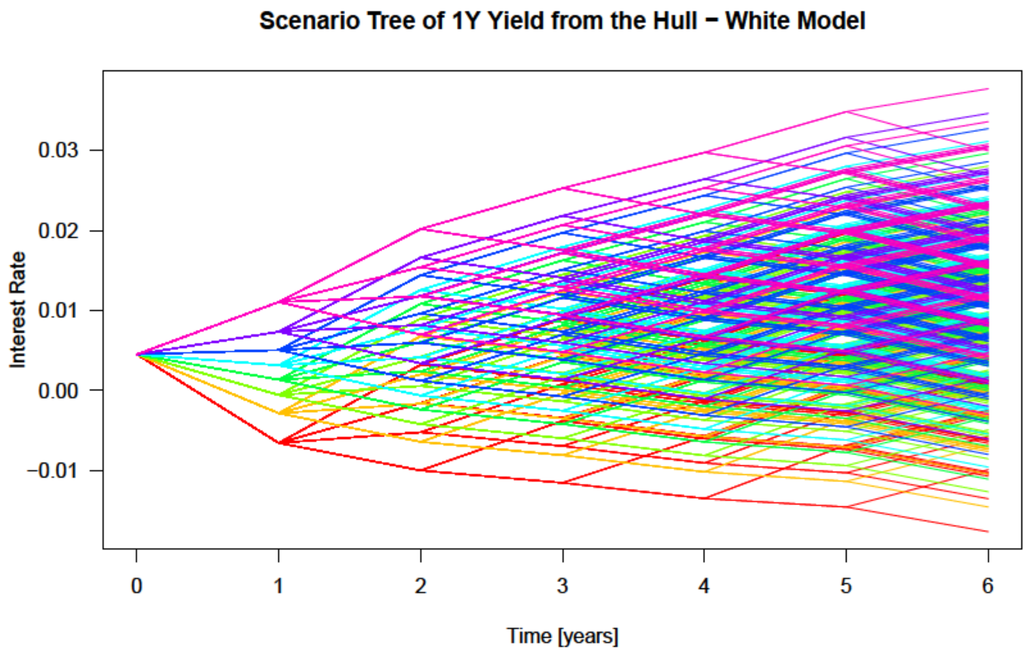

Figure 3.2: Scenario values of one year interest rate in the Czech market in a tree structure used in the optimisation problem.

3.2 Demand for Loans of the Leasing Company

The second model we will need to develop in order to be able to generate scenarios is a model which describes the randomness of demand for loans of the leasing company. We will focus to describe the demand for loans in the Czech market. Therefore, we have downloaded data of the total volume of closed leasing loans in the Czech market which are published by Czech National Bank. These are yearly data starting in 2005 and ending in 2016 – altogether 12 consecutive observations. The volumes are split into three categories, loans with a maturity less than one year, with a maturity between 1 and 5 years and with a maturity longer than 5 years. The data are showed in Table 3.1.

As we do not have many observations and our model should tell us the demand for loans at a given time with a given maturity, we have to employ some simplifying assumptions in order to specify a usable model. It will be formulated as follows. We will assume that clients can borrow a loan only with maturity one, two, three, four or five years. Demand $d_{t_i,t_j}, t_j \in (t_i+1, \ldots, t_i+5)$ will have a gamma distribution with a mean μ_i and a shape parameter a, so it will not depend on the maturity t_j. We will use the parametrization of the gamma family with the shape parameter a and the scale parameter s whose probability density function has the following form.

$$f_{\Gamma(a,s)}(x) = \frac{s^{-a}}{\Gamma(a)} x^{a-1} e^{-x/s}, \qquad x \geq 0,$$

where $\Gamma(\cdot)$ is the gamma function. The relationship between mean μ of a gamma distribution and parameters a, s is that $\mu = a \cdot s$. We will model the demand

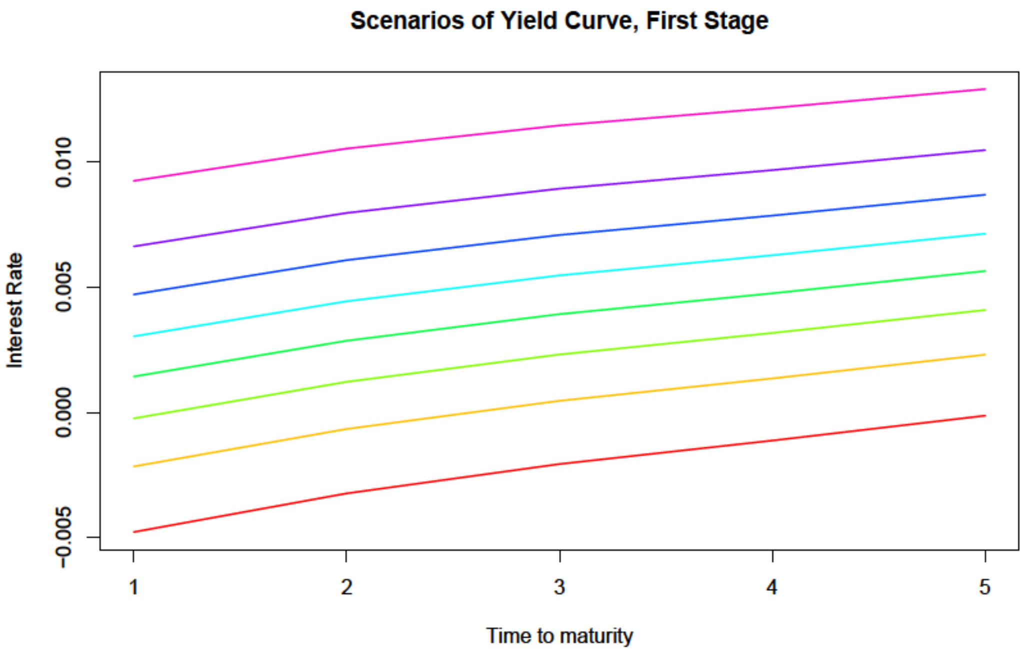

Figure 3.3: Yield curves of the first–stage nodes exhibit increasing rate with maturity.

to depend only on the value of one year interest rate $y^1_{t_i}$ such that

$$\mu_i = \exp\{\beta_0 + \beta_1 y^1_{t_i}\},$$

where $\beta_0, \beta_1 \in \mathbb{R}$ are parameters. Their estimates can be obtained by fitting a generalized linear model of a gamma family with a log link to our data set. The fit was done so we took the sum of volumes in each year divided by a factor of five as a response and the one year rate as an explanatory variable. Even though this approach is not completely statistically correct, we believe it can give us reasonable estimates, especially with such a simple data we have. The model estimates were the following:

$$\hat{\beta}_0 = 10.362, \qquad \hat{\beta}_1 = -0.020, \qquad \hat{a} = 85.6. \tag{3.31}$$

To illustrate the distribution of the demand for loans and the effect of the one year yield on it, we present Figure 3.4. It shows the density of a distribution of the demand for loans implied by the model with two different values of one year yield. The interpretation of the coefficient β_1 is that an increase in the one year interest rate of 1% causes the expected value of the demand for loans to move by a factor e^{β_1}. As the estimate of β_1 is negative, the effect is that the volume decreases by $e^{-0.02} \doteq 0.98$. This slight decrease can be seen in the figure.

As this model was fitted to the total demand for loans in the Czech market, in the program, we will consider a company which has a share of $\phi \in (0, 1)$ of the market. So under the parametrization of the gamma family we stated above, the demand $d_{t_i,t_j}, t_j \in (t_i + 1, \dots t_i + 5)$ will have a distribution of

$$d_{t_i,t_j} \sim \phi\Gamma(\hat{a}, \hat{s}_i), \qquad \hat{s}_i = \frac{1}{\hat{a}} \exp\{\hat{\beta}_0 + \hat{\beta}_1 y^1_{t_i}\}, \tag{3.32}$$

where estimates $\hat{a}, \hat{\beta}_0$ and $\hat{\beta}_1$ are given as in (3.31). The value of the one year interest rate at time t_i will be determined from the node value of this yield as given

Year	$m < 1$ year	$1 \geq m < 5$ years	$m \geq 5$ years	1 year rate [%]
2016	32 951	117 757	32 633	0.44
2015	29 651	106 791	30 248	0.46
2014	24 325	97 345	26 162	0.51
2013	25 375	93 079	24 390	0.60
2012	27 010	97 053	29 718	0.87
2011	28 455	100 973	28 764	1.73
2010	27 827	98 430	31 450	1.80
2009	27 682	90 550	25 370	2.13
2008	38 045	106 664	26 164	3.93
2007	32 603	100 464	13 887	4.23
2006	31 789	82 911	16 483	2.81
2005	30 411	80 982	15 367	2.53

Table 3.1: Volumes of leasing loans closed in a given year with maturity m in millions of CZK and corresponding 1 year PRIBOR, source: ARAD database of CNB.

by the Hull – White model, which was described in the previous section.

Sampling of scenarios will be completely different to the technique adopted in the Hull – White model. In this case, we will employ random generation of scenarios from the known distribution. That was decided from the reason that the distribution of demand is known only after we obtain the node value for the one year yield, so choosing quantiles of the distribution is out of question as we only need to obtain one value for demand with a given maturity. We will assume that within a single node, demands for different maturities are independent. Hence once we calculate the value of the one year yield in the node, we generate five samples from the corresponding gamma distribution to obtain demands with one, two, three, four and five years maturity. That is a very simple approach, hence for real analysis, it would be worth to make a more detailed statistical survey where we would learn how demand for leasing loans behave in more detail.

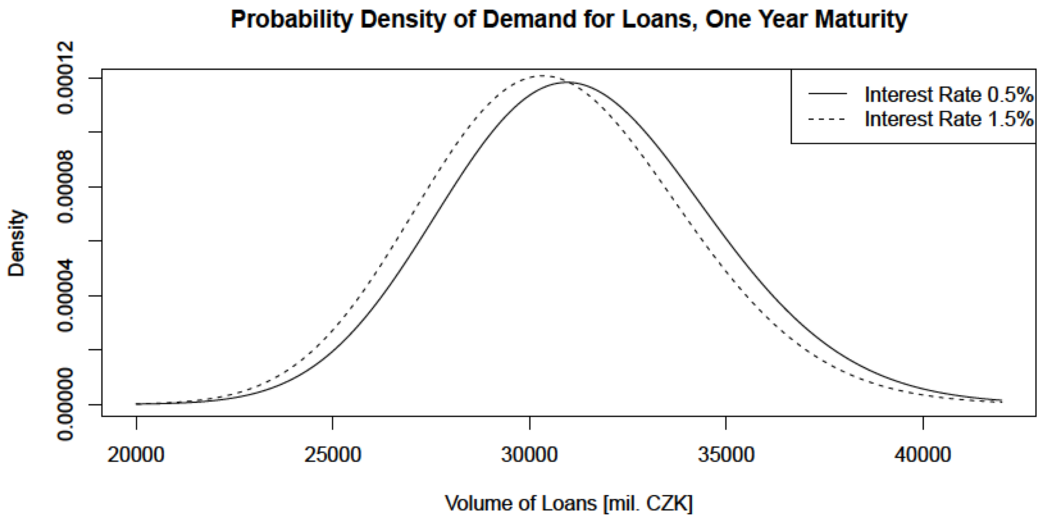

Figure 3.4: The density of demand for loans for two different values of interest rate.

4. Model Analysis

In this chapter, we will complete the model specification by defining the final details of our model and by specifying the values of its parameters. Subsequently, we will analyse this model and present the results of the stochastic programming problem for all risk constraints with a plethora of values of the parameters employed.

4.1 Model Specification and Parameter Settings

Let us now introduce the model in a more concrete form, which can be used for the analysis. I would like to stress that the following specifications will mean a change in the notation. All symbols will still have their previous meanings, but the indexation will be a bit different. We do this for more simple notation of the optimisation problem we will be analysing later in the chapter.

As we hinted at the beginning of Chapter 3, our decision stages will take place with yearly period. We will assume that we are now at year 0 and that the decisions can be done once a year at times $0, 1, \dots, n-1$, where n is the investment horizon. In the analysis, we will assume $n = 6$. Clients will have a possibility to get a loan with maturity j years, where $j \leq m = 5$. The demand for such a loan at time i with time–to–maturity j will be denoted $d_{i,j}$. Here, please note the difference in the notation with the previous chapters, as demand was before denoted by two time indices. The first denoted time at which the deal was closed and the second denoted time at which the deal matured. Now, we employ an equivalent notation with the first index denoting the time when the deal was agreed, while the second index denotes the length of the deal. Hence we have demand

$$d_{i,j} \geq 0, \qquad i = 0, \dots, n-1, \;\; j = 1, \dots, m.$$

Similarly, the leasing company will be permitted to borrow money from the bank at the same times with the same maturities, yielding

$$x_{i,j} \geq 0, \qquad i = 0, \dots, n-1, \;\; j = 1, \dots, m.$$

Next, changing the notation of time from $(t_0, t_1, \dots, t_n, \dots)$ to $(0, 1, \dots, n, \dots)$ we can define

$$\bar{r}_{i,j}(\bar{\omega}_i^s) = \sum_{l=i+1}^{i+j} \exp\{-r_i^l(\bar{\omega}_i^s) \cdot l\}, \qquad 0 \leq i < n, \;\; 1 \leq j \leq m, \;\; s \in S_i.$$

The object $\bar{r}_{i,j}(\bar{\omega}_i^s)$ determines the proportion of a loan closed at time i maturing in j years which will be paid in a single instalment. The motivation for its definition can be seen in (2.2). Similarly, we will denote

$$\bar{s}_{i,j}(\bar{\omega}_i^s) = \sum_{l=i+1}^{i+j} \exp\{-s_i^l(\bar{\omega}_i^s) \cdot l\}, \qquad 0 \leq i < n, \;\; 1 \leq j \leq m, \;\; s \in S_i$$

and

$$p_{i,j}(\bar{\omega}_i^s) = P(i, i+j, \bar{\omega}_i^s) = \exp\{-y_i^j(\bar{\omega}_i^s) \cdot j\}, \qquad 0 \leq i \leq n, \;\; 0 < j \leq 5, \;\; s \in S_i.$$

Employing such a notation, the model formulated in (2.14) can be rewritten as:

$$
\begin{aligned}
&\max_{x_{i,j}(\bar{\omega}_i^s)} \quad \frac{1}{|S_n|}\sum_{s\in S_n} V_n(\bar{\omega}_n^s) \qquad (4.1)\\
\text{s.t. } & R_k(\bar{\omega}_{k-1}^s) = \sum_{i=[k-m]^+}^{k-1}\sum_{j=k-i}^{m} \frac{d_{i,j}(a_i(\bar{\omega}_{k-1}^s))}{\bar{r}_{i,j}(a_i(\bar{\omega}_{k-1}^s)))}, \qquad 1\le k\le n, s\in S_{k-1},\\
& Q_k(\bar{\omega}_{k-1}^s) = \sum_{i=[k-m]^+}^{k-1}\sum_{j=k-i}^{m} \frac{x_{i,j}(a_i(\bar{\omega}_{k-1}^s))}{\bar{s}_{i,j}(a_i(\bar{\omega}_{k-1}^s)))}, \qquad 1\le k\le n, s\in S_{k-1},\\
& D_k(\bar{\omega}_k^s) = \sum_{j=1}^{m} d_{k,j}(\bar{\omega}_k^s), \quad X_k(\bar{\omega}_k^s) = \sum_{j=1}^{m} x_{k,j}(\bar{\omega}_k^s), \;\; 0\le k<n, s\in S_k,\\
& B_0 = X_0(\omega_0^s) - D_0(\omega_0^s), \quad s\in S_0,\\
& B_k(\bar{\omega}_k^s) = \frac{B_{k-1}(a_{k-1}(\bar{\omega}_k^s))}{p_{k-1,1}(a_{k-1}(\bar{\omega}_k^s))} - E_{k-1} + X_k(\bar{\omega}_k^s) - Q_k(a_{k-1}(\bar{\omega}_k^s))\\
&\qquad + R_k(a_{k-1}(\bar{\omega}_k^s)) - D_k(\bar{\omega}_k^s), \quad 1\le k<n, \;\; s\in S_k,\\
& B_n(\bar{\omega}_{n-1}^s) = \frac{B_{n-1}(\bar{\omega}_{n-1}^s)}{p_{n-1,1}(\bar{\omega}_{n-1}^s)} - E_{n-1} + R_n(\bar{\omega}_{n-1}^s) - Q_n(\bar{\omega}_{n-1}^s), \;\; s\in S_{n-1},\\
& A_n(\bar{\omega}_n^s) = \sum_{i=[n-m+1]^+}^{n-1}\sum_{j=n-i+1}^{m}\sum_{l=n-i+1}^{j} p_{n,l+i-n}(\bar{\omega}_n^s)\frac{d_{i,j}(a_i(\bar{\omega}_n^s))}{\bar{r}_{i,j}(a_i(\bar{\omega}_n^s)))}, \quad s\in S_n,\\
& L_n(\bar{\omega}_n^s) = \sum_{i=[n-m+1]^+}^{n-1}\sum_{j=n-i+1}^{m}\sum_{l=n-i+1}^{j} p_{n,l+i-n}(\bar{\omega}_n^s)\frac{x_{i,j}(a_i(\bar{\omega}_n^s))}{\bar{s}_{i,j}(a_i(\bar{\omega}_n^s)))}, \quad s\in S_n,\\
& V_n(\bar{\omega}_n^s) = A_n(\bar{\omega}_n^s) - L_n(\bar{\omega}_n^s) + B_n(a_{n-1}(\bar{\omega}_n^s)), \quad s\in S_n,\\
& B_k(\bar{\omega}_k^s) \ge 0, \quad 0\le k<n, \;\; s\in S_k, \qquad x_{i,j}^s \ge 0, \quad 0\le i\le j, \;\; s\in S_i,
\end{aligned}
$$

where we again employed the new notation for time $(0, 1, \ldots, n, \ldots)$ instead of the notation $(t_0, t_1, \ldots, t_n, \ldots)$ which has been seen in the previous sections. The symbol $[\cdot]^+ = \max(0, \cdot)$ denotes the positive part of a real number. Next, we will reformulate the equations which specify the value of the benchmark strategy. We would like again to stress here that there are no decisions to be made so given the node values of demand and interest rate, the value of the benchmark strategy's portfolio is fixed. We can calculate it as follows:

$$
\begin{aligned}
Q_k^0(\bar{\omega}_{k-1}^s) &= \sum_{i=[k-m]^+}^{k-1}\sum_{j=k-i}^{m} \frac{d_{i,j}(a_i(\bar{\omega}_{k-1}^s))}{\bar{s}_{i,j}(a_i(\bar{\omega}_{k-1}^s)))}, \qquad 1\le k\le n, \;\; s\in S_{k-1},\\
X_k^0(\bar{\omega}_k^s) &= D_k(\bar{\omega}_k^s), \quad 0\le k<n, \;\; s\in S_k, \qquad B_{t_0}^0 = 0, \qquad (4.2)\\
B_k^0(\bar{\omega}_{k-1}^s) &= \frac{B_{k-1}^0(\bar{\omega}_{k-1}^s)}{p_{k-1,1}(\bar{\omega}_{k-1}^s)} - E_{k-1} - Q_k^0(\bar{\omega}_{k-1}^s) + R_k(\bar{\omega}_{k-1}^s), \;\; 1\le k\le n, \;\; s\in S_{k-1},\\
L_n^0(\bar{\omega}_n^s) &= \sum_{i=[n-m+1]^+}^{n-1}\sum_{j=n-i+1}^{m}\sum_{l=n-i+1}^{j} p_{n,l+i-n}(\bar{\omega}_n^s)\frac{d_{i,j}(a_i(\bar{\omega}_n^s))}{\bar{s}_{i,j}(a_i(\bar{\omega}_n^s)))}, \quad s\in S_n,\\
V_n^0(\bar{\omega}_n^s) &= A_n(\bar{\omega}_n^s) - L_n^0(\bar{\omega}_n^s) + B_n^0(a_{n-1}(\bar{\omega}_n^s)), \quad s\in S_n.
\end{aligned}
$$

To complete the model definition, we will reformulate the risk constraints which were introduced in Section 2.3. For the first three constraints, we will need a level of significance $\alpha \in (0, 1)$. We also mentioned earlier that it is a usual habit to consider α close to zero for the chance constraints and close to one for the Value–at–Risk and the conditional Value–at–Risk constraints. Moreover, symbol M will denote some sufficiently large (larger than any potential loss) number.

- The chance constraint:

$$\begin{aligned} V_n^0(\bar{\omega}_n^s) - V_n(\bar{\omega}_n^s) \leq M \cdot z^s, \quad z^s \in \{0,1\}, \quad s \in S_n, \\ \sum_{s \in S_n} z^s \leq \alpha \cdot |S_n|. \end{aligned} \tag{4.3}$$

- The Value–at–Risk constraint with threshold u_α:

$$\begin{aligned} -V_n(\bar{\omega}_n^s) - u_\alpha \leq M \cdot z^s, \quad z^s \in \{0,1\}, \quad s \in S_n, \\ \sum_{s \in S_n} z^s \leq (1-\alpha) \cdot |S_n|. \end{aligned} \tag{4.4}$$

- The conditional Value–at–Risk constraint with threshold v_α:

$$\begin{aligned} z^s \geq -V_n(\bar{\omega}_n^s) - a, \quad z^s \geq 0, \quad s \in S_n, \\ a + \frac{1}{1-\alpha}\frac{1}{|S_n|}\sum_{s \in S_n} z^s \leq v_\alpha, \quad a \in \mathbb{R}. \end{aligned} \tag{4.5}$$

- The second–order stochastic b–dominance constraint for dominating the benchmark portfolio plus a fixed amount b:

$$V_n(\bar{\omega}_n^{s_i}) - b \geq \sum_{j=1}^{|S_n|} w_{ij} V_n^0(\bar{\omega}_n^{s_j}), \quad s_i \in S_n, \tag{4.6}$$

$$w_{ij} \geq 0, \ \sum_{i=1}^{|S_n|} w_{ij} = 1, \ \sum_{j=1}^{|S_n|} w_{ij} = 1.$$

The value of the parameters of the risk constraints will be specified later on as we will study the results of the problem for a variety of parameter values. We will look how optimal solution changes with different limits for the VaR and the CVaR constraints, a different value of α in the chance constraint and also with a different value of b in the SSD constraint. The process of obtaining scenarios was described in Chapter 3, where we should add that the market share of the leasing company as in (3.32) was considered to be 1%. The last sets of parameters yet unspecified were introduced in Section 2.1. These include the spread $s(\tau)$ and the mark–up $m(\tau)$, which determine the value for which clients and the leasing company can borrow, and the cost of running the company $E_k, k = 0, \ldots, n-1$.

The values of the spread and the mark–up were determined from a real data of a company CS Autoleasing in order to ensure that we work with realistic numbers. We have analysed the differences between the rates the company gets from the bank and the market rates to obtain the following spreads (in %):

$$s(1) = 0.41, \ s(2) = 0.49, \ s(3) = 0.56, \ s(4) = 0.58, \ s(5) = 0.59.$$

The sequence confirms what we expected as the spread increases with maturity. It might be considered very small, which could be a consequence of the fact that CS Autoleasing borrows money from Ceska Sporitelna, which is its owner. Calculating the mark–up appeared to be more complicated as the rates clients get depend quite heavily on the client itself and on the product he buys. It could also be more correct to subtract from the mark–up the amount the leasing company charges the client due to its risk profile. That is because this money serves as an insurance against a default of the client, which we do not consider in the analysis. However, the mark–up is basically used only for calculating the final amount of money which the company makes and its value does not really give any advantage to any of the strategies. For that reason, we decided to increase a bit the costs of running the company (so it includes a budget for covering unfulfilled loans) and leave the rates unchanged. The mark–up was then obtained as the difference between the weighted average of rates clients get and the rates the company gets from the bank. We set (in %)

$$m(1) = 4.3, \;\; m(2) = 5.9, \;\; m(3) = 4.4, \;\; m(4) = 4.2, \;\; m(5) = 4.2.$$

One might be surprised by the value of the mark–up for loans with two year maturity, but the rates charged by the leasing company were the highest for two year loans in four out of five years we analysed. We had quite a strong argument to respect the numbers we were given. Finally, the costs of running the company were set 50 mil. CZK in the first year, 100 mil. CZK in the second year and 125 mil. CZK in other years. The decision for such costs stemmed from two reasons. First, we had to make sure that the benchmark strategy does not run out to debt in the first periods. Earnings of the company in the first stage are lower than in the other stages because fewer loans are repaid by that time, which must have been taken into account. Second, we wanted to make sure that the benchmark strategy does not accumulate too much money on its bank account, as then that money would loose value over time of the program which would favour the optimal strategy. Under such settings, we have created a real and a fairly competitive environment for the benchmark strategy. So if we find an optimal strategy which can be considered better than the benchmark strategy, it will be down only to more efficient borrowing.

4.2 Results and Sensitivity Analysis

In the following section, we will present the results of our model of the leasing company for different risk constraints. At the beginning, we will analyse a model without them to find out which parameter values make sense to consider for the risk constraints. Thereafter, we will show the results of the stochastic program for all the considered values of the parameters and analyse how the company's decisions change if we employ more strict limits.

4.2.1 No–Risk Constraint Problem

First, we run the stochastic program with the formulation of the problem as given in (4.1) together with (4.2). The optimal solution then states how a manager should behave if his only objective is to maximize expected profit. The inclusion of (4.2) allows us to compare the results of the optimal strategy with the value of the benchmark portfolio. The linear program had 7008 variables and 4563 equations. The model formulation was specified in GAMS, while the solution was provided by CPLEX in less than a second. It suggested us to borrow entirely in one year loans in the first four years of the program, while in the fifth year, which was the final decision stage, it chose to borrow mostly for two years. The expected value of such a strategy was 316.46 mil. CZK, while the benchmark achieved a mean return of 294.47 mil. CZK.

However, this number tells only a part of the story about the distribution of returns of the optimal portfolio. We might be interested in many other characteristics which describe the riskiness of such a strategy; for example how many times the optimal strategy performs worse than the benchmark strategy. We might want to learn about the comparison of VaR and CVaR at different level of significance, which could suggest us what value we should use for the risk constraints in subsequent models too. These questions will be further discussed. Now, for more detailed comparison of the two strategies, we show in Figure 4.1 their histogram of returns (= values of the portfolio).

There, one can see a clear cost for obtaining a greater mean value of the portfolio by the optimal strategy. It is the uncertainty in its final value, as now, the value of the portfolio has got much greater variance. However, the lower tails of both strategies are similar, with the lowest values being just above 230 mil. CZK. It is also of interest to know when the benchmark strategy betters the opti-

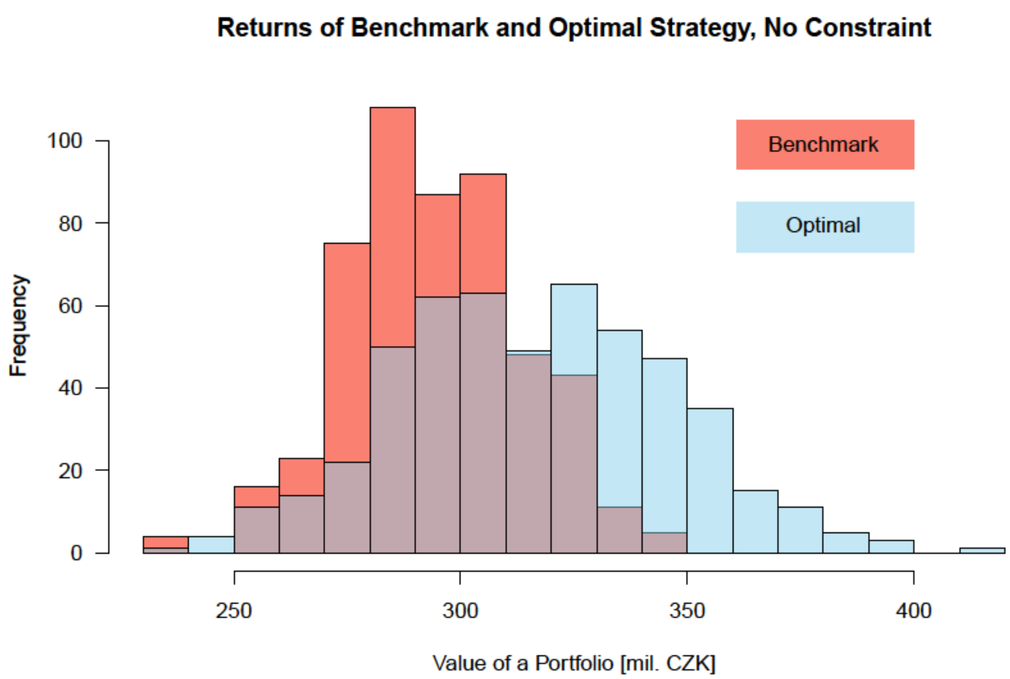

Figure 4.1: Comparison of returns of the benchmark and the optimal strategy in the no–risk constraint problem.

mal strategy. We would expect the optimal portfolio to loose most when interest rates increase, as in that case borrowing money in the next period becomes more expensive than it is today. This surmise is supported by Figure 4.2, where differences between final portfolio values of the optimal and of the benchmark strategy for every scenario are shown against the final–stage one–year interest rate. We would like to add that the color of the points in that figure corresponds to the color of scenarios in Figures 3.2 and 3.3.

In Figure 4.2 one can see a clear dependence between the difference in return and the value of interest rate. Roughly speaking, once the interest rate exceeds two percent, the optimal strategy looses against the benchmark strategy. The relationship between the difference of portfolio values and one year interest rate looks to be linear, as the optimal strategy gains/looses circa 20 mil. CZK per 1% change in interest rate. Summing this up, the benchmark strategy betters the optimal strategy in little bit less than 19% of cases. This hints us that we should consider parameter α in the chance constraint to be between $\langle 0, 0.19\rangle$. That is because this parameter controls how often we permit the optimal strategy to do worse than the benchmark strategy, so any value larger than 0.19 implies that the chance constraint will not be active.

Similarly, we need to find a reasonable range for the parameters in other risk constraints. For VaR and CVaR, we will consider the level of significance to be strictly 0.95. The Value–at–Risk of the benchmark portfolio at this level is -262.32 mil. CZK, while the corresponding quantity of the optimal portfolio is -269.10 mil. CZK. We can see, that this risk measure ranks the optimal portfolio

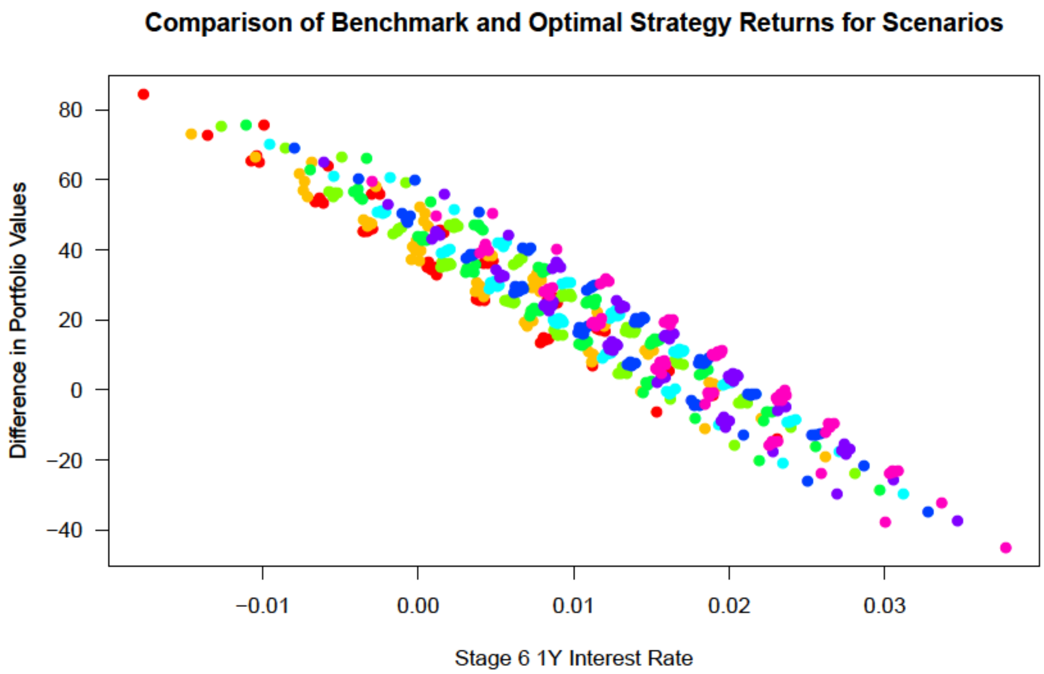

Figure 4.2: Differences in portfolio values plotted against one–year interest rate in the final stage.

to be less risky than the benchmark portfolio. Therefore, if we restricted the considered portfolios to be the ones which have VaR smaller than the benchmark portfolio, the new optimal solution would still be the currently optimal portfolio. In our analysis, we will consider portfolios with even smaller VaR than the VaR of the currently optimal portfolio. The lower bound for VaR will be determined based on the computational difficulties, which might arise in solving mixed integer programs, and by its feasibility.

Values of CVaR of the benchmark portfolio and of the optimal portfolio are -254.88 and -256.71 mil. CZK respectively. Also in this case, the benchmark portfolio is concluded to be more risky than the optimal portfolio. For conditional Value–at–Risk, it will make sense to consider potential losses smaller than -256 mil. CZK. Because the CVaR constraint can be rewritten as a linear program, the lower bound for the loss will be determined by the feasibility of the problem, as we should not experience any computational difficulties.

Finally, we would like to touch the SSD constraint and discuss what values of parameter b come into considerations. In general, the most interesting value is $b = 0$, which corresponds to finding a portfolio which SSD dominates the benchmark portfolio. The goal could also be to find b as large as possible so the problem is still feasible. This could be an achievable task as the stochastic program with the SSD constraint can be rewritten as a linear program.

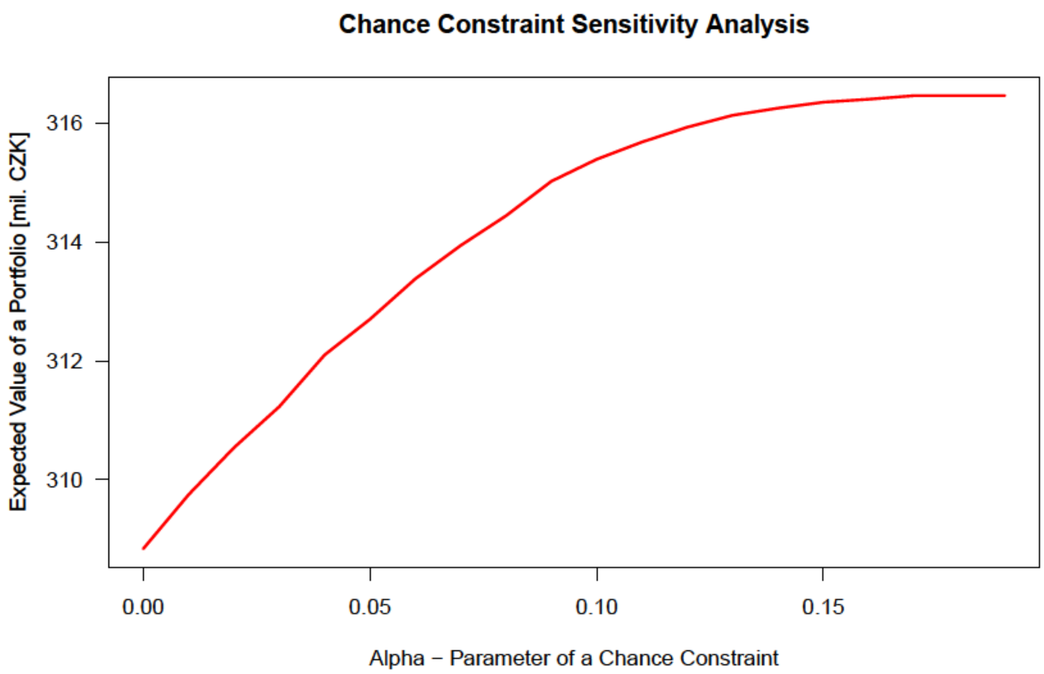

Figure 4.3: Dependence of the expected value of the optimal solution on the value of α in the chance constraint.

4.2.2 Chance Constraint

In this section, we will consider the model formulation to consist from the main part as in (4.1) and (4.2) together with the chance constraint as given in (4.3). This constraint includes a parameter α, which has the meaning that we allow the model suggested strategy to get beaten by the benchmark strategy in $100 \cdot \alpha\%$ cases at maximum. We have seen that it only makes sense to consider values of $\alpha \in \langle 0, 0.19 \rangle$. In the case of $\alpha = 0.19$, this constraint will not be active and the solution of the problem will be the same as without it. On the other hand, setting $\alpha = 0$ means that we want the new strategy not to loose to the benchmark at every scenario. The chance constraint enlarges the original formulation by introducing new 512 binary variables and 513 equations. Even though the relative increase in the number of variables is not high, the fact that they are binary results in a massive increase in the computational difficulty.

We analysed the problem for 20 different values of α. These points formed an equidistant sequence with endpoints 0 and 0.19 with the difference of 0.01. To summarize the results, we give the two following figures. First, in Figure 4.3, we show how the expected value of the optimal portfolio changes with the value of the parameter α. Logically, for smaller values of α (which corresponds to a more strict condition) we obtain lower expected values of the optimal strategies. This decrease in the expected value can be interpreted as the cost of making sure that we perform better than the benchmark portfolio. Ultimately, for $\alpha = 0$, the expected value of the strategy is 308.84 mil. CZK, which is still around 14.4 mil. CZK higher than the one of the benchmark strategy.

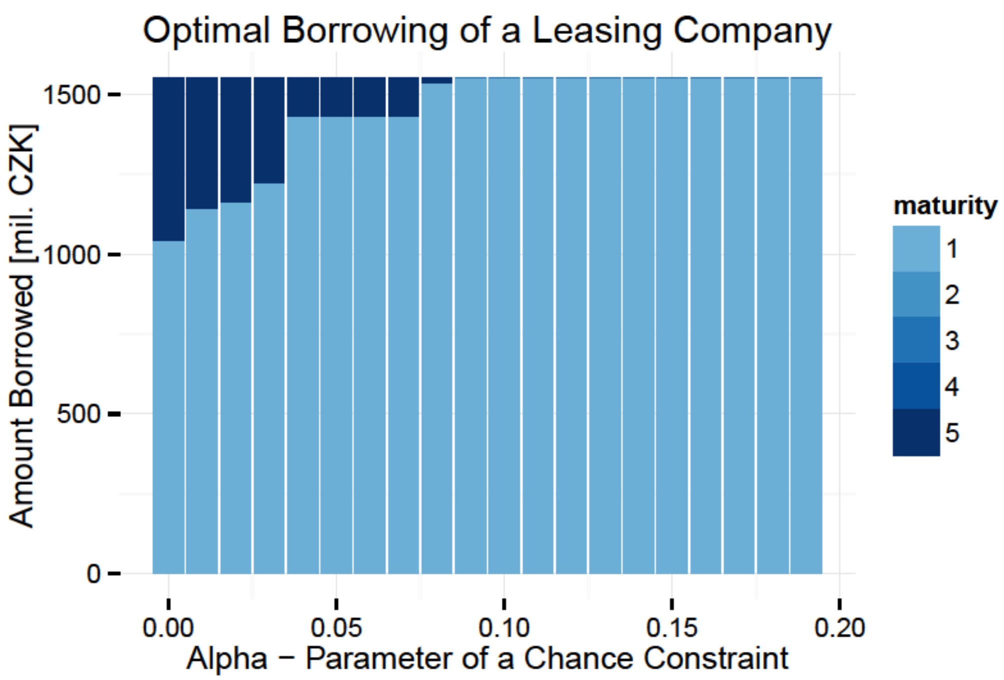

Figure 4.4: Suggested borrowings by the optimal solution with different value of α in the chance constraint.

Second, Figure 4.4 shows the first stage decisions of the problem's optimal solution for every considered value of α in the chance constraint. One can see that for high values of α, one does not need to adjust borrowings (in comparison to the no–risk constraint optimal strategy) in the first stage as it is enough to wait until next information is observed. In contrast, as α moves closer to 0, the solution suggests to close some five year loans with the bank so the strategy is more robust to unfavourable interest rate movements. When we studied the structure of borrowings in more detail, we found out that the model proposes behaviour one would expect. For scenarios with in general lower interest rates, it suggested to focus more on borrowing one year loans. On the other hand, when interest rates increased, the solution recommended to borrow greater proportion of money in longer term loans.

4.2.3 Value–at–Risk Constraint

The next constraint we will be analysing will be the Value–at–Risk constraint introduced in Section 2.3.2. In this constraint, we must choose the level of significance α, which is usually considered close to one, and the VaR limit $u_\alpha \in \mathbb{R}$ so a loss greater than u_α happens only with a low probability $1-\alpha$. We consider our loss function to be the negative value of the portfolio; hence this condition can be translated such that we require the value of the portfolio to be greater than or equal to $-u_\alpha$ in at least $100 \cdot \alpha$ percent of cases. We will consider only $\alpha = 0.95$. Therefore for example setting $u_\alpha = -280$ mil. CZK means that we require our strategy to do worse than 280 mil. CZK at 5% cases at maximum.

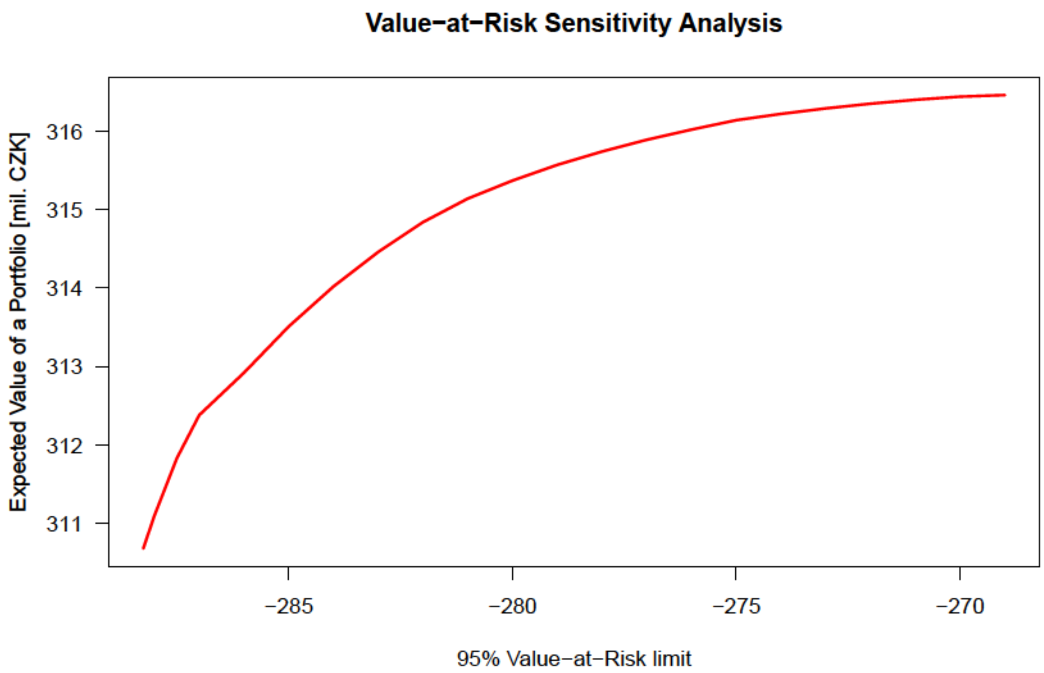

Figure 4.5: Dependence of the expected value of the optimal solution on the value of u_α in the Value–at–Risk constraint.

As one can see from (4.3) and (4.4), the computational complexities of the VaR constraint and the chance constraint are the same. To illustrate the results of a problem specified by the model (4.1) and the constraint (4.4) we again provide two figures. In Figure 4.5 we show a development of the maximal expected value from all strategies meeting the given VaR limit. We found the lowest limit for which the stochastic program of the asset–liability model had a feasible solution to be -288.25 mil. CZK. The x-axis therefore shows the optimal portfolio returns for all reasonable VaR limits.

We should add that all optimal strategies (for all VaR limits) suggested in the first stage to close loans with the bank only with one year maturity. That is a difference to the chance constraint, where some strategies suggested to borrow also for five years. The reason for this is that closing long–term loans has two consequences on the final value of the portfolio. First, it increases the portfolio value in the worst scenarios (as these have increasing value of interest rate so borrowing for long term saves money in the future) and second it decreases the overall expected value as we loose money by closing loans with greater interest. The chance constraint forces the worst scenarios to do much better, hence it forces the strategy to close these long term deals. On the other hand, the VaR constraint does not care about the value of the very worst returns, hence it could afford to borrow only for one year in the first stage. Consequently, Value–at–Risk optimal strategies generate generally higher optimal expected value than the chance constraint strategies.

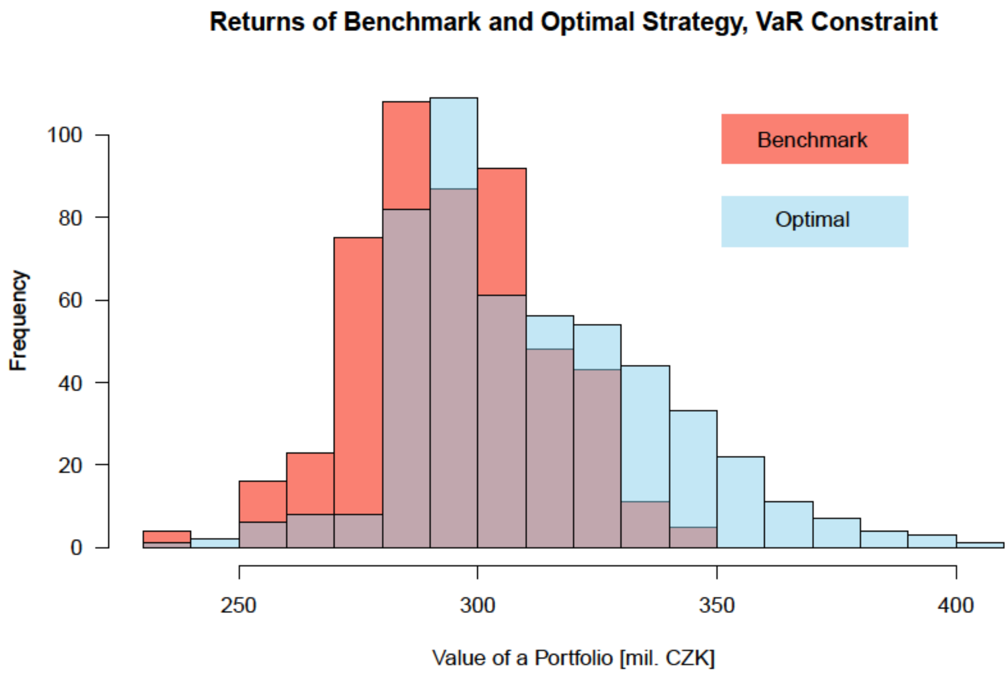

Figure 4.6: Histogram of portfolio values at the investment horizon according to the optimal solution of a VaR constraint with the limit of $u_\alpha = -287$ mil. CZK.

The fact that even for very low VaR limits the optimal strategy still produces some bad portfolio returns is illustrated in Figure 4.6. There, we show a comparison of histograms of portfolio returns of the benchmark strategy and of the optimal strategy with the VaR limit of -287 mil. CZK. We should also add that the 0.95 VaR of the benchmark strategy is -162.32 mil. CZK — 25 mil. CZK higher than the considered VaR limit for the asset–liability model. Moreover, under such a VaR limit, we were able to find a strategy which outperforms the expected return of the benchmark strategy by 18 mil. CZK.

4.2.4 Conditional Value–at–Risk Constraint

The third constraint we will be discussing is the conditional Value–at–Risk constraint specified by (4.1) and (4.5). It has the advantage of a linear representation in the stochastic program and hence the optimal solution of the problem is found quickly and efficiently. The corresponding problem has been solved by CPLEX in a matter of seconds, which suggests that we could afford larger scenario tree to describe the stochastic variables more accurately. We again considered $\alpha = 0.95$ as the level of significance for this constraint and the highest CVaR limit of $v_\alpha = -256$ mil. CZK. The lowest CVaR limit for which there was a feasible solution was -277.25 mil. CZK, which is around 23 mil. CZK lower than the CVaR of the benchmark strategy. To illustrate, how the optimal expected value of a portfolio changes with different CVaR limits, we present Figure 4.7.

The shape of the curve in Figure 4.7 is not surprising as we again see decreasing expected value of the optimal solution with decreasing limit of CVaR.

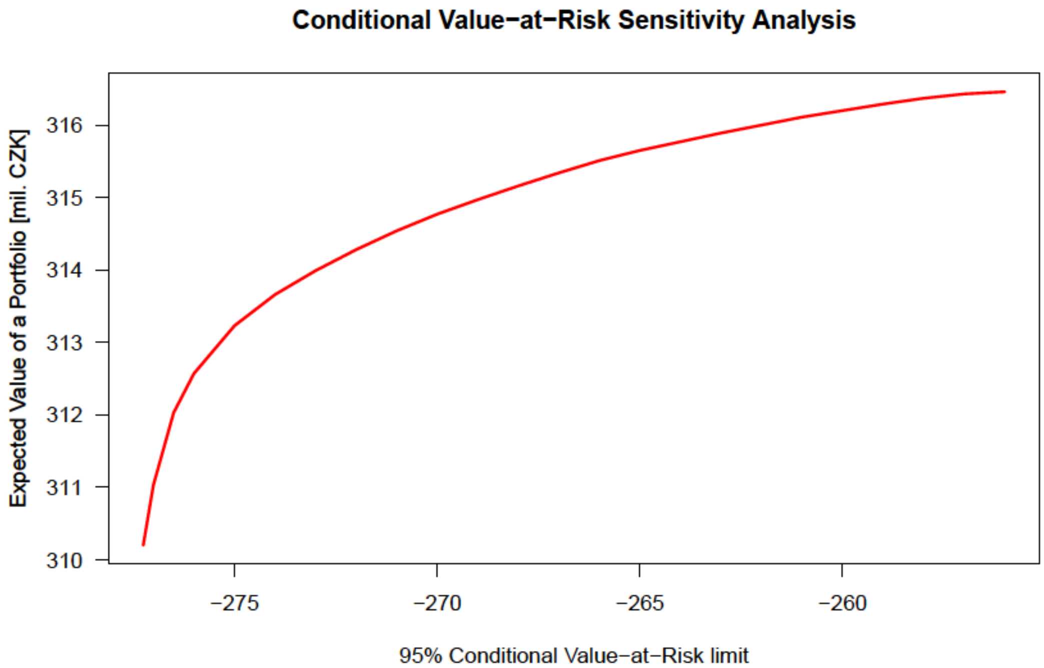

Figure 4.7: Dependence of the expected value of the optimal solution on the value of v_α in the conditional Value–at–Risk constraint.

However, the curvature looks to be greater here than in Figures 4.3 and 4.5, which implies that lowering the CVaR limit for high CVaR limits is much less costly than for low CVaR limits. The difference between the optimal expected values of portfolios with CVaR limits -256 mil. CZK and -266 mil. CZK is just 0.95 mil. CZK, while between limits -267 mil. CZK and -277 mil. CZK it is 4.31 mil. CZK.

The optimal strategies suggested by the model proposed to close in the first stage only one year loans. This again underlines how costly it is to borrow with longer maturity and as long as there is not an urgent need to improve the worst case results, the solution tries to avoid it. In Figure 4.8, we show a comparison of returns of the benchmark strategy and of the optimal strategy with the CVaR limit of -275 mil. CZK. This limit basically eliminated returns smaller than 270 mil. CZK, while it still managed to keep relatively high number of returns greater than 350 mil. CZK — that is more than the best return of the benchmark strategy.

4.2.5 Second–Order Stochastic Dominance Constraint

The final constraint which we have employed in our model was the second–order stochastic b–dominance constraint which we introduced in Section 2.3.4. The parameter b describes how much we want to dominate the benchmark strategy. Nevertheless, we are mostly interested in the case $b = 0$. We have also mentioned that incorporating such a constraint introduces many new variables and inequalities into the model formulation; so even thought the enlarged program is still linear, we might experience computational difficulties. The corresponding linear

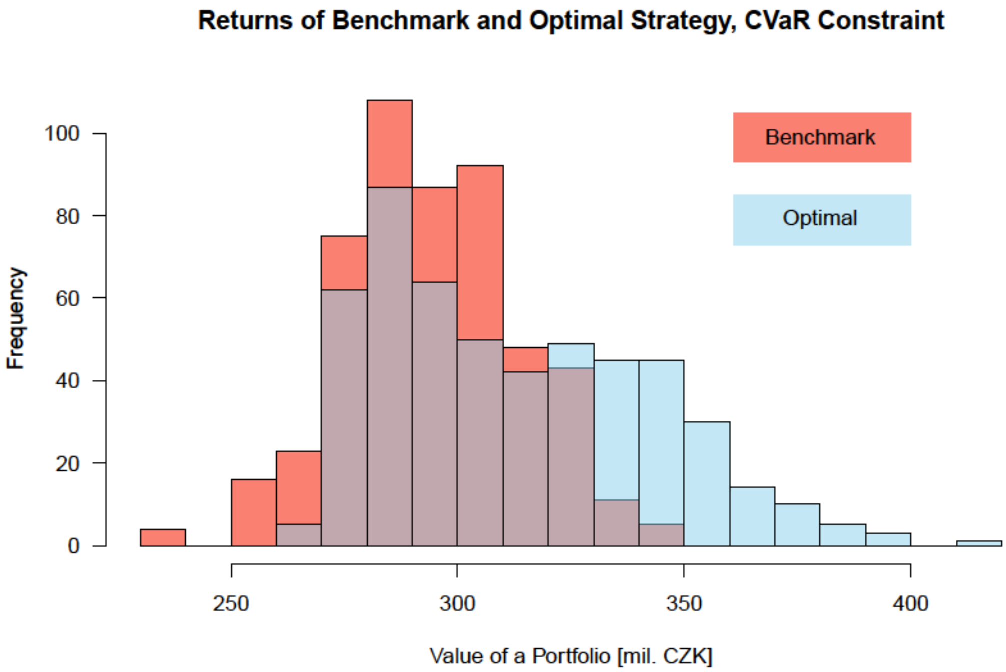

Figure 4.8: Histogram of values of a portfolio at the investment horizon according to the optimal solution of the CVaR constraint with a CVaR limit set to -275 mil. CZK.

program grew substantially and with the SSD constraint, it reached 269152 variables and 6099 equations. However, even with such a large scale model definition, it took in general only minutes to CPLEX to find the optimal solution.

To summarize the effect of the parameter b on the value of the optimal solution, we present Figure 4.9. The values of b we analysed were between $\langle 0, 15\rangle$, where $b = 15$ mil. CZK was the highest integer for which the problem was feasible. Similarly as with CVaR and VaR, all optimal strategies suggested to borrow only for one year in the first stage. An illustration of what the second–order stochastic dominance constraint causes and how parameter b affects the return distribution, we show Figure 4.10. There, one can see a comparison of histograms of the benchmark strategy and of the optimal strategy obtained as a solution of the program with $b = 12$. The SSD constraint says that every risk–averse investor would prefer the optimal strategy to the benchmark strategy even if we added to the benchmark strategy a sure bonus of $b = 12$ mil. CZK. Given that 12 is also the width of bins in the histogram, one can nicely see an effect of such a constraint on the return distribution of the optimal strategy.

To conclude this chapter, we have seen what were the largest returns which could be generated by the leasing company under different risk and second order stochastic dominance constraints. We saw that they bettered the benchmark strategy in all cases. This was a very surprising finding as we have tried to formulate the model in the most realistic way as possible. The amount by how much the benchmark was beaten by the optimal solution with the chance constraint where $\alpha = 0$ underlines the inefficiency of the benchmark strategy.

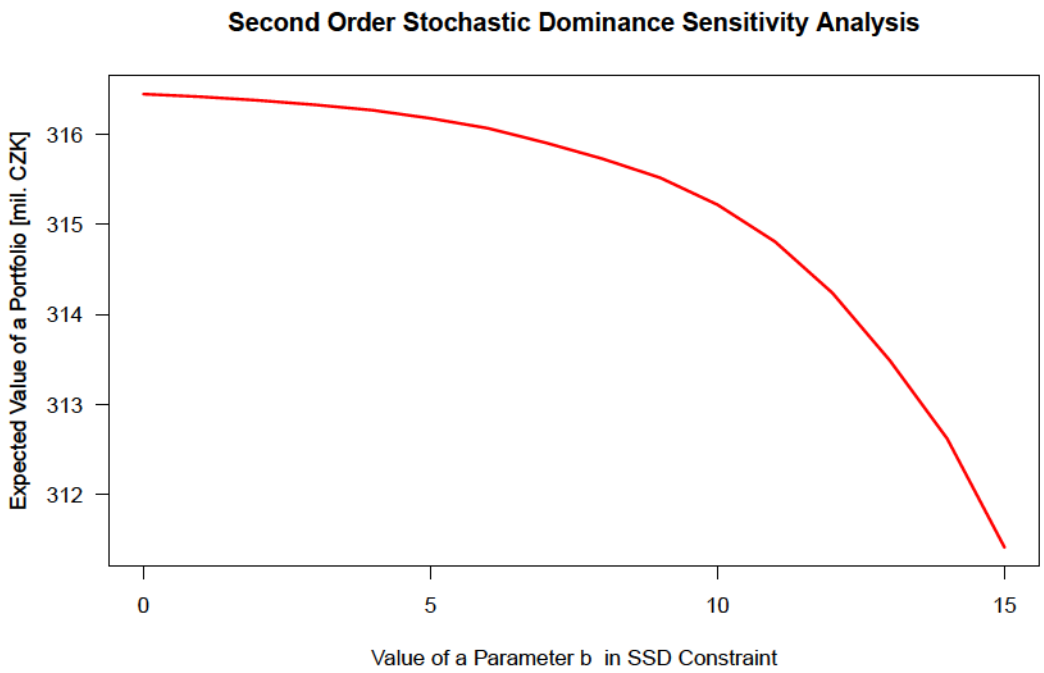

Figure 4.9: Dependence of the expected value of the optimal solution on the value of b in the second–order stochastic dominance constraint.

Finally, we would like to comment on why the suggested strategy was to close only one year loans in the initial stage in all VaR, CVaR and SSD constraints. We think it is a consequence of a balance between the riskiness of a position and its expected return. The stochastic program can determine what is the optimal balance and it makes decisions so this balance is reached. The fact, that (almost) all optimal strategies proposed to borrow only for one year in the first stage and hence to open the position as much as possible can be attributed to the fact that we start with no active deals. The suggestion to borrow only for one year in the initial stage must not be translated in the way that under every constraint the optimal strategy is to borrow only for one year. To support this statement, we present Figure 4.11, where the optimal borrowing in the third stage node (year two) with the highest interest rate is shown for the CVaR constraint. There, the program decides to close a combination of one and five year loans with the ratio highly dependent on the CVaR limit.

If we tried to implement this policy to an existing company, it would be required to consider current assets and liabilities of the company too. From there, we would obtain initial values of cash flows which would determine our current position and this would probably imply more varying decisions in the initial stage for different risk constraints and their parameters. The reason why we decided to model the company without any predetermined cash flows was the fact that we wanted to limit input parameters which do not have any real source. This would have been the case of initial cash flows, as no such data were available.

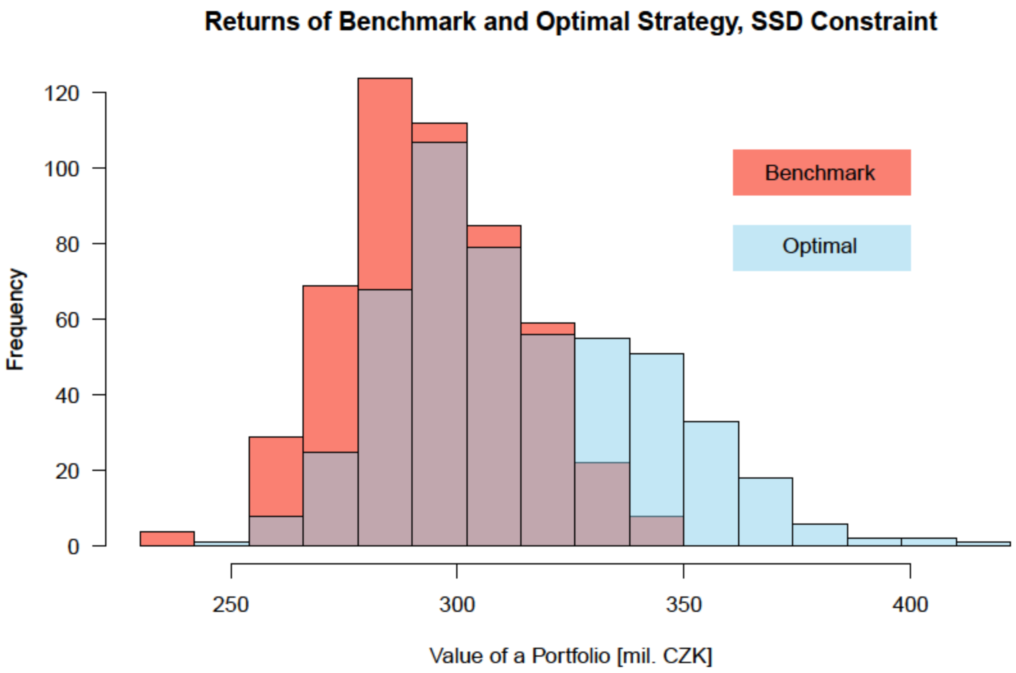

Figure 4.10: Histogram of values of a portfolio at the investment horizon according to the optimal solution of a SSD constraint which dominates the benchmark by $b = 12$ mil. CZK.

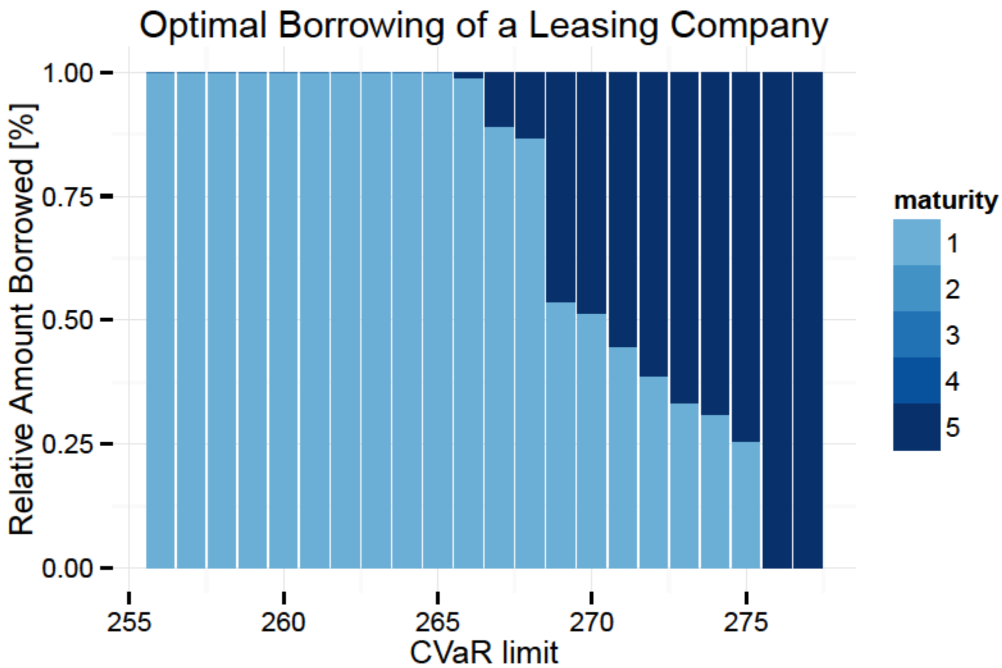

Figure 4.11: Suggested borrowings by the optimal solution for different value of v_α in the CVaR constraint in the third stage scenario with the highest interest rate.

5. Stress Testing of Optimal Solutions

The concept of stress testing is frequently used in the financial industry under various different meanings. The general idea behind it is however relatively clear, as we aim to find out what would happen to the object of interest in the case of unexpected shocks to the market environment. Some work on stress testing in the context of multi–stage stochastic programming has been done by Dupačová (2006b) and Dupačová and Kozmík (2015), who used a contamination technique to asses the effect of new scenarios on the optimal solution. However, the results of Dupačová (2006b) hold only for multi–stage stochastic programs where set of decisions is independent of the probability distribution, as shown in Dupačová (2006a). That is not the case of our asset–liability program. Our approach to the stress testing will be therefore bit different. In the first part of this chapter, we will describe the methodology we will use, while the second part will be devoted to the analysis itself.

5.1 Methodology of Stress Testing

Let us recall the main reason why financial institutions do stress testing. Their aim is to find out, if they are financially stable enough to overcome a deep financial crisis. In general, the goal of such testing is not to find an allocation of their assets and liabilities so they perform well in such situations. If one wanted to adjust his decisions so they take into account possible crisis scenarios, one would simply add them into the scenario tree. Therefore we will not manipulate with our original scenario tree which was implemented in the stochastic program analysed in Chapter 4. Instead, we will focus on answering the below–stated question.

Assuming that we adopt some optimal strategy as suggested by the multi–stage program, what would happen if a deep financial crisis came?

The events, which in reality happen, take place in the following order. At time zero, we decide to adopt some strategy and we borrow money according to it. Then a crisis comes. After we observe it, we can adjust our decisions to mitigate the negative effects of the crisis on the value of our portfolio. Finally, we ask what is the value of the portfolio at the time horizon. This heuristically described stress test can be employed in the multi–stage stochastic program in the following manner.

- Solve the asset–liability stochastic program for some risk parameters, obtain the optimal borrowing strategy in the first period, borrow money as proposed by the strategy at time zero.
- Then a crisis comes. We move to a crisis scenario at time one, where we can adjust our borrowing strategy.

- We generate a new scenario tree which grows from the crisis scenario node at time one with the same horizon of $n = 6$ years as the initial tree (so the new tree is one year shorter).
- We solve a new stochastic program with the new tree to find the best borrowing strategy in the crisis scenario. We learn what is the new expected value of our strategy in such a crisis scenario.

The question arises how to construct the crisis scenario. In our problem, we face two types of uncertainties. The first in the value of interest rates and the second in the amount of demand for loans. As demand affects the optimal and the benchmark strategy in the same way, the more interesting case is clearly a rapid change in interest rates. Decrease of interest rates should favour the optimal strategy, while increase would make it loose money, as one can see from Figure 4.2. On the other hand, the benchmark strategy should be relatively insensitive to changes in interest rates. Therefore our aim will be to investigate how the expected value of the optimal strategy changes with increasing interest rates. Banks in the Czech republic create crisis interest rate scenarios by adding (or subtracting) 2% to (from) the yield curve. In our problem, we will generate crisis scenarios by fixing the value of the short rate in the first year, with the most severe scenario corresponding to an increase in the short rate of 2%.

We also need to specify which optimal strategies we want to analyse. Because our motivation is to compare possible strategies with the benchmark strategy, their specification will be made with respect to this view. Hence the constraints will be set in the way that if there is a feasible solution, then the optimal solution of such a problem can be considered better than the benchmark strategy according to the corresponding risk measure or the dominance constraint. Especially the chance constraint with $\alpha = 0$ appears to us as very strong, but once $\alpha \neq 0$, it is impossible to claim which strategy is better, as both could be. Let us suppose that the management of the company chooses between pursuing the following goals:

1. A strategy which maximizes the expected value of the portfolio — we will call it the no–risk constraint strategy.
2. A strategy suggested by the chance constraint with $\alpha = 0$ — the chance constraint strategy.
3. A strategy which is proposed by the second–order stochastic dominance constraint with $b = 0$ — the SSD constraint strategy.
4. A strategy aiming to maximize the expected value while having VaR at $\alpha = 0.95$ lower than the benchmark strategy — the VaR constraint strategy.
5. A strategy with a goal to maximize the expected value under a condition of having CVaR at $\alpha = 0.95$ lower than the benchmark strategy — the CVaR constraint strategy.

The main question now is how these strategies would compare to the benchmark strategy under a crisis scenario of interest rate increase. To investigate this in more detail, we will consider a sequence of crisis scenarios each corresponding to a different level of interest rate increase.

5.2 Stress Test Analysis

In the following, we will proceed to the stress test as described in the previous section. First, we will explain the construction of crisis scenarios and how they compare to the scenarios in the original tree. The expected value of the short rate in the first year is 0.35%, while its highest value at that stage in the original tree is 1.1%. We will consider the short rate in the stressed scenarios to take values from the following sequence

$$\{0\%, 0.1\%, 0.2\%, \ldots, 2.4\%\}, \tag{5.1}$$

so the value of the short rate for the most severe scenario is more than 2% higher than its expected value, which corresponds to the usual shock applied in banks. We could add that from the model, the probability that the short rate is going to be greater than 2.4% is just 0.000234, which can be thought as a very extreme event. A severe scenario which might theoretically happen can be considered to be the 99% quantile of the short rate. That has a value of 1.7%. Given a scenario value of the short rate, the Hull – White model can be used to generate the subsequent tree in the same way as the original tree as described in Chapter 3. To illustrate how such a tree looks, we present Figure 5.1. We can see that after the initial stage, there is a rapid increase in the short rate, which represents our crisis scenario. Afterwards, we assume interest rates to behave

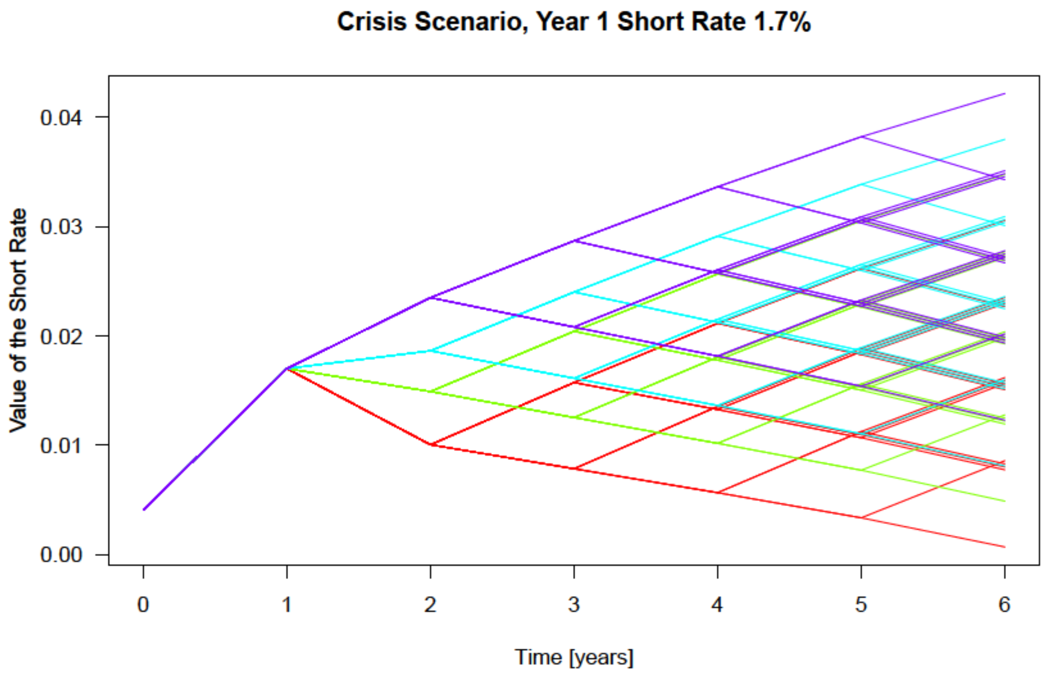

Figure 5.1: Scenario tree for stress test with the value of the short rate 1.7% in the first stage.

according to the model, which implies that the rest of the tree is similar to what we could see in Figure 3.2. The fact that one does not see much of a mean reversion is a consequence of increasing forward rates for latter stages and a relatively small estimate of the mean reversion parameter α in the Hull – White model.

Next, we will go through the steps we mentioned in the previous section, where we presented the stress test's methodology. We have learnt in Section 4.2 that the optimal investment strategies in the first stage were for the chance constraint with $\alpha = 0$ to borrow around a third in five year loans and the rest in one year loans, while all other optimal strategies corresponding to the constraints we listed above suggested to borrow everything in one year loans. Let us assume that the management of the company acted as given strategy proposed. Thereafter the crisis (which corresponds to the short rate having the value from the sequence (5.1)) comes. At stage two, the management realise their situation and they react to it by adjusting their decisions. We still assume their follow the strategy specified in the initial stage. They solve the corresponding optimization problem with a scenario tree as in Figure 5.1 and look at the result describing the expected value of the optimal portfolio at the time horizon. These are plotted for all strategies and for all considered values of the second stage short rate in Figure 5.2.

First, we would like to comment on the feasibility of strategies. Imagine that the board agree on the VaR constraint strategy, which means that they limit strategies to have VaR lower than the benchmark strategy. They borrow in the initial stage according to the optimal solution of such a model but after-

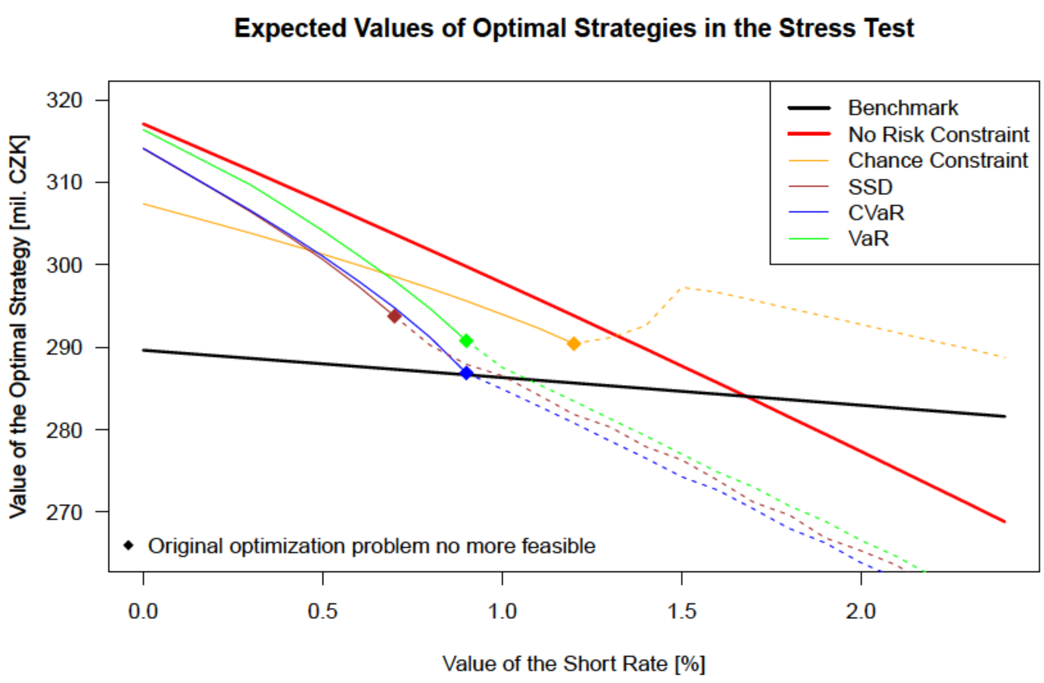

Figure 5.2: The expected value of the optimal solution of the asset–liability model for different level of stress test and various strategies. Dashed lines show optimal values of the programs after relaxing the (infeasible) risk constraints.

wards a crisis comes. If the crisis is deep enough, it can happen that there is no way how to adjust borrowings so their portfolio has VaR lower than the benchmark. That is down to the fact that the crisis causes the optimal strategy to loose too much money. In other words, for some (large) values of the short rate, the strategy which the management set, becomes infeasible. However, in such a situation, the management of the company still have to make a decision how to borrow money. We have decided that if such a situation happens, the management will employ the optimal strategy of a problem with the most strict risk constraint with some feasible solution. Hence in the case of VaR, if it happens that it is not possible to beat the benchmark VaR, we choose the VaR limit to be the lowest possible for which there is a feasible solution. Similar adjustment is made for all strategies. For example in the chance constraint strategy, when constraint $\alpha = 0$ becomes infeasible (which is a very strict constraint meaning we do not loose to the benchmark at every scenario), we set α the lowest possible but such that the problem is feasible. In Figure 5.2, we mark the point where we are forced to relax the initial assumption on the strategy of the management unit for every risk constraint by a small rhombus in the corresponding color. Expected values of the optimal solutions of such relaxed programs are then drawn by a dashed line.

Figure 5.2 summarizes how valuable are these five year loans in the case of unexpected interest rate movements. One can see that the optimal strategies which borrowed everything in one year loans suffer a lot once the value of the short rate exceeds 1.1%, which was the largest value of the short rate in the original scenario tree. We should not be surprised by this fact because these strategies were not trained for such situations. Given the decision on initial borrowings, one could feel that these strategies will be very sensitive to a large increase in interest rate. If we wanted these strategies to perform better in such scenarios, we would have to include them into the original scenario tree or to impose some additional constraints in the model formulation, which would ensure that borrowings would be more balanced. For example some condition on the duration of assets and liabilities could be useful for this purpose.

On the other hand, the chance constraint strategy looks to cope well with the crisis. We were surprised that the problem with the chance constraint was feasible even for the value of the short rate 1.2%, as this is greater than the largest value of the short rate in the original tree. To illustrate the distribution of returns in crisis scenarios, we present Figure 5.3 where the comparison of returns of the benchmark strategy and the chance constraint strategy is shown. The value of the short rate considered in this figure is 1.2%. We think it is fair to say that the optimal chance constraint strategy is far superior to the benchmark strategy. This strong statement is justified by Figure 5.2 where we learn that the implied strategy betters the return of the benchmark strategy even for extreme scenarios, which were not included in the original tree.

We would also like to comment on the large jump of the orange dashed line around the x-axis value of 1.5% in Figure 5.2, which the reader must have been wondering about. That is a consequence of relaxing the parameter α in the chance constraint. At the value of the short rate 1.5%, α rapidly increased from 6.3% to 22%, which led to an increase in the maximal expected value. Summing this up, for increasing value of the short rate, we must have increased α to make the problem feasible. This implies that the benchmark improves its performance relative to the optimal chance constraint strategy in the worst scenarios. This indicates that the left tail of the optimal chance constraint strategy becomes heavier than the left tail of the benchmark strategy. Therefore, the chance constraint strategy could be considered more risky than the benchmark. In exchange, we obtain an increase in the expected value of the chance constraint strategy, which to some extent compensates for the increase in its riskiness.

In this chapter, we have seen how one can approach the so called post–optimality analysis, which studies properties of optimal solutions. This can provide us with valuable information about the suitability of the model formulation. The stress test we adopted answers questions which one could expect to be asked by someone who could employ a strategy proposed by the model. Given the results of this analysis, the manager would either get more belief in the optimal solution or he would identify properties where he wants the solution to improve. Consequently, we could reformulate the problem to achieve desired results.

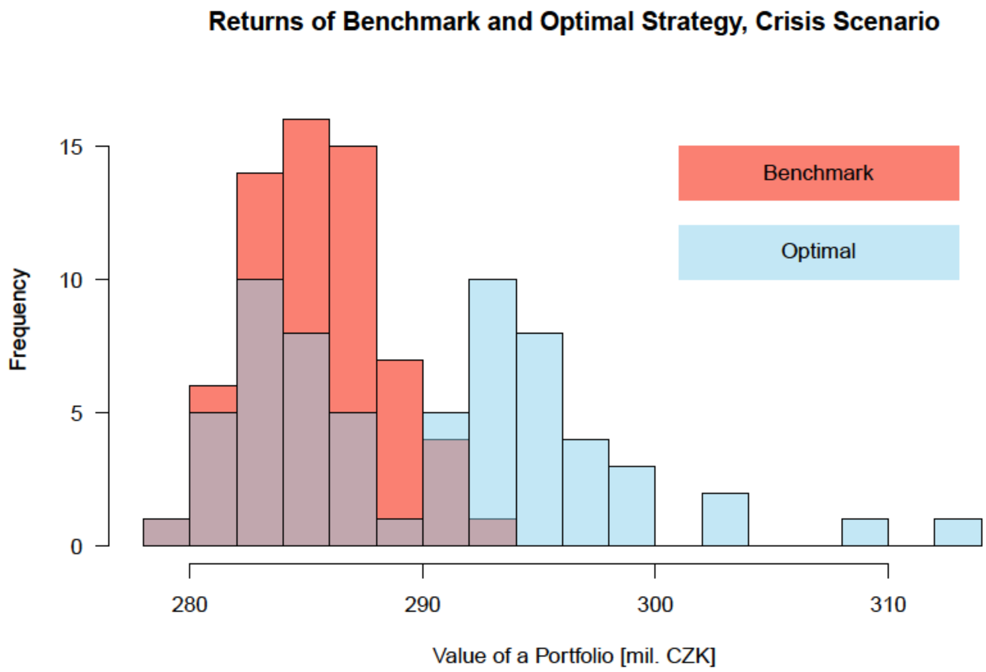

Figure 5.3: Histogram of values of portfolios of the benchmark strategy and the optimal chance constraint strategy given the value of the short rate to be 1.2% in the second stage.

Conclusion

The ultimate goal of this thesis was to show an application of stochastic programming to asset–liability management. We formulated a multi–stage stochastic program inspired by a business model of a leasing company, which as far as we know was not yet subject to modelling by means of stochastic programming. We focused on showing how optimal decisions obtained as a solution of the optimisation problem can improve the company's performance. The core of the model was formed by the dynamics of cash–flows of the company. Still, the risk constraints, which were employed to reflect the risk aversion people usually posses, should be considered at least as equally important. They give to the optimal solution desirable properties and a strong interpretation.

We paid a special attention to creating as most realistic model as possible. It was an absolute necessity to use market data to obtain input parameters of the model. Moreover, for scenario generation of interest rates, we developed a new calibration procedure for one–factor short–rate models. Our approach uses more information available at the market than any other well–known method. It is based on the maximum likelihood, which however before did not produce satisfactory estimates of the model's parameters. The generalization of the construction of the likelihood we introduced corrected the maladies which were previously associated with this method. Moreover, we derived analytical expressions for the changes our generalized approach implied to the likelihood function so the adjustments do not complicate the numerically already difficult optimisation of the likelihood. The choice of the Hull – White model for the scenario generation was driven by the fact that the implied rates fit precisely the currently observed rates, and hence we think the model gives the most realistic predictions between all affine–term structure short rate models.

Consequently, we believe that the results we have reached in Chapters 4 and 5 were obtained in a competitive environment, which did not favour to any of the considered strategies. Most notably, we found in Chapter 4 that the here–and–now solution to borrow money only for one year generates much higher return than the benchmark strategy; and moreover under suitable behaviour in latter stages it is possible to achieve much better risk parameters of the portfolio. The inefficiency of the benchmark is underlined by both the amount by how much it is beaten by the optimal strategy obtained from the stochastic program with the chance constraint with $\alpha = 0$ and by the maximum value of the parameter $b = 15$ for which we found an SSD b–dominating portfolio. The solution of the chance–constraint program says that we can make on average 14.4 mil. CZK greater value of a portfolio after six years, while being sure that we are not going to do worse than we would do if we kept our original strategy. On the other hand, finding an SSD 15–dominating portfolio implies that we are able to propose a strategy which would be preferred to the benchmark strategy by every risk–averse non–satiated investor even if he would be obliged to pay 15 mil. CZK at the investment horizon for following such a strategy. The facts we found in Chapter 4, that all optimal strategies suggested by the stochastic program for various risk constraints are far superior to the benchmark strategy capture the applicability of stochastic programming in the financial sector.

If any institution ever considers implementing such a method in practise, we think they would also ask what performance they can expect from their strategy if things do not go as they planned. The stress test we have introduced in Chapter 5 answers such a question. Most notably, the Figure 5.2 provides a clear graphical answer on the dependence of the optimal solution's expected value and the year–one interest rate. Such a figure can be of a great importance when choosing between optimal strategies or when assessing the suitability and the strictness of risk constraints. For example after seeing how the optimal strategies which borrow only one year loans in the initial stage perform in the crisis scenario with the value of the short rate 1.7%, we might wish to reformulate the model so our strategy is more robust to such fluctuations. This could lead us to introduce a condition to the model which would force the considered strategy not to do worse than the benchmark in such a crisis scenario. It is worth to note that the methodology we have applied in the stress test have not been used before in any application of stochastic programming within asset–liability management.

There are many ways how to extend the model we presented in this thesis. The power of stochastic programming lies in its ability to model various needs of managers and to take into account different regulations and specific market conditions. We could for example consider multiple criteria in the objective function, as we might wish to consider short–term profit too. Extensions of risk constraints can be also added because managers might require our strategy to meet some limits in the course of the time, so they do not loose belief in the model. Another interesting way how to improve the program might be to add a possibility to invest into market instruments, like interest rate caps and floors. These could be used to hedge against a possible increase in interest rate and consequently to improve the company's performance. However, we think that implementation of such adjustments should be driven mainly by a manager's wish, as there is no general agreement which of them should be the most preferred. Anyway, the idea how this would be done is rather similar to what we have presented in this thesis and consequently we decided to leave this work to a curious reader or to an ambitious business analyst.

Bibliography

Artzner, P., Delbaen, F., Eber, J. M., and Heath, D. (1999). Coherent Measures of Risk. *Mathematical Finance*, 9(3):203–228.

Birge, J. and Louveaux, F. (1997). *Introduction to Stochastic Programming.* Springer Verlag, New York.

Brigo, D. and Mercurio, F. (2001). *Interest Rate Models - Theory and Practice with Smile, Inflation and Credit (2nd ed. 2006).* Springer Verlag, Berlin.

Carino, D. R., Kent, T., Myers, D. H., Stacy, C., Sylvanus, M., Turner, A. L., Watanabe, K., and Ziemba, W. T. (1994). The Russell-Yasuda Kasai Model: An Asset/Liability Model for a Japanese Insurance Company Using Multistage Stochastic Programming. *Interfaces*, 24(1):29–49.

Chen, R. and Scott, L. (1993). Maximum Likelihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates. *The Journal of Fixed Income*, 3:14–31.

Cox, J., Ingersoll, J. E., and Ross, S. A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, 53:385–407.

de Oliveira, A., Filomena, T., Perlin, M., Lejeune, M., and de Macedo, G. (accepted 2017). A Multistage Stochastic Programming Asset-Liability Management Model: An Application to the Brazilian Pension Fund Industry. *Optimization and Engineering*, pages 1–20.

Dert, C. (1995). *Asset Liability Management for Pension Funds: A Multistage Chance Constrained Programming Approach.* PhD thesis, Erasmus University, Rotterdam.

Dupačová, J. (2006a). Contamination for Multistage Stochastic Programs. In Huskova, M. and Janzura, M., editors, *Proc. of Prague Stochastics 2006*, pages 91–101. Matfyzpress.

Dupačová, J. (2006b). Stress Testing via Contamination. In Marti, K., Ermoliev, Y., Makowski, M., and Pflug, G., editors, *Coping with Uncertainty: Modeling and Policy Issues*, pages 29–46. Springer, Berlin.

Dupačová, J., Growe-Kuska, N., and Römisch, W. (2003). Scenario Reduction in Stochastic Programming. *Mathematical Programming*, 95(3):493–511.

Dupačová, J. and Kozmík, V. (2015). Stress Testing For Risk-Averse Stochastic Programs. *Acta Math. Univ. Comenianae*, LXXXIV(2):205–217.

Dupačová, J. and Polívka, J. (2009). Asset-Liability Management for Czech Pension Funds Using Stochastic Programming. *Annals of Operations Research*, 165:5–28.

Hadar, J. and Russell, W. (1969). Rules for Ordering Uncertain Prospects. *American Economic Review*, 59:25–34.

Hardy, G., Littlewood, J. E., and Polya, G. (1934). *Inequalities.* Cambridge University Press, Cambridge.

Harrison, J. and Pliska, S. (1981). Martingales and Stochastic Integrals in the Theory of Continuous Trading. *Stochastic Processes and their Applications*, 11:215–260.

Heitsch, H. and Römisch, W. (2003). Scenario Reduction Algorithms in Stochastic Programming. *Computational Optimization and Applications*, 24:187–206.

Hoyland, K. (1998). *Asset Liability Management for a Life Insurance Company: A Stochastic Programming Approach.* PhD thesis, Norwegian University of Science and Technology.

Hull, J. (2008). *Options, Futures, and Other Derivatives (7th ed. 2008).* Pearson Prentice Hall, Englewood Cliffs.

Hull, J. and White, A. (1990). Pricing Interest-Rate Derivative Securities. *The Review of Financial Studies*, 3:573–592.

J.P. Morgan Risk Metrics (1995). Technical Document, 4th Edition. Morgan Guaranty Trust Company, New York.

Klein Haneveld, W. K., Streutker, M. H., and van der Vlerk, M. H. (2010). An ALM Model for Pension Funds Using Integrated Chance Constraints. *Annals of Operations Research*, 177(1):47–62.

Kopa, M. and Chovanec, P. (2008). A Second-Order Stochastic Dominance Portfolio Efficiency Measure. *Kybernetika*, 44:243–258.

Kuosmanen, T. (2004). Efficient Diversification According to Stochastic Dominance Criteria. *Management Science*, 50(10):1390–1406.

Kusy, M. I. and Ziemba, W. (1986). A Bank Asset and Liability Management Model. *Operations Research*, 34(3):356–376.

Levy, H. (1992). Stochastic Dominacne and Expected Utility: Survey and Analysis. *Management Science*, 38(4):555–593.

Ogryczak, W. and Ruszczynski, A. (2002). Dual Stochastic Dominance and Related Mean-Risk Models. *SIAM Journal of Optimization*, 13:60–78.

Pflug, G. C. (1999). Some remarks on the value-at-risk and the conditional value-at-risk. In Uryasev, S. P., editor, *Probabilistic Constrained Optimization: Methodology and Applications*. Kluwer.

Rockafellar, R. T. and Uryasev, S. (2000). Optimization of Conditional Value-at-Risk. *Journal of Risk*, 2:21–41.

Rockafellar, R. T. and Uryasev, S. (2002). Conditional Value-at-Risk for General Loss Distributions. *Journal of Banking and Finance*, 26(7):1443–1471.

Telser, L. (1955). Safety First and Hedging. *The Review of Economic Studies*, 23(1):1–16.

Varadhan, R., University, J. H., Borchers, H. W., and Research., A. C. (2011). *dfoptim: Derivative-free Optimization*. R package version 2011.8-1.

Vasicek, O. (1977). An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics*, 5:177–188.

List of Figures

1.1 A scenario tree of the private investor problem. 5

3.1 Monthly observations of the yield curve derived from the swap rates on PRIBOR from January 2009 to November 2016. 36
3.2 Scenario values of one year interest rate in the Czech market in a tree structure used in the optimisation problem. 37
3.3 Yield curves of the first–stage nodes exhibit increasing rate with maturity. 38
3.4 The density of demand for loans for two different values of interest rate. 40

4.1 Comparison of returns of the benchmark and the optimal strategy in the no–risk constraint problem. 45
4.2 Differences in portfolio values plotted against one–year interest rate in the final stage. 46
4.3 Dependence of the expected value of the optimal solution on the value of α in the chance constraint. 47
4.4 Suggested borrowings by the optimal solution with different value of α in the chance constraint. 48
4.5 Dependence of the expected value of the optimal solution on the value of u_α in the Value–at–Risk constraint. 49
4.6 Histogram of portfolio values at the investment horizon according to the optimal solution of a VaR constraint with the limit of $u_\alpha = -287$ mil. CZK. 50
4.7 Dependence of the expected value of the optimal solution on the value of v_α in the conditional Value–at–Risk constraint. 51
4.8 Histogram of values of a portfolio at the investment horizon according to the optimal solution of the CVaR constraint with a CVaR limit set to -275 mil. CZK. 52
4.9 Dependence of the expected value of the optimal solution on the value of b in the second–order stochastic dominance constraint. 53
4.10 Histogram of values of a portfolio at the investment horizon according to the optimal solution of a SSD constraint which dominates the benchmark by $b = 12$ mil. CZK. 54
4.11 Suggested borrowings by the optimal solution for different value of v_α in the CVaR constraint in the third stage scenario with the highest interest rate. 55

5.1 Scenario tree for stress test with the value of the short rate 1.7% in the first stage. 58
5.2 The expected value of the optimal solution of the asset–liability model for different level of stress test and various strategies. Dashed lines show optimal values of the programs after relaxing the (infeasible) risk constraints. 59

5.3 Histogram of values of portfolios of the benchmark strategy and the optimal chance constraint strategy given the value of the short rate to be 1.2% in the second stage. 61

List of Tables

3.1 Volumes of leasing loans closed in a given year with maturity m in millions of CZK and corresponding 1 year PRIBOR, source: ARAD database of CNB. 39