FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

DOCTORAL THESIS

RNDr. Tomáš Rusý

Asset-Liability Management: Application of Stochastic Programming with Endogenous Randomness and Contamination

Department of Probability and Mathematical Statistics

Supervisor of the doctoral thesis:	doc. RNDr. Ing. Miloš Kopa, PhD.
Study programme:	Probability and Statistics, Econometrics and Financial Mathematics
Study branch:	Mathematics

Prague 2021

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources. It has not been used to obtain another or the same degree.

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 .... date ....

Author's signature

At this place, I would like to thank a few people and institutions who significantly contributed to the development of my doctoral 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 models and to reach desired results.

Second, a big thanks goes to Dr. Kamil Kladívko, of University of Örebro, who cooperated with me on the development of the calibration method of the Hull-White model and provided me with various interesting points end remarks, especially from the practical point of view.

I would also like to acknowledge the grant agencies GAČR and GAUK who through grants P402/12/G097, 18-05631S, 19-28231X and 258318 contributed funding for the research project.

Last, but not least I would like to thank my family for supporting me through my PhD studies at Charles University.

Thank you.

Title: Asset-Liability Management: Application of Stochastic Programming with Endogenous Randomness and Contamination

Author: RNDr. Tomáš Rusý

Department: Department of Probability and Mathematical Statistics

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

Abstract:

This thesis discusses a stochastic programming asset-liability management model that deals with decision-dependent randomness and a subsequent contamination analysis. The main model focuses on a pricing problem and the connected asset-liability management problem describing the typical life of a consumer loan. The endogeneity stems from the possibility of their customer rejecting the loan, the possibility of the customer defaulting on the loan and the possibility of prepayment which are all affected by the company's decision on interest rate of the loan. Another important factor, which plays a major role for liabilities, is the price of money in the market. There, we focus on the scenario generation procedure and develop a new calibration method for estimating the Hull-White model [Hull and White, 1990a] under the real-world measure. We define the method for the general class of one-factor short-rate models and perform an extensive analysis to assess the estimation performance and properties. Further, we extend the contamination approach of Dupačová [1986, 1996] to models with decision-dependent randomness. This gives us a tool for investigating stability of stochastic programs with decision-dependent randomness with respect to changes in the underlying probability distributions. That represents an important step before deploying any model to production. In this thesis, we first extend the current results by developing a tighter lower bound applicable to wider range of problems. Thereafter, we define contamination for decision-dependent randomness stochastic programs and prove various lower and upper bounds. We split the cases into two separate sub-classes based on whether the feasibility set is fixed or probability-distribution-dependent and discuss several tractable formulations. The method is illustrated on the aforementioned example of the consumer loan stochastic program as well as on its extended version with implemented risk constraint.

Keywords: maximum likelihood estimation of short-rate models, stochastic programming, decision-dependent randomness, contamination

Contents

Introduction

Stochastic programming is a well established area of mathematical optimization, whose main results have been published in multiple manuscripts. Between the most important, we should include Birge and Louveaux [1997], Ruszczyński and Shapiro [2003] and Kall and Mayer [2005], and further Shapiro et al. [2014] and Pflug and Pichler [2016]. Its main objective is to provide a framework to solve optimization problems under uncertainty. Such problems are tackled in various fields, among these, we should mention transportation and logistics, medicine, biology and also finance, especially the area of financial planning and control. The first application of stochastic programming in finance occurred in 1986 when Kusy and Ziemba published their famous paper dealing with a bank asset-liability model [Kusy and Ziemba, 1986]. Two years later, Dempster and Ireland [1988] introduced a different model that focused on the immunisation of liabilities and for the first time took into account the risks that are associated with financial problems. Carino et al. [1994] presented the very first asset-liability model for an insurance company, which was later followed by many others. There have also been several applications focused on insurance companies [Hoyland, 1998, Pliska and Ye, 2007, Broeders et al., 2009] and pension funds [Dert, 1995, Consigli and Dempster, 1998, Geyer and Ziemba, 2008, Dupačová and Polívka, 2009]; these are the two types of financial institutions where stochastic programming was used most often.

These works, however, do not exhaust the potential of applications for stochastic programming in finance, and especially in asset–liability management. For example, Consigli [2008] investigated the problem of an individual investor who needed to undertake investment decisions and manage his consumption. Kopa et al. [2018] analysed the effect of new, modern risk constraints on the optimal solution as they applied a second-order stochastic dominance constraint. An important aspect of the modelling exercise is to focus on the accuracy of model formulation. Vitali et al. [2017] and Moriggia et al. [2019] paid a lot of attention to this aspect, considering a variety of investment possibilities. Recently, Consigli et al. [2020] studied optimal decisions from the household point of view and Zapletal et al. [2020] investigated optimal policies in emission management of a steel company.

So far, in all the stochastic programming models of asset-liability management that we have seen, authors look for optimal decisions of a single agent (insurance company, pension fund, etc) in the market. Such market is always considered large enough to justify assuming independence of evolving market prices from agent's strategies, and thus risk sources are treated as exogenous. In our work, we focus on a stochastic programming formulation of an asset-liability management problem which includes a pricing decision. Such a decision is a natural part of the business of every company, as companies need to price their products. Yet, given that it brings significant complications to the model solution, it has not, according to our best knowledge, been tackled in other academic works.

In financial problems, this decision-dependent uncertainty is typically observed within a bilateral relationship. One party acts as a price setter and its pricing decision affects the demand for some good of the second party. In mathematical terms, this can be translated to observing a change in the underlying probability distribution of the demand (endogenous randomness). Apart from demand distribution, probabilities of some subsequent events such as default and prepayment can also be affected. Other decisions, such as marketing decisions or branding investments, also indirectly affect the company's books and lead to decision-dependent uncertainty.

The issue of endogenous uncertainty was first addressed by Pflug [1990] who investigated a general Markovian process in which states depended on the decision parameter of the model. In the following years, more models with decision-dependent randomness have emerged; these can be generally divided into two types. In the first type, decisions taken in the program help to determine uncertainty in the model. This case was tackled, for example, by Jonsbraten et al. [1998], Goel and Grossmann [2004] and Tarhan et al. [2009] who all described models for finding the optimal strategy in an offshore oil field development plan. The general idea was that companies could decide to run exploratory analyses on potential oil fields, obtain a better estimate of how much oil is available and determine where it is more profitable to set up a plant. In the second type, as is the case of our problem, decisions affect scenario probabilities. This type of endogenous uncertainty was thoroughly investigated, for example, in network flows problems [Ahmed, 2000, Held and Woodruff, 2005, Vishwanath et al., 2004]. Under this setting, the probabilities of scenarios depend directly on decisions taken and they can be also considered as variables in the program.

In one part of the thesis, we formulate a stochastic programming model describing a consumer loan that a company gives to its customer. This acknowledges a relationship between the two parties. In this situation, the customer's behaviour is affected by the company's actions which induces the decision-dependent relationship. Studying such an financial application of stochastic programming is one of the contributions of this thesis.

Stochastic programming models aim to find solutions of optimization problems which involve uncertainty. For this purpose, one usually uses the "best estimate" of the stochastic distribution depicting the uncertainty. However, people, for example from the financial industry, quickly started to ask what would happen if the underlying distribution was different. These questions were mainly driven by observed crises regularly challenging the financial industry. For this reason, investigating the stability and robustness of the stochastic programs became an important part of the modelling exercise. Consequently, various methods were proposed in this direction, for example, the min-max approach of Žáčková [1966] and methods based on the local stability results for non-linear parametric programming nicely summarized in works of Robinson [1987], Bonnans and Shapiro [1998], Römisch [2003] or more recently in Shapiro et al. [2014].

The min-max approach relies on a definition of the ambiguity set. This set defines the stochastic distributions considered in the distributionally robust program. It can be constructed via either moment matching [Wagner, 2008, Zymler et al., 2013, Zhang et al., 2018] or some statistical distance such as $\phi-$divergence [Jiang and Guan, 2016], Wasserstein metric [Esfahani and Kuhn, 2018, Blanchet and Murthy, 2019] or norm-based distances [Jiang and Guan, 2018]. Robustness of decision-dependent randomness stochastic programs via the min-max approach has been recently discussed in several papers. Theoretical foundations via variational theory were presented by Royset and Wets [2017], while Zhang et al. [2016] show Holder continuity properties for a specific class of distributionally robust problems. Furthermore, Noyan et al. [2018] developed a single-stage problem where the ambiguity set is determined by Wasserstein metric. Yu and Shen [2020] studied a multi-stage problem, where they opted to use first and second moment matching method to define the ambiguity set. A similar approach was also used by Basciftci et al. [2021] in their study of a distributionally robust facility location problem. Other possibilities on how to construct the ambiguity set under the decision-dependent setting are investigated by Luo and Mehrotra [2020] who formulated five types of ambiguity sets defined by simple measure and moment inequalities, bounds on moment constraints, Wasserstein metric, $\phi-$divergence and based on Kolmogorov-Smirnov test.

In the final part of the thesis, we focus on the contamination approach of Dupačová [1986, 1996], which combines stochastic and parametric programming. It investigates the behaviour of the optimal value function of a stochastic program, when we move from the initial distribution P to a contaminating, or stress-testing distribution Q. Roughly speaking, we define a contamination parameter $t \in [0, 1]$ and then distribution $P_t = (1 - t)P + tQ$. For each t, we also define a stochastic program by substituting for the distribution P the distribution P_t in the original program. Then, we investigate how the optimal value changes with t. We aim to construct lower and upper bounds for the optimal value function (as a function of t). This gives us an approximation of the optimal objective value for any level of contamination without a need to solve it for each t. Such bounds have been constructed for a variety of stochastic programs. Dupačová [1996] developed them in the case of a stochastic program with a fixed set of feasible decisions. These were applied for example in Dupačová and Polívka [2007] and Dupačová and Kozmík [2015a,b]. In the latter work, there was also an extension to multi-stage and risk-averse stochastic programs. Dupačová and Kopa [2012] further derived bounds for the case of probability-dependent set of feasible decisions and applied them to test the robustness of mean-CVaR optimization model and stochastic programs with second-order stochastic dominance constraints. Dupačová and Kopa [2014] then extended the work to application on programs with first-order stochastic dominance constraints.

Another recent work discussing contamination or stress-testing is that of Moriggia et al. [2019], who invented nodal contamination. In their numerical study, they modify values in scenarios (merge two same-size scenario trees). The advantage of their approach is that it neither increases computational complexity of the original problem nor breaks equiprobability of scenarios, which might be required in some applications. Another technique, designed especially for stress-testing, was described in Rusý and Kopa [2018]. They investigate the robustness of the solution in their multi-stage program by fixing the here-and-now decisions, obtained as a solution of the stochastic program with probability distribution P, and then, modify scenario values in the subsequent stages. With such an approach, they investigate what would happen if the here-and-now solution was applied while the realization of the random elements in the first stage was not considered in the original scenario tree characterized by the distribution P.

In this thesis, we first extend results for contamination in exogenous randomness stochastic programs by proving tighter lower bound applicable to a wider range of problems. Then, we define contamination also for the endogenous uncertainty or decision-dependent randomness stochastic programs. To our best knowledge, this was previously only shortly mentioned by Dupačová [2006], where a concise definition and discussion is provided. Yet, situations where decisions affect random elements are frequently seen in the financial industry, but due to their complicated nature, not much is known about their properties. In general, any pricing, marketing or advertising decision has a direct effect on the demand for some good of a company and hence it needs to be considered in problems of welfare maximization. Decision-dependent randomness problems represent an important class of problems tackled by most financial companies.

However, before we move to the formulation of a decision-dependent asset-liability management stochastic program and contamination of stochastic programs, we first develop a new calibration method for estimating one-factor short-rate models. This is a prerequisite for applications of stochastic programming in finance, as we work with money and interest rate represents the cost of capital so it needs to be taken into account. These costs are not deterministic and hence forecasting models need to be used to create scenarios. The class of one-factor short-rate models is then frequently used for such tasks.

The main idea of our approach lies in introducing postcalibration errors and maximising the likelihood of model implied yields in periods subsequent to the calibration time. The short rate, which we consider as the sole yield curve factor, is modelled as a latent process. We infer the short rate, jointly with postcalibration errors, by inverting the so-called postcalibration equation. By assuming normality for the postcalibration errors we are able to formulate a profile log-likelihood function, which is easy to maximise by using standard numerical algorithms.

Classical time-homogeneous one-factor short-rate models, such as Vašíček model [Vasicek, 1977] or Cox–Ingersoll–Ross (CIR) model [Cox et al., 1985b] proved to be useful for capturing yield curve dynamics, but their use in pricing is limited. The reason is that time-homogeneous models do not match market yield curve data perfectly. This drawback has been overcome by time-inhomogeneous models pioneered by Ho and Lee [1986] and Hull and White [1990b]. Time-homogeneity is implied by assuming constant model parameters and means that the transition probability of the short rate and thus the whole yield curve is independent of time. A time-inhomogeneous short-rate model is a relaxation of a time-homogeneous model from constant to time-varying parameters.[1]

Hull and White [1990b] suggest to replace a constant unconditional mean of the one-factor Vašíček model by a function of time and market forward rates. This modification is known as the Hull-White model. Thanks to its tractability and ability to fit yield curve data exactly, the model has become popular among practitioners. This is the reason why we select this model for numerical investigation and compare the performance of our method to other calibration and

[1] [Brigo and Mercurio, 2001, Chapter 3] call time-inhomogeneous models "exogenous" as the market yield curve enters the model in the form of time-varying parameter. Time-homogeneous models are then called "endogenous" as the model yield curve may differ from the market yield curve. [Hull, 2008, Chapter 30] denotes time-homogeneous models as "equilibrium", whereas time-inhomogeneous models as "no-arbitrage". The rationale for this classification is that the CIR model is a byproduct of a general equilibrium complete economy model developed in Cox et al. [1985a], and "no-arbitrage" reflects the perfect fit of any market yield curve by a time-inhomogeneous model.

estimation approaches.

We use a time series of yield curves and estimate model parameters under both risk-neutral and real-world probability measures. Our estimation approach thus overcomes drawbacks that may arise when calibration methods are used. The standard practice in identification of no-arbitrage models is to calibrate the model to market prices of vanilla interest rate derivatives, such as swaptions (see Section 1.4.7). However, calibration methods are unsuitable in the following two cases.

First, for many currencies, market prices of interest rate derivatives are not available, and straightforward use of calibration methods is therefore not possible [Vojtek, 2004, Witzany, 2010]. Second, calibration methods are based on a snapshot of market prices prevailing at a particular point in time, and thus provide parameter values under the risk-neutral measure only. Since the market price of interest rate risk is absent in the risk-neutral measure, the real-world interest rate trajectory can be substantially different from its risk-neutral path, see Hull et al. [2014] for a detailed discussion. Therefore, risk assessment and portfolio management requires that the model is identified under both risk-neutral and real-world probability measures. See Vedani et al. [2017] for pitfalls in life insurance, Jain et al. [2019] for issues in credit risk, or Hull and White [2018] for a general discussion.[2] Fergusson [2017] studies the use of the Hull-White model for actuarial valuation under the real-world measure, and derives the model implied long-term yield.

Our likelihood approach builds upon estimation methods developed for time-homogeneous short-rate models. These methods have received a fair amount of attention. A strand of literature focuses on maximum likelihood estimation of diffusion processes that serve as an observable proxy of the short rate. Fergusson and Platen [2015] provide a thorough theoretical and empirical investigation of maximum likelihood estimation to the short-rate process used in the Vašíček, CIR and 3/2 model. Fergusson [2020] considers maximum likelihood estimation of square root (CIR) and Bessel processes, while Aït-Sahalia [2002] develops a maximum likelihood approximation, based on Hermite polynomials, to a general diffusion process.

The maximum likelihood method is related to a strand of literature that considers estimation of arbitrage free yield curve models. In this case, the model parameters are estimated from a time series of yield curves, and the short rate is typically extracted as a latent process. Chen and Scott [1993] propose a maximum likelihood estimator for the CIR class of models, whereas Ang and Piazzesi [2003] use a maximum likelihood approach to estimate a Gaussian macro-finance model. As an alternative to the likelihood methods Hamilton and Wu [2012] suggest a minimum-chi-square estimator. However, the existing papers focus dominantly on time-homogeneous models. In fact, we are only aware of Harms et al. [2016, 2018], who develop a recursive recalibration technique for time-inhomogeneous models. Their consistent re-calibration approach (CRC) introduces a new class of Heath-Jarrow-Morton models called CRC models. The distinguishing feature of a CRC model is that all of the model parameters are time-dependent or stochastic.

[2] To address risk-management applications, Hull and White [2018] incorporate a real-world measure scenario analysis into their tree procedures. However, an estimate of the market price of risk still needs to be provided externally.

Harms et al. [2018], who mainly focus on the CRC theory, implement a CRC version of the extended Vašíček model by using realised covariance estimators, and 3-month yield as the short-rate proxy. Harms et al. [2016] considers a three-factor discrete-time CRC extended Vašíček model, and uses a combination of Kalman filtering, CRC recursion and realised covariance estimator to infer the model dynamics.

In academic literature, a method for calibrating the popular Hull-White model under both measures, so it can be used for derivatives pricing and yield curve forecasting, has been missing. We develop such a method which can be moreover applied to a large class of one-factor short-rate models. We see this as another contribution of this thesis.

This thesis is organised as follows. In the first chapter, we describe the new calibration method for the one-factor short-rate models. First, in Section 1.2, we present the methodological core of the method and define the maximum likelihood estimation approach. Further, in Section 1.3 the Hull-White model is formulated under both measures. In Section 1.4 empirical performance of our estimation approach is investigated on EUR interest rate data, including in-sample and out-of-sample performance of the estimated model, and comparison with calibration methods. Results mentioned in this chapter have been submitted for publication in Kladívko and Rusý [2021].

Further, in Chapter 2, we develop a program describing the life-cycle of a loan which a company provides to an individual customer, which we believe is the first application of stochastic programming with endogenous randomness in finance. We present the model formulation in detail in Section 2.1. Among other things, we also discuss what decisions can be taken and how they affect the uncertainties the company faces. We describe scenarios, their implementation into the program and all constraints which are part of the model formulation. Thereafter, in Section 2.2, we show the results of this program for one parameter's setting to illustrate the optimal decisions and how the loan value would evolve depending on the company's decisions. Furthermore, we discuss the effect of the customer's properties on the model solution. We also mention the losses which are incurred by the company if it does not behave in the optimal way. The model has been published in Kopa and Rusý [2021b].

Finally, in the last Chapter 3 we discuss the contamination approach for analysing robustness of stochastic programs. In Section 3.1, we review current works on contamination for the programs with exogenous randomness. There, we also show that the lower bound developed by Dupačová and Kopa [2012] for contamination in cases with probability-dependent set of feasible decisions can be improved. Next in Section 3.2, we define the meaning of contamination for decision-dependent randomness stochastic programs. We develop lower and upper bounds and show that these can be applied in certain cases with endogenous randomness. Thereafter, in Section 3.3, we apply the developed bounds on the asset-liability management stochastic program introduced in Chapter 2. Moreover, we extend the program for a CVaR constraint to show an application when the feasibility set is probability-distribution-dependent as well. This part uses results submitted for publication in Kopa and Rusý [2021a].

1. Interest Rate Modelling

In the first part of this thesis, we present a new calibration method for the one-factor short-rate models, a class of models which is frequently used in the financial industry for modelling interest rates. In stochastic optimization problems, it can be applied for generation of interest rate scenarios. The method has been submitted for publication in Kladívko and Rusý [2021].

We begin this chapter with establishing the notation in Section 1.1. Next, in Section 1.2, we continue with introduction of the new calibration method for the one-factor short-rate models. First, in Section 1.2.1, we introduce the one-factor short-rate models. Thereafter, in Section 1.2.2, we formulate the short rate density and then in Section 1.2.3 we add the post-calibration errors. From there, we can develop the likelihood function for market data, which is formulated in Section 1.2.4. Further, in Section 1.3 the Hull-White model is formulated under both risk-neutral and real-world measures. In Section 1.4 empirical performance of our estimation approach is investigated on EUR interest rate data, including in-sample and out-of-sample performance of the estimated model, stability of the method, and comparison with calibration methods. We stress the Section 1.4.5, where the formulae relevant for scenario generation are shown.

1.1 Interest Rate Basics

Let $P(t,T)$ denote the price at time t of a *zero-coupon bond* with a unit notional principal paid out at the maturity date T. The continuously compounded *yield* of a zero-coupon bond maturing in $T-t$ periods ahead is related to the bond price by

$$y(t,T) = -\frac{\ln P(t,T)}{T-t}, \quad t \leq T. \tag{1.1}$$

The set of yields $\{y(t,T),\ t \leq T\}$ is known as a *yield curve* at time t.

The yield curve can be unambiguously expressed in terms of *forward rates.* The forward rate

$$f(c,t,T) = -\frac{1}{T-t} \ln \frac{P(c,T)}{P(c,t)}, \quad c \leq t \leq T, \tag{1.2}$$

is an interest rate fixed, by a no-arbitrage argument, at time c between the expiry time t and maturity time T.

For the modelling purposes a theoretical concept of limiting rates is useful. The *instantaneous forward rate* is obtained by collapsing the maturity time to the expiry time:

$$f(c,t) := \lim_{T \to t^+} f(c,t,T) = -\frac{\partial \ln P(c,t)}{\partial t}. \tag{1.3}$$

We will call the set $f(c,\cdot) := \{f(c,T),\ c \leq T\}$ the *instantaneous forward curve* at time c. The yield of an instantaneous maturity bond defined by

$$r(t) := f(t,t), \tag{1.4}$$

or equivalently by $r(t) := \lim_{T \to t^+} y(t,T)$, is known as the *short rate.*

1.2 Maximum Likelihood Estimation of One-Factor Short-Rate Models

In the following section, we will present the idea of the construction of the likelihood function of one-factor short-rate models. We will introduce the *postcalibartion errors* and specify, how they enter the relationship between the model implied yield curve and the observed rates. Subsequently, we will show how one can derive the density of the yield curve for such parametrisations. From there, the likelihood function can be formulated.

1.2.1 One-Factor Short-Rate Models

First, we would like to shortly mention the difference between the real world measure and the risk neutral measure, which are generally used for specifying the short rate dynamics. The latter includes the market price of risk — a quantity derived by Vasicek [1977], which summarizes how much investors charge for holding risky positions. Their relationship was exploited in number of places, e.g. in a book of Brigo and Mercurio [2001]. In our aim to express the density of historical rates, we have to work with the model specified under the less frequently used real world measure. In general, we consider the short rate to have dynamics

$$\mathrm{d}r_c(t) = \mu_c(t, r(t))\mathrm{d}t + \sigma(t, r(t))\mathrm{d}W^{\mathbb{P}}(t),\ \ c \leq t, \tag{1.5}$$

$r_c(c) = f^M(c, c)$, $W^{\mathbb{Q}}(c) = 0$, where $\mu(\cdot, \cdot)$ and $\sigma(\cdot, \cdot)$ are some sufficiently smooth deterministic functions and $W^{\mathbb{P}}$ is a standard Brownian motion under the real-world measure $\mathbb{P}$. We denote c the calibration time of the model. The interpretation of such an equation is that the change of the short rate is driven by one deterministic factor and one random factor. Numerous short-rate models have already been described and analysed. We should mention the Vasicek [Vasicek, 1977], the Hull - White [Hull and White, 1990a] and the Cox - Ingersoll - Ross [Cox et al., 1985a] models which are frequently used in practice. We restrict ourselves to the models which imply so called affine term structure of zero-coupon bond prices. Formally, it means that the time t model implied value of a zero-coupon bond paying 1 unit at time T and calibrated at time c can be written as

$$P_c(t, T) = \exp\{A_c(t, T) - B(T - t)r_c(t)\}, \tag{1.6}$$

for some arbitrary functions $A_c(t, T)$ and $B(T - t)$. Functions A and B can also depend on the short rate model parameters, but for simplicity these will be omitted in the notation. Annualized and continuously compounded time t yield-to-maturity $t + \tau, \tau > 0$, which we denote $z_{c,t}(\tau)$ can be expressed as

$$z_{c,t}(\tau) = -\frac{1}{\tau} \log P_c(t, t + \tau) = a_c(t, \tau) + b(\tau)r_c(t), \tag{1.7}$$

where $a_c(t, \tau) = -A(t, t + \tau)/\tau$ and $b(\tau) = B(\tau)/\tau$. The model definition implies that the short rate $r_c(t)$ is given information at time s a random variable. Hence conditioned on the knowledge of information at time $s, z_{c,t}(\tau)$ is also a random variable, as $z_{c,t}(\tau)$ is a function of $r_c(t)$.

We assume a data set of n zero-coupon yields with time-invariant tenors $\tau \in \{\tau_1, \ldots, \tau_n\}$.[1] The yields observed on the market at time t are collected into the vector

$$\mathbf{y}^M(t) = \Big(y^M(t, t+\tau_1), \ldots, y^M(t, t+\tau_n)\Big)^\top,$$

where $\top$ denotes matrix transposition. The model implied yields as well as functions $a_c(t,\cdot)$ and $b(\cdot)$ are collected into n-dimensional vectors

$$\begin{aligned}\mathbf{y}_c(t) = \Big(y_c(t, t+\tau_1), \ldots, y_c(t, t+\tau_n)\Big)^\top, \\ \mathbf{b} = \Big(b(\tau_1), \ldots, b(\tau_n)\Big)^\top, \\ \mathbf{a}_c(t) = \Big(a_c(t, \tau_1), \ldots, a_c(t, \tau_n)\Big)^\top,\end{aligned}$$

to write the model implied yields at time t in the vector form

$$\mathbf{y}_c(t) = \mathbf{a}_c(t) + \mathbf{b} r_c(t). \tag{1.8}$$

1.2.2 Postcalibration Equation and Yield Curve Density

The main idea of our method is to maximise the likelihood of yields in periods subsequent to the calibration time. In order to do that we formulate the *postcalibration equation*

$$\begin{aligned}\mathbf{y}^M(t) &= \mathbf{y}_c(t) + \mathbf{W}\boldsymbol{\varepsilon}(t) \\ &= \mathbf{a}_c(t) + \mathbf{b} r_c(t) + \mathbf{W}\boldsymbol{\varepsilon}(t), \quad c < t,\end{aligned} \tag{1.9}$$

where $\boldsymbol{\varepsilon}(t)$ is an $(n-1)$–dimensional vector of *postcalibration errors*, and $\mathbf{W}$ is an $n \times (n-1)$ *error weighting matrix.* The postcalibration errors realised at time t measure the discrepancy between the market and model implied yields relative to the calibration time c.

We model the short rate as a latent factor and recover it jointly with the postcalibration errors from the vector of market yields by inverting (1.9). This system of equations can be written in the matrix multiplication form

$$\mathbf{y}^M(t) = \mathbf{a}_c(t) + \mathbf{H} \begin{pmatrix} r_c(t) \\ \boldsymbol{\varepsilon}(t) \end{pmatrix}, \quad c < t, \tag{1.10}$$

where $\mathbf{H} := \begin{pmatrix} \mathbf{b} & \mathbf{W} \end{pmatrix}$ is an $n \times n$ *transformation matrix.* Assuming invertibility of $\mathbf{H}$, we obtain that

$$\begin{pmatrix} r_c(t) \\ \boldsymbol{\varepsilon}(t) \end{pmatrix} = \mathbf{H}^{-1}(\mathbf{y}^M(t) - \mathbf{a}_c(t)), \quad c < t. \tag{1.11}$$

We call $r_c(t)$ the *inferred short rate.*

We next use the well-known change of variable technique for monotone transformations of continuous random vectors to obtain the *conditional density of*

[1] The estimation approach we develop can be modified in a straightforward way to time-varying tenors. We opted against such a generalisation because this would complicate the notation.

market yields, $g_{\mathbf{y}^M}\Big(\mathbf{y}^M(t)|\,\mathbf{y}^M(s)\Big)$, as the product of the model-implied short-rate density $g_r^{\mathbb{P}}(r_c(t)|\,r_c(s))$ and postcalibration errors density $g_{\boldsymbol{\varepsilon}}\Big(\boldsymbol{\varepsilon}(t)\Big)$:

$$g_{\mathbf{y}^M}\Big(\mathbf{y}^M(t)|\,\mathbf{y}^M(s)\Big) = \frac{1}{|\det \mathbf{H}|} g_r^{\mathbb{P}}\Big(r_c(t)|\,r_c(s)\Big) g_{\boldsymbol{\varepsilon}}\Big(\boldsymbol{\varepsilon}(t)\Big), \quad c \leq s < t, \tag{1.12}$$

where $|\det \mathbf{H}|$ denotes the absolute value of the determinant of the transformation matrix $\mathbf{H}$.

Before we formulate the likelihood function based on the conditional density of market yields, we first specify the error weighting matrix $\mathbf{W}$, and the distribution of the postcalibration error vector $\boldsymbol{\varepsilon}(t)$.

1.2.3 Postcalibration Errors Density and Transformation Matrix

We assume that the postcalibration errors are multivariate normal with zero mean vector and regular $(n-1)\times(n-1)$ variance-covariance matrix $\boldsymbol{\Sigma}$. Hence, the density function of $\boldsymbol{\varepsilon}(t)$ is

$$g_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}(t)) = \frac{1}{(2\pi)^{\frac{n-1}{2}}(\det \boldsymbol{\Sigma})^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}\boldsymbol{\varepsilon}(t)^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}(t)\right\}. \tag{1.13}$$

We further suggest the following specification of the error weighting matrix $\mathbf{W}$, and thus the transformation matrix $\mathbf{H}$:

$$\mathbf{H} := \Big(\mathbf{b} \quad \mathbf{W}\Big) = \begin{pmatrix} b(\tau_1) & 1 & 0 & \cdots & 0 \\ b(\tau_2) & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ b(\tau_{n-1}) & 0 & 0 & \cdots & 1 \\ b(\tau_n) & \frac{1}{\sqrt{n-1}} & \frac{1}{\sqrt{n-1}} & \cdots & \frac{1}{\sqrt{n-1}} \end{pmatrix}. \tag{1.14}$$

This specification of $\mathbf{W}$ has some convenient properties.

It follows from the last row of $\mathbf{W}$ that the postcalibration error of the τ_n–yield is given as a weighted sum of the $\boldsymbol{\varepsilon}(t)$ elements, with equal weights of $1/\sqrt{n-1}$. These weights provide a plausible variance of the τ_n-yield postcalibration error. For example, if $\boldsymbol{\Sigma}$ were diagonal with all elements equal to σ_ε^2, then the variance of τ_n-yield error would be equal to σ_ε^2 as well, and the covariance between the τ_n-yield error and any other error would be of order $O(1/\sqrt{n-1})$, and thus converging to zero with increasing number of yields.

The transformation matrix is sparse enough to allow for the inverse and determinant of $\mathbf{H}$ in a closed-form, which in turn improves the speed and convergence of the likelihood maximisation algorithm. After a sequence of linear algebra manipulations we obtain the inverse of $\mathbf{H}$:

$$\mathbf{H}^{-1} = \frac{1}{\widetilde{b}} \begin{pmatrix} -\frac{1}{\sqrt{n-1}} & -\frac{1}{\sqrt{n-1}} & \cdots & -\frac{1}{\sqrt{n-1}} & 1 \\ \widetilde{b} + \frac{b(\tau_1)}{\sqrt{n-1}} & \frac{b(\tau_1)}{\sqrt{n-1}} & \cdots & \frac{b(\tau_1)}{\sqrt{n-1}} & -b(\tau_1) \\ \frac{b(\tau_2)}{\sqrt{n-1}} & \widetilde{b} + \frac{b(\tau_2)}{\sqrt{n-1}} & \cdots & \frac{b(\tau_2)}{\sqrt{n-1}} & -b(\tau_2) \\ \vdots & & \ddots & & \vdots \\ \frac{b(\tau_{n-1})}{\sqrt{n-1}} & \frac{b(\tau_{n-1})}{\sqrt{n-1}} & \cdots & \widetilde{b} + \frac{b(\tau_{n-1})}{\sqrt{n-1}} & -b(\tau_{n-1}) \end{pmatrix}, \tag{1.15}$$

where

$$\widetilde{b} := b(\tau_n) - \frac{1}{\sqrt{n-1}} \sum_{i=1}^{n-1} b(\tau_i). \tag{1.16}$$

By applying the Laplace expansion along the first column and the last row of $\mathbf{H}$ we obtain the absolute value of the determinant of $\mathbf{H}$:

$$\begin{aligned} |\det \mathbf{H}| &= \left| \sum_{i=1}^{n-1} (-1)^{i+1} b(\tau_i) \cdot \frac{1}{\sqrt{n-1}} (-1)^{n-1+i} \cdot 1 + (-1)^{n+1} b(\tau_n) \cdot 1 \right| \\ &= \left| b(\tau_n) - \frac{1}{\sqrt{n-1}} \sum_{i=1}^{n-1} b(\tau_i) \right| = |\widetilde{b}|. \end{aligned} \tag{1.17}$$

The necessary and sufficient condition for regularity of $\mathbf{H}$ is $\widetilde{b} \neq 0$, which does not appear to be a limiting restriction.

One could prefer to choose a yield with a different tenor than τ_n to have the postcalibration error set as the linear combination of the other errors. This is straightforward to do by multiplying the error weighting matrix $\mathbf{W}$ by a permutation matrix $\mathbf{P}$ that swaps the rows of $\mathbf{W}$ as desired. The inverse and determinant of the permuted transformation matrix $\mathbf{H_P} := \begin{pmatrix} \mathbf{b} & \mathbf{PW} \end{pmatrix}$ are given by

$$\begin{aligned} \mathbf{H_P^{-1}} &= \begin{pmatrix} \mathbf{b} & \mathbf{PW} \end{pmatrix}^{-1} = \left(\mathbf{P} \begin{pmatrix} \mathbf{P}^{-1}\mathbf{b} & \mathbf{W} \end{pmatrix} \right)^{-1} = \begin{pmatrix} \mathbf{P}^{-1}\mathbf{b} & \mathbf{W} \end{pmatrix}^{-1} \mathbf{P}^{-1}, \\ |\det \mathbf{H_P}| &= |\det \begin{pmatrix} \mathbf{b} & \mathbf{PW} \end{pmatrix}| = |\det \mathbf{P}| \cdot |\det \begin{pmatrix} \mathbf{P}^{-1}\mathbf{b} & \mathbf{W} \end{pmatrix}| = |\det \begin{pmatrix} \mathbf{P}^{-1}\mathbf{b} & \mathbf{W} \end{pmatrix}|. \end{aligned}$$

The matrix $\begin{pmatrix} \mathbf{P}^{-1}\mathbf{b} & \mathbf{W} \end{pmatrix}$ is obtained by swapping elements of the $\mathbf{b}$ vector according to $\mathbf{P}^{-1}$, whereas $\mathbf{P}^{-1} = \mathbf{P}^\top$. The inverse and determinant of $\mathbf{H_P}$ then follow from (1.15) and (1.17), respectively.

1.2.4 Data Set and Likelihood Function

The time series of yield curves

$$\{\mathbf{y}^M(t_k) = (y^M(t_k, t_k + \tau_1), \ldots, y^M(t_k, t_k + \tau_n))^\top\}_{k=1}^K \tag{1.18}$$

consists of K yield curve observations at times $t_1, \ldots, t_K$, whereas each yield curve has n yields with tenors $\tau_1, \ldots, \tau_n$.

The *conditional likelihood function* of the parameter vector $\boldsymbol{\theta}$ which includes all parameters of the short-rate model and the variance-covariance matrix $\boldsymbol{\Sigma}$ is given by

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\Sigma}) &= \prod_{k=1}^{K-1} g_{\mathbf{y}^M} \left(\mathbf{y}^M(t_{k+1}) | \mathbf{y}^M(t_k) \right) \\ &= \prod_{k=1}^{K-1} \frac{1}{|\det \mathbf{H}|} g_r^{\mathbb{P}} \left(r_{c_k}(t_{k+1}) | r_{c_k}(t_k) \right) g_{\boldsymbol{\varepsilon}} \left(\boldsymbol{\varepsilon}(t_{k+1}) \right), \quad c_k \in \{t_1, \ldots, t_k\}. \end{aligned} \tag{1.19}$$

Note that the likelihood function is conditioned on the first observation $\mathbf{y}^M(t_1)$, and that the density factorisation (1.12) has been used.

It is computationally convenient to work with the *log-likelihood function* given by

$$
\begin{aligned}
&\ln\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\Sigma}) \\
&= \sum_{k=1}^{K-1} \Big(-\ln|\det \mathbf{H}| + \ln g_r^{\mathbb{P}}\Big(r_{c_k}(t_{k+1}) | r_{c_k}(t_k)\Big) + \ln g_{\varepsilon}\Big(\boldsymbol{\varepsilon}(t_{k+1})\Big)\Big) \\
&\propto -(K-1)\ln|\widetilde{b}| + \sum_{k=1}^{K-1}\left(-\frac{1}{2}\ln v^2_{t_{k+1}|t_k} - \frac{1}{2v^2_{t_{k+1}|t_k}}\Big(r_{c_k}(t_{k+1}) - m^{\mathbb{P}}_{t_{k+1}|t_k,c_k}\Big)^2\right) \\
&\qquad - \frac{K-1}{2}\ln|\det \boldsymbol{\Sigma}| - \sum_{k=1}^{K-1} \boldsymbol{\varepsilon}(t_{k+1})^{\top}\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}(t_{k+1}).
\end{aligned} \tag{1.20}
$$

We further simplify the log-likelihood function by using the maximum likelihood estimator of $\boldsymbol{\Sigma}$. This results in the *profile log-likelihood function* of the parameter vector $\boldsymbol{\theta}$ given by

$$
\begin{aligned}
\ell(\boldsymbol{\theta}) &= \max_{\boldsymbol{\Sigma}} \ln\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\Sigma}) \\
&\propto \sum_{k=1}^{K-1}\left(-\frac{1}{2}\ln v^2_{t_{k+1}|t_k} - \frac{1}{2v^2_{t_{k+1}|t_k}}\Big(r_{c_k}(t_{k+1}) - m^{\mathbb{P}}_{t_{k+1}|t_k,c_k}\Big)^2\right) \\
&\qquad - (K-1)\ln|\widetilde{b}| - \frac{K-1}{2}\ln\Big|\det\hat{\boldsymbol{\Sigma}}\Big|,
\end{aligned} \tag{1.21}
$$

where

$$
\hat{\boldsymbol{\Sigma}} = \frac{1}{K-1}\sum_{k=1}^{K-1}\boldsymbol{\varepsilon}(t_{k+1})\boldsymbol{\varepsilon}(t_{k+1})^{\top}. \tag{1.22}
$$

The idea of maximising the likelihood of the observed market yields and retrieving the short rate as a latent variable comes from Chen and Scott [1993]. While our estimation method is based on the same idea it differs in several important aspects. Chen and Scott's method is developed for time-homogeneous models which do not fit the yield curve perfectly, and thus lead to measurement errors rather than postcalibration errors. Chen and Scott formulate the likelihood for just four tenors, and assume that the yield with the shortest tenor is measured perfectly. This assumption implies one-to-one map between this yield and the short rate, and hence restricts the short rate's flexibility, and also its latent variable interpretation. Lastly, the likelihood maximisation of Chen and Scott's model is numerically more involved as the profile log-likelihood is not available.

1.2.5 Confidence Intervals

In this section we suggest two types of confidence intervals for the model parameters $\boldsymbol{\theta}$. For the concepts and proofs associated with the confidence intervals construction we refer to Hamilton [1994, Chapters 5 nad 14].

The true parameter values are collected in the vector $\boldsymbol{\theta}_0$, and we consider the null hypothesis,

$$
\mathrm{H}_0: \quad \boldsymbol{\theta} = \boldsymbol{\theta}_0,
$$

against the alternative

$$\mathrm{H_A}: \quad \boldsymbol{\theta} \neq \boldsymbol{\theta}_0.$$

First, we consider confidence intervals based on the likelihood ratio test. It follows from standard results on the likelihood ratio test that the log-likelihood ratio $2\Big(\ell(\hat{\boldsymbol{\theta}}) - \ell(\boldsymbol{\theta}_0)\Big)$ has under the null hypothesis asymptotically $\chi^2(m)$ distribution, where m is the number of short-rate model parameters. The $1-p$ confidence region of the parameter vector $\boldsymbol{\theta}$ under the H_0 is given by the set

$$\Big\{\boldsymbol{\theta}: \quad 2\Big(\ell(\hat{\boldsymbol{\theta}}) - \ell(\boldsymbol{\theta})\Big) \leq \chi^2_{1-p}(m)\Big\}, \tag{1.23}$$

where $p \in (0,1)$ is the significance level, and $\chi^2_{1-p}(m)$ is the $1-p$ percentile of χ^2-distribution with m degrees of freedom. The confidence intervals can then be calculated from an outer approximation of the confidence region by a multi-dimensional rectangle.

Second, we construct confidence intervals based on the Wald test. These are calculated from a quadratic approximation of the likelihood function. The standard results state that under the null hypothesis

$$\sqrt{K}\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right) \longrightarrow \mathcal{N}(0, I\left(\boldsymbol{\theta}_0)^{-1}\right). \tag{1.24}$$

The distribution of the parameter estimates vector is approximated by normal distribution with expected value $\boldsymbol{\theta}_0$ and variance-covariance matrix $K^{-1}I(\boldsymbol{\theta}_0)^{-1}$. The matrix $I(\boldsymbol{\theta}_0)$ is known as the Fisher information matrix, which is consistently estimated by the observed Fisher information matrix $J_K(\hat{\boldsymbol{\theta}})$. This matrix is obtained by calculating the negative Hessian of the likelihood function evaluated at the achieved maximum, that is

$$J_K(\hat{\boldsymbol{\theta}}) = \frac{1}{K}\left[-\frac{\partial^2}{\partial\boldsymbol{\theta}_i \partial\boldsymbol{\theta}_j}\ell(\boldsymbol{\theta})\right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}. \tag{1.25}$$

The $1-p$ confidence region of the parameter vector $\boldsymbol{\theta}$ under the H_0 is given by the set

$$\Big\{\boldsymbol{\theta}: \quad (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^\top \left[-\frac{\partial^2}{\partial\boldsymbol{\theta}_i \partial\boldsymbol{\theta}_j}\ell(\boldsymbol{\theta})\right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \leq \chi^2_{1-p}(m)\Big\}, \tag{1.26}$$

where $p \in (0,1)$ is the significance level, and $\chi^2_{1-p}(m)$ is the $1-p$ percentile of χ^2-distribution with m degrees of freedom. For the maximum likelihood problem we resort to numerical evaluation of the Hessian as an analytical expression is not known. Consequently, we obtain the standard errors of individual parameter estimates as the square root of the corresponding diagonal elements in the negative Hessian matrix.

1.3 The Hull-White Model

1.3.1 The Hull-White Model under the Risk-Neutral Measure $\mathbb{Q}$

Hull and White [1990b] suggest several extensions of one-factor time-homogeneous interest rates models with constant parameters to time-inhomogeneous models with time-varying parameters. We will focus on their most popular extension known as the Hull-White model (Hull and White [1994]).

In the Hull-White model the constant unconditional mean of the Vašíček model [Vasicek, 1977, Section 5] is replaced by a function of time and market forward rate curve. The model nests on the short-rate process

$$\mathrm{d}r_c(t) = (\mu_c(t) - \alpha r_c(t))\,\mathrm{d}t + \sigma \mathrm{d}W^{\mathbb{Q}}(t),\ c \leq t, \tag{1.27}$$

$r_c(c) = f^M(c,c)$, $W^{\mathbb{Q}}(c) = 0$, where σ is the volatility parameter and $W^{\mathbb{Q}}$ is a standard Brownian motion under the risk-neutral measure $\mathbb{Q}$. The short rate is attracted to the $\mu_c(t)/\alpha$ path, where α controls the force of attraction. The time-varying parameter $\mu_c(t)$ is given by

$$\mu_c(t) = \frac{\partial f^M(c,t)}{\partial t} + \alpha f^M(c,t) + \frac{\sigma^2}{2\alpha}\left(1 - e^{-2\alpha(t-c)}\right), \quad c \leq t, \tag{1.28}$$

where $f^M(c,t)$ is the instantaneous forward rate prevailing on the market at time $c \leq t$. Note that we use the superscript M to distinguish a market observable quantity from its model implied counterpart.

We say that the model is *calibrated* at time c when the market instantaneous forward rate curve $f(c,\cdot)$ is used to set the time-varying parameter $\mu_c(\cdot)$. Instantaneous forward rates are not quoted directly, but they can be inferred from market data by yield curve fitting methods.

The short rate is normally distributed with conditional expected value given as $m^{\mathbb{Q}}_{t|s,c} := \mathsf{E}^{\mathbb{Q}}\left[r_c(t)|\, r_c(s)\right]$, and variance, $v^2_{t|s} := \mathrm{var}\left[r_c(t)|\, r_c(s)\right]$, given by

$$\begin{aligned} m^{\mathbb{Q}}_{t|s,c} &= r_c(s)e^{-\alpha(t-s)} + f^M(c,t) - f^M(c,s)e^{-\alpha(t-s)} \\ &\quad + \frac{\sigma^2}{2\alpha^2}\left(1 - e^{-\alpha(t-s)} + e^{-2\alpha(t-c)} - e^{-\alpha(t-2c+s)}\right), \end{aligned} \tag{1.29}$$

$$v^2_{t|s} = \frac{\sigma^2}{2\alpha}\left(1 - e^{-2\alpha(t-s)}\right), \quad c \leq s < t. \tag{1.30}$$

The expected value (1.29) is adapted to the natural filtration of the short-rate process, but since it is also a function of the calibration time, we include c in the notation.

The Hull-White zero-coupon bond pricing function is given by

$$P_c(t,T) = \exp\{A_c(t,T) - B(t,T)r_c(t)\}, \tag{1.31}$$

where

$$B(t,T) = \frac{1 - e^{-\alpha(T-t)}}{\alpha}, \tag{1.32}$$

$$\begin{aligned} A_c(t,T) &= -f^M(c,t,T)(T-t) + B(t,T)f^M(c,t) \\ &\quad - B(t,T)^2\frac{\sigma^2}{4\alpha}\left(1 - e^{-2\alpha(t-c)}\right), \end{aligned} \tag{1.33}$$

and where $-f^M(c,t,T)(T-t) = \ln P^M(c,T) - \ln P^M(c,t)$.

It is important to note that at the calibration time c the Hull-White model fits the market yield curve perfectly for any value of α and σ. The equality

$P_c(c,T) = P^M(c,T)$ for all T can be readily seen by using $f(c,c) = r_c(c)$ and $P(c,c) = 1$ when evaluating the bond price (1.31) at time $t = c$. However, at some later time $t > c$, the Hull-White bond price $P_c(t,T)$ can be different from the market price $P^M(t,T)$. Recalibrating the model at time t to $f^M(t,\cdot)$ can be viewed as marking the model to market.

1.3.2 The Hull-White Model under the Real-World Measure $\mathbb{P}$

As we have discussed in the previous paragraph the Hull-White model fits the market yield curve perfectly at the calibration time for any value of α and σ, and therefore it is not possible to identify these two parameters from a single observation of the yield curve. Our estimation method uses a time series of yield curves. Time series data are, however, generated under the real-world measure $\mathbb{P}$, and thus we need to bring the Hull-White model under the real-world measure. The transformation of measures follows from the Girsanov theorem and it is carried out by introducing

$$\mathrm{d}W^{\mathbb{Q}}(t) = \mathrm{d}W^{\mathbb{P}}(t) + \lambda(t)\mathrm{d}t, \tag{1.34}$$

where $W^{\mathbb{P}}$ is a standard Brownian motion under the real-world measure, and λ is an Itô process that is called the *market price of interest rate risk* (λ is subject to the Novikov condition, see [Øksendal, 2007, Section 8.6] for details).

In this thesis we assume that the market price of risk is governed by the constant, that is, $\lambda(t) := \lambda$. Then it follows that the short-rate dynamics under the real-world measure are

$$\mathrm{d}r_c(t) = \left(\mu_c(t) + \lambda\sigma - \alpha r_c(t)\right)\mathrm{d}t + \sigma \mathrm{d}W^{\mathbb{P}}(t). \tag{1.35}$$

We can immediately see that the change of measure introduces a parallel shift of the time-varying parameter $\mu_c(t)$, and thus introduces an additional term to the conditional expected value, which is given by

$$m^{\mathbb{P}}_{t|s,c} = m^{\mathbb{Q}}_{t|s,c} + \frac{\lambda\sigma}{\alpha}\left(1 - e^{-\alpha(t-s)}\right) \tag{1.36}$$

under the real-world measure. The real-world transition density of $r_c(t)$ conditional on $r_c(s)$ is

$$g^{\mathbb{P}}_r(r_c(t)|\, r_c(s)) = \frac{1}{\sqrt{2\pi v^2_{t|s}}} \exp\left\{ -\frac{\left(r_c(t) - m^{\mathbb{P}}_{t|s,c}\right)^2}{2v^2_{t|s}} \right\}, \quad c \leq s < t. \tag{1.37}$$

A more flexible specification of the market price of risk has been suggested in the literature; Ang and Piazzesi [2003], Hamilton and Wu [2012], Joslin et al. [2011] and Nyholm and Vidova-Koleva [2012] among many others postulate a linear relationship between the price of risk and the level of the short rate. Specifically, the affine form $\lambda(t) := \lambda + \lambda_r r_c(t)$, where λ and λ_r are constants, is assumed. The short rate remains normally distributed under the affine market price of risk, and therefore our maximum likelihood method can accommodate this price of risk specification by a straightforward generalisation. We note that the $\lambda_r r_c(t)$

term will be absorbed by the parameter α, which thus, for $\lambda_r \neq 0$, takes different values under the two measures. Nonlinear, time-dependent, or exogenous (dependent on other sources of uncertainty) forms of the market price of risk would require a separate analysis.

1.3.3 Calibration Time

For yield curve data sampled at times $t_1, t_2, \ldots, t_K$ various strategies to the calibration time c_k can be considered. In the next section we calibrate to each available yield curve, that is for the yield curve observed at time t_{k+1}, we set $c_k = t_k$ and consider the postcalibration equation

$$\mathbf{y}^M(t_{k+1}) = \mathbf{a}_{t_k}(t_{k+1}) + \mathbf{b}r_{t_k}(t_{k+1}) + \mathbf{W}\boldsymbol{\varepsilon}(t_{k+1}). \tag{1.38}$$

The short-rate conditional moments (1.36) and (1.30) are then modified to

$$\begin{aligned} m^{\mathbb{P}}_{t_{k+1}|t_k,t_k} &= f^M(t_k, t_{k+1}) + \frac{\sigma^2}{2\alpha^2}\Big(1 - e^{-\alpha(t_{k+1}-t_k)}\Big)^2 \\ &\quad + \frac{\lambda\sigma}{\alpha}\Big(1 - e^{-\alpha(t_{k+1}-t_k)}\Big), \end{aligned} \tag{1.39}$$

$$v^2_{t_{k+1}|t_k} = \frac{\sigma^2}{2\alpha}\Big(1 - e^{-2\alpha(t_{k+1}-t_k)}\Big). \tag{1.40}$$

The conditional mean is simplified due to the instantaneous rate identity $r_{t_k}(t_k) = f^M(t_k, t_k)$.

1.4 Empirical Study

In this section we use the suggested maximum likelihood method to estimate the Hull-White model from time series of EUR yields. We analyse the model performance from both in-sample and out-of-sample perspective, discuss estimation challenges, and also compare our estimation approach with calibration to swaption prices.

1.4.1 Data

We use Euribor and swap benchmark rates calculated by IBA[2] and downloaded from Bloomberg. Euribor rates with tenors one week (1W), one, two, three and six months (1M, 2M, 3M, 6M), and swap par rates with tenors one to ten years and 15 years (1Y–10Y, 15Y) are bootstrapped by cubic splines [Hagan and West, 2006] to obtain continuously compounded zero-coupon yields under the Act/Act day-count convention.[3] The cubic splines are also used to estimate forward rate curves. We check that the spline method in use, which fits the market rates perfectly, provides stable and smooth forward curves.

First, in Section 1.4.2 and 1.4.3 we illustrate estimation results based on daily data over a two-year period from 1 September 2017 to 30 August 2019. Second,

[2]The ICE Benchmark Administration, see `https://www.theice.com/iba/ice-swap-rate`.

[3]Euribor rates are zero-coupon rates under the Act/360 convention. The swap rates in use are annual coupon par rates under the 30/360 convention.

in Section 1.4.4 we expand the data set until 21 January 2021 and assess the stability of parameter estimates by running the estimation on a rolling window with a weekly frequency. Third, in Section 1.4.6 we conduct an out-of-sample forecasting exercise by using monthly data over 20 years from January 2001 until January 2021. Fourth, in Section 1.4.7 we use daily data from the beginning of 2009 until the end of 2016 to compare parameter estimates obtained by our maximum likelihood method with parameter values obtained by calibrating the Hull-White model to market prices of swaptions.

1.4.2 Estimating the Model Parameters

In this section we estimate the Hull-White model on daily data spanning a two-year period from 1 September 2017 to 30 August 2019.[4] The data set of zero-coupon yields is captured in Figure 1.1, where the vertical dashed line on 30 August 2019 denotes the end of the sample and thus the date to which the parameters are estimated.

The yields are persistently trending downwards to the negative territory (all of the yields are negative in the last 12 days of our data set). The yield curve is upward sloping during the considered period with the yield curve slope decreasing over time. However, the shorter end of the yield curve becomes inverted, as the slope, measured by $y(1\text{Y}) - y(1\text{W})$, takes a negative value on 2 July 2019. The correlation between the swap-based yields is very strong (a pairwise correlation is at least 0.98 for yields with tenors from four to 15 years), while the correlation is moderate between the Euribor- and swap-based yields (a pairwise correlations ranges from 0.47 to 0.96).

We select the 13 most liquid tenors (1M, 2M, 3M, 6M, 1Y–7Y, 10Y, 15Y) to form the likelihood function, and use the transformation matrix specification (1.14). We use the derivative-free Nelder-Mead algorithm implemented in the R package `nmkb`[5] to maximise the profile log-likelihood function (1.21)–(1.22) subject to the postcalibration equation (1.38). We assume $\Delta t := t_{k+1} - t_k$ constant for all k, that is, we work with sampling of yields as if it was equidistant. For daily data we set $\Delta t = 1/252$, for monthly data in Section 1.4.6 we set $\Delta t = 1/12$. The numerical maximisation converges to an optimum and returns parameter estimates reported in Table 1.1. Note that the reported values, in particular the volatility parameter σ, are for interest rates measured in percent.

The point estimates of the Hull-White model parameters are given in the second column (Point Est). The 95% confidence intervals based on the likelihood ratio test, LR intervals hereafter, are reported in the third column (LR 95% CI). The LR intervals are based on the smallest three-dimensional rectangles that cover the confidence set (1.23). Finally, the 95% confidence intervals based on the Wald test, the Wald intervals hereafter, are presented in the fourth column (Wald 95% CI). The Hessian in (1.26) is calculated by using the R package `numDeriv`[6]. The Wald intervals are by construction symmetric, whereas the LR intervals are not. Also note that the LR intervals are wider, for all parameters, than the Wald

[4] According to our discussions with risk management practitioners, using daily interest rate data from past two years is believed to reflect the prevailing interest rate regime.

[5] See `https://rdrr.io/cran/dfoptim/man/nmkb.html`.

[6] See `https://cran.r-project.org/web/packages/numDeriv/index.html`.

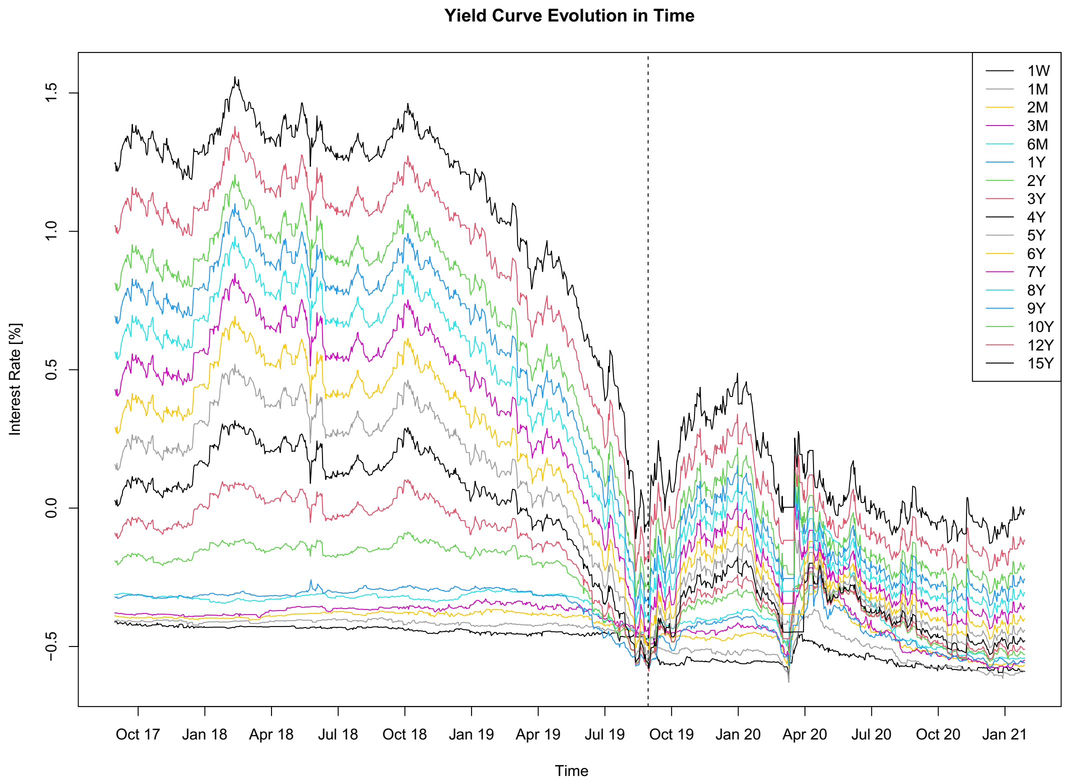

Figure 1.1: Euribor-based yields (1W, 1M, 2M, 3M, 6M) and swap-based yields (1Y, 2Y, 3Y, 4Y, 5Y, 6Y, 7Y, 8Y, 9Y, 10Y, 12Y, 15Y) used in the Hull-White model estimation. The vertical dashed line denotes 30 August 2019. The time series of yields until this date is used to estimate the Hull-White model in Section 1.4.2. Parameter estimates based on a rolling window that covers the whole sample period are presented in Section 1.4.4

Param.	Point Est.	LR 95% CI	Wald 95% CI
α	0.050	(0.015, 0.093)	(0.024, 0.076)
σ	0.107	(0.098, 0.117)	(0.100, 0.113)
λ	-0.283	$(-2.143, 1.531)$	$(-1.649, 1.083)$

Table 1.1: The maximum likelihood estimates of the Hull-White model parameters for interest rates measured in percent. Point Est. denotes the point estimate. LR denotes the 95% confidence intervals based on the likelihood ratio, Wald denotes the 95% confidence intervals based on the Wald test.

intervals. Recall that the LR intervals are calculated as an outer approximation of the confidence region. We note that the estimate of the market price of interest rate risk is fairly uncertain, and $\lambda = 0$ can not be rejected on the 5% significance level.

The maximum likelihood estimate of the variance-covariance matrix of post-calibration errors is reported in the form of correlations and standard deviations in Table 1.2. More precisely, what is reported are results of $\mathbf{W}\hat{\mathbf{\Sigma}}\mathbf{W}^{\top}$, where $\hat{\mathbf{\Sigma}}$ is given by (1.22), and $\mathbf{W}$ is the error weighting matrix given in (1.14). In order to save space, we report values for roughly every other tenor; the unreported values

τ	1M	3M	1Y	3Y	5Y	7Y	10Y	15Y
1M	0.70	1.00	0.82	−0.24	−0.35	−0.24	−0.10	0.05
3M	1.00	0.66	0.87	−0.17	−0.29	−0.19	−0.04	0.11
1Y	0.82	0.87	0.54	0.25	0.07	0.16	0.31	0.45
3Y	−0.24	−0.17	0.25	0.90	0.96	0.95	0.95	0.94
5Y	−0.35	−0.29	0.07	0.96	1.66	0.99	0.96	0.91
7Y	−0.24	−0.19	0.16	0.95	0.99	2.17	0.99	0.95
10Y	−0.10	−0.04	0.31	0.95	0.96	0.99	2.69	0.98
15Y	0.05	0.11	0.45	0.94	0.91	0.95	0.98	3.21

Table 1.2: The maximum likelihood estimate of the variance-covariance matrix of the postcalibration errors in the form of correlations and standard deviations, which are reported on the diagonal of the matrix.

are approximately equal to the average of two neighbouring values. The Euribor-related errors (1M, 2M, 3M, 6M) are negatively correlated with the swap-related errors (3Y, 4Y, 5Y, 6Y, 7Y, 10Y) as the Hull-White model attempts to capture the correlation structure between Euribor- and swap-based yields. However, the correlation between the Euribor-related errors and 15-year yield related error has positive sign. The reason is that the postcalibration error of the 15-year yield is given as the weighted sum of all remaining postcalibration errors. The standard deviation of postcalibration errors follows the volatility pattern in the yields as it increases with the tenor. The average standard deviation of the errors is just 1.1 basis point, which is about one order smaller than the average volatility of yields.

We next assess how well the imposed modelling assumptions hold in the estimated model. First, we discuss the inferred short rate, the instantaneous forward rate, and assess whether the difference between the inferred short-rate process and its conditional mean (1.39) is normally distributed and serially uncorrelated. Second, we examine the cross-correlation and normality of the postcalibration errors.

The inferred short rate and the instantaneous forward rate is recorded in the top-left panel of Figure 1.2. These two rates move within the range of the one-week and six-month Euribor-based yields. The top-left panel also records Eonia, which is the EUR overnight lending reference rate. The overnight tenor suggests Eonia to be used as an empirical proxy for the short rate. However, we argue that Eonia is rather disconnected from the yield curve we aim to model. First, Eonia is only weakly correlated with the other yields; the average of the pairwise correlation coefficients is 0.11, whereas it is at least 0.55 for the other yields (with Eonia excluded from the average). Second, Eonia values exhibit spikes which are not observed in the other yields. The dynamics of the inferred short rate and instantaneous forward rate are rather aligned with the dynamics of Euribor- and swap-based yields. Indeed, the average of the pairwise correlation coefficients between these yields and the inferred short rate or instantaneous forward rate is 0.65.[7] This is not surprising since the short rate is the only factor that drives the evolution of the whole yield curve.

[7] The correlation between Eonia and the inferred short rate is −0.03, and −0.06 between Eonia and the instantaneous forward rate (the correlation coefficients are not significant).

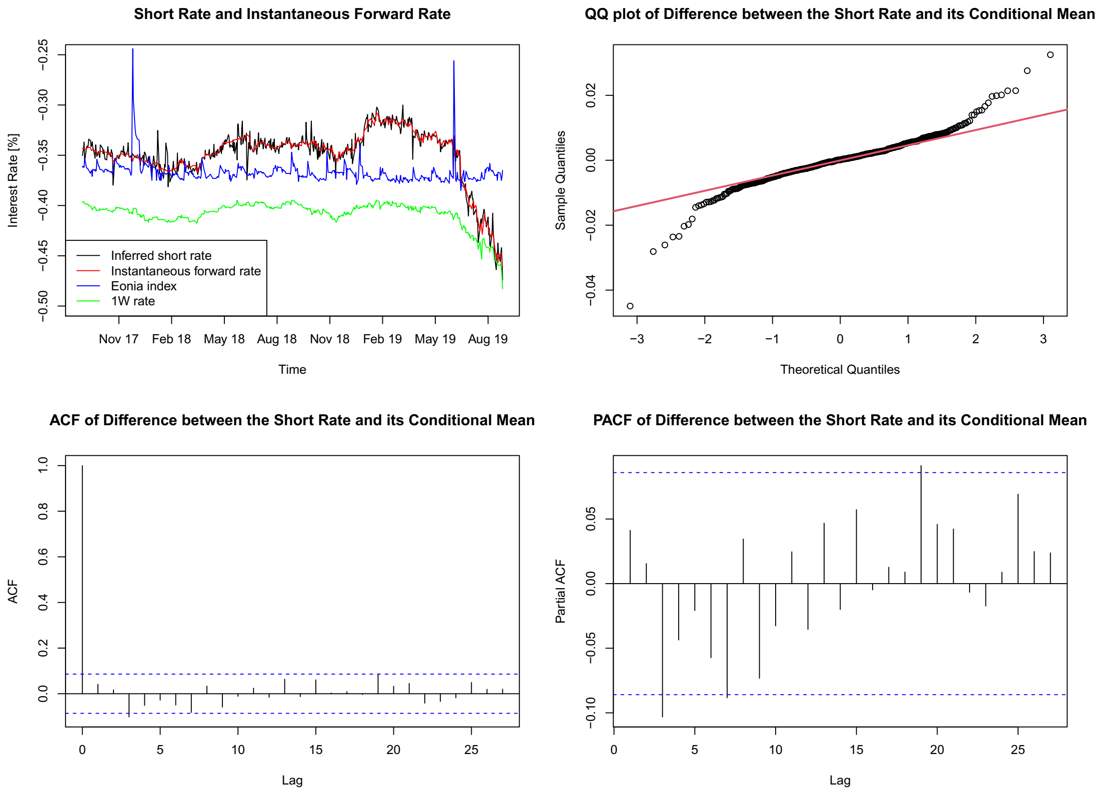

Figure 1.2: The inferred short rate, the instantaneous forward rate, Eonia index, and one-week rate in the top-left panel. The Normal Q–Q plot of the difference between the inferred short rate and its conditional mean in the top-right panel. The autocorrelation function (ACF) of difference between the short rate and its conditional mean in the bottom-left panel, and the partial autocorrelation function (PACF) for the same difference in the bottom-right panel.

The differences between the inferred short rate and its conditional mean are assumed to be realisations of the Brownian motion increments, and thus independently and identically normally distributed. The Normal Q–Q plot in the top-right panel shows that this difference is symmetric, but exhibits heavier tails than those of normal distribution. Indeed, as it is often the case in asset pricing models, normality is rejected by Jarque-Bera and Shapiro-Wilk tests at the 1% level. Serial correlation of the increments is examined in the bottom panels of Figure 1.2 by plotting the autocorrelation function (ACF) and partial autocorrelation function (PACF). We can reject serial correlation structure at the 5% level.

Regarding the postcalibration errors we note that their lagged cross-correlations are relatively small with the highest value of 0.20 for lag 1 cross-correlation (significant at the 5% level). Furthermore, the normality of the postcalibration errors is rejected at the 1% level. These two empirical inconsistencies with our modelling assumptions could be possibly nullified by introducing an autocorrelation structure and a heavy-tailed distribution (a natural candidate would be the t-distribution) for the postcalibration errors. We do not consider these two modifications here.

1.4.3 The Likelihood Function Maximisation

We conclude our numerical investigation with an examination of the likelihood function. We always fix one of the Hull-White model parameters at its maximum likelihood point estimate and vary the remaining two parameters to see their effect on the value of the profile log-likelihood function (1.21)–(1.22). The results are presented in Figure 1.3 in the form of heat maps. In the top pair of panels $\lambda = -0.283$, while σ and α are varied. In the middle pair of panels $\sigma = 0.107$ while α and λ are varied. Finally, in the lower pair of panels $\alpha = 0.050$ while λ and σ are varied. The heat maps suggest that the profile log-likelihood function has a well localised global maximum. We can clearly pin down the maximum value of 84 014 for all three pairs of varied parameters in the right-hand side panels (values smaller than 84 004 are in the same brownish colour).

The heat maps in the left-hand side panels cover a larger interval of likelihood values in order to illustrate monotonicity of the gradient of the likelihood function. Thus, an optimisation algorithm should be able to find its way to the optimal value from a wide range of initial values. The ease of the likelihood maximisation is an appealing property. It has been well documented that likelihood maximisation of time-homogeneous short-rate models is challenging (for example, Ang and Piazzesi [2003], Joslin et al. [2011], or Kim [2009]). The principle obstacles being the flatness of the likelihood surface in the parameters which determine the unconditional mean of yields. There is a large empirical evidence that interest rates are strongly persistent and thus close to nonstationarity.[8] The inability to reject the nonstationarity indicates that the unconditional mean of the interest rate process is not well defined, and thus challenging to estimate, see Hamilton and Wu [2012] for an interesting discussion of this topic. This issue is non-existent in no-arbitrage models as the mean of yields is time varying and calibrated to market yield curve data. This may explain why our likelihood maximisation of the Hull-White model is straightforward, but more investigation is needed.

1.4.4 Stability of Parameter Estimates over Time

In this section we expand the data set until 29 January 2021, and reestimate the model parameters on a weekly basis by using the same approach as described in Section 1.4.2. In particular, for each reestimation we use two years of data and the same 13 tenors in the likelihood. The evolution of the parameter estimates is presented in Figure 1.4; the red curve provides the points estimates, whereas the area shaded in grey marks the 95% confidence intervals based on the Wald test (see Section 1.2.5).

In total, 74 parameter sets were estimated. Note that the left-most estimates are those reported in Table 1.1. We can observe that the parameter estimates do not exhibit any abrupt changes that would signal instability of our estimation method. We can also note that the width of confidence intervals is relatively stable over time. The α estimates are markedly stable around 0.05 with a relatively wide confidence interval. We can observe a small increase of $\hat{\alpha}$ with the beginning of the Covid-19 pandemics in March 2020. The impact of the pandemics is clearly

[8] Indeed, Dickey-Fuller tests do not reject nonstationarity of yields in our data set at the 5% significance level.

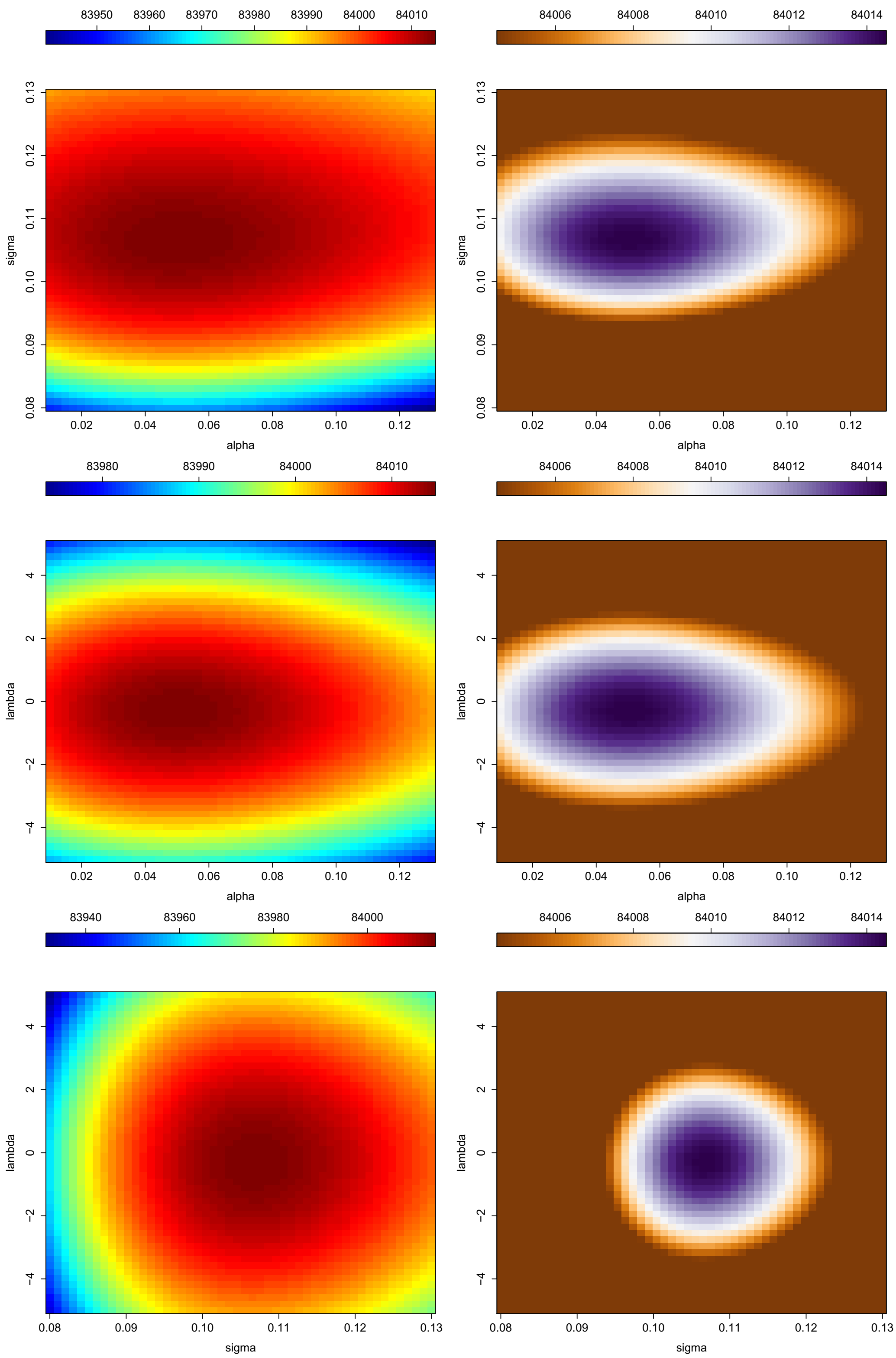

Figure 1.3: Heat maps of profile log-likelihood function values. In the top pair of panels $\lambda = -0.283$ while σ and α are varied. In the middle pair of panels $\sigma = 0.107$ while α and λ are varied. In the lower pair of panels $\alpha = 0.050$ while λ and σ are varied. The heat maps in the left-hand side panels cover a larger interval of likelihood values compared to the heat maps in the right-hand side panels.

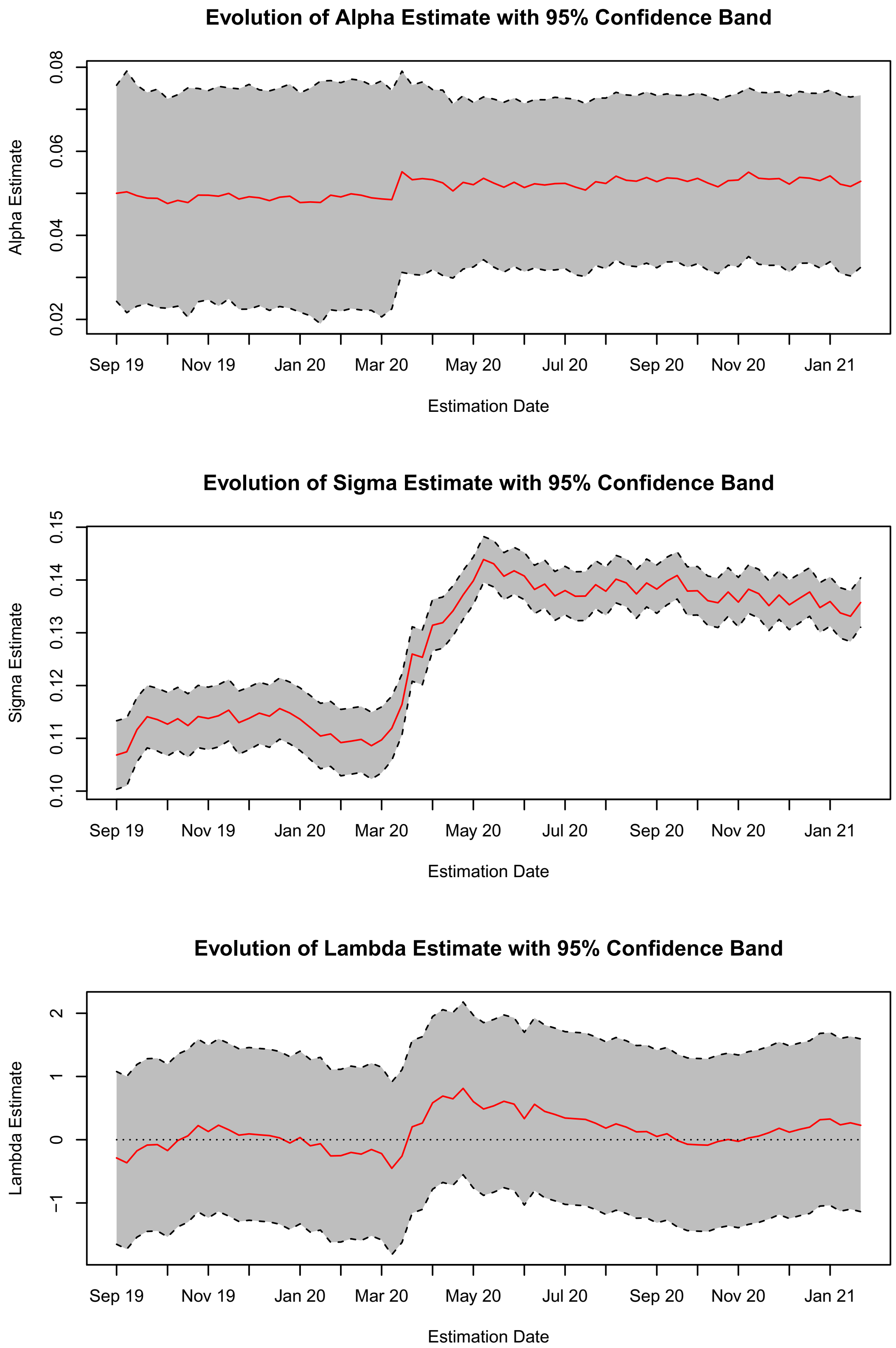

Figure 1.4: Evolution of parameters estimates over 74 weeks from 30 August 2019 until 29 January 2021. The point estimates are recorded by the red curve. The 95% confidence intervals based on the Wald test are shaded in grey.

manifested in the σ and λ estimates. The estimate of the volatility parameter σ has increased about almost 50% during the late spring of 2020 and remained on this level until the end of our sample. The estimate of the market price of risk parameter λ also quickly increased to positive values during the first month of the pandemics and has declined since then. We also note that $\lambda = 0$ cannot be rejected during the considered period. However, as we illustrate in Section 1.4.6, the point estimate of λ is crucial for the real-world performance of the model.

1.4.5 Model Forecasts, Scenario Generation

We now elaborate on yields forecasting from the Hull-White model. A point forecast is based on the conditional expected value as this type of forecast minimizes the mean squared error [Hamilton, 1994, Chapter 4]. The out-of-sample forecast from the Hull-White model of a τ-year yield at the h-year horizon is then given by

$$\begin{aligned} &\hat{y}_t^{\mathbb{P}}(t+h, t+h+\tau) \\ &\quad := \mathsf{E}^{\mathbb{P}}\left[a_t(t+h,\tau) + b(\tau)r_t(t+h)\Big|r_t(t), f^M(t,\cdot)\right] \\ &\quad = f^M(t, t+h, t+h+\tau) + b(\tau)^2\frac{\tau\sigma^2}{4\alpha}\left(1-e^{-2\alpha h}\right) \\ &\qquad + b(\tau)\frac{\sigma^2}{2\alpha^2}\left(1-e^{-\alpha h}\right)^2 + b(\tau)\frac{\lambda\sigma}{\alpha}\left(1-e^{-\alpha h}\right). \end{aligned} \tag{1.41}$$

The forecast is generated under the real-world measure at time t when the model was recalibrated to $f^M(t,\cdot)$. The formula follows directly from the Hull-White yield function (1.31)–(1.33) and the conditional expected value of the short-rate process (1.36). As we estimate the model under both measures, we can also generate forecasts under the risk-neutral measure $\mathbb{Q}$. It follows from the Girsanov transformation of measures that the risk-neutral forecast is given by

$$\hat{y}_t^{\mathbb{Q}}(t+h, t+h+\tau) = \hat{y}_t^{\mathbb{P}}(t+h, t+h+\tau) - b(\tau)\frac{\lambda\sigma}{\alpha}\left(1-e^{-\alpha h}\right). \tag{1.42}$$

The sign of the difference between the real-world and risk-neutral forecast is thus decided by the sign of the market price of risk parameter λ.

We set 30 August 2019 as the forecast origin date, that is the date to which we have estimated the model parameter in Section 1.4.2, and present yield curve forecasts at horizons from one to five years in Figure 1.5. The left-hand side panel presents forecasts under the real-world measure $\mathbb{P}$, whereas the right-hand side panel presents forecasts under the risk-neutral measure $\mathbb{Q}$. Observe that the yield curve shape at the one-year forecast horizon is similar to the initial yield curve shape, but with less pronounced short-end inversion and convexity. The expected yield curves are increasing and concave beyond the two-year forecast horizon. Since we have estimated $\widehat{\lambda} = -0.283$, the difference between the real-world and risk-neutral forecast is negative, and the risk-neutral expectations are higher (for any tenor and at any horizon) than real-world expectations. For example, the forecast of one-year yield at the five-year horizon is -0.43% under the $\mathbb{P}$ measure, whereas it is -0.30% under the $\mathbb{Q}$ measure. We note that the real-world forecasts should be used if future values of yields per se are of interest, for example, in

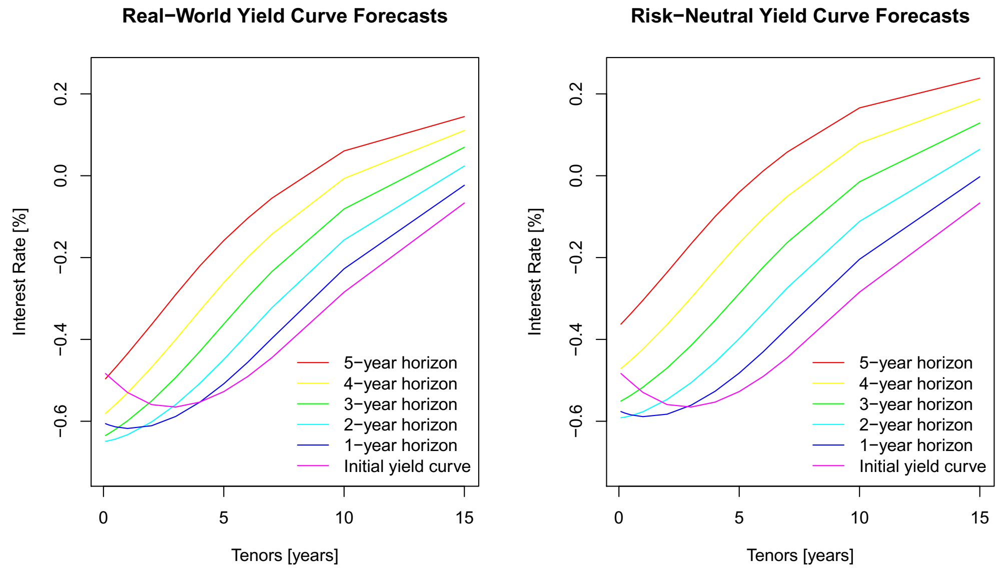

Figure 1.5: The Hull-White model yield curve expectations at forecast horizons from one to five years. The forecast origin date is 30 August 2019. The left-hand side panel captures the real-world expectations, while the risk-neutral expectations are in the right-hand side panel.

macroeconomic policy decisions, trading strategies, or risk-management. In these forecasting applications the interest often lies in forecasting the whole distribution of future yields, in particular, quantiles of the future yields. Therefore, we next focus on these distributional or density forecasts.

An appealing feature of the Hull-White model is that thanks to its conditional normality the density forecasts are available in closed form. Specifically, the real-world measure q-quantile of the forecasted yield is given by

$$\hat{y}_t^{\mathbb{P}}(t+h, t+h+\tau) + \sqrt{\text{var}\Big[y_t(t+h, t+\tau+h)\Big]} \times u_q, \tag{1.43}$$

where u_q is the q-quantile of the standard normal distribution, and $\text{var}\Big[y_t(t+h, t+\tau+h)\Big]$ is the variance of the out-of-sample forecast. It follows from the Girsanov transformations of measures that the forecast variance remains the same under the measure change, and therefore the risk-neutral density forecast is obtained by replacing $\hat{y}_t^{\mathbb{P}}$ with $\hat{y}_t^{\mathbb{Q}}$ in (1.43). The variance of the out-of-sample forecast is based on the postcalibration equation (1.9). Specifically, the variance is given by

$$\text{var}\Big[y_t(t+h, t+h+\tau)\Big] = b(\tau)^2 \frac{\sigma^2}{2\alpha}\left(1 - e^{-2\alpha h}\right) + \frac{h}{\Delta t}(\mathbf{W}\hat{\mathbf{\Sigma}}\mathbf{W}^\top)_{i,i},$$

where the first term on the right-hand side comes from the variance of the short-rate process, the second term is the variance of the postcalibration error, which is based on our assumption of a serially uncorrelated postcalibration process. The variance of the postcalibration error that corresponds to the τ-period yield is selected by the index i, i in $\mathbf{W}\hat{\mathbf{\Sigma}}\mathbf{W}^\top$. The factor $h/\Delta t$ scales the the postcalibration error variance to adjust the possibly different length of the forecast horizon, h, and sampling frequency, Δt, used in estimation.

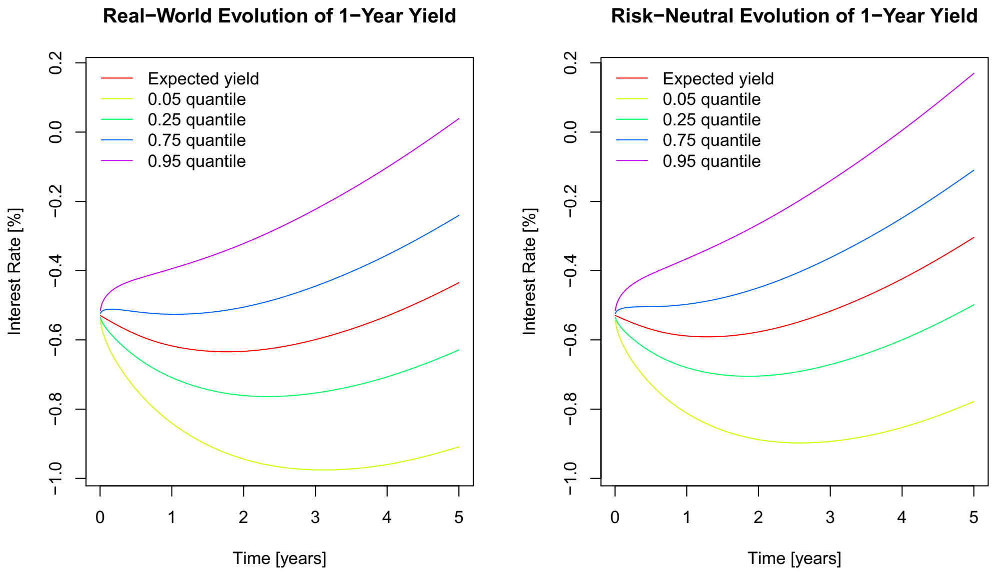

Figure 1.6: Density forecast on 30 August 2019 of one-year yield. The real-world density is provided at the left-hand side panel, while the risk-neutral density is given at the right-hand side panel.

Density forecasts of the one-year yield for the next five years are provided in Figure 1.6. As in Figure 1.5, we can observe that the real-world density forecasts in the left-hand side panel are lower than the risk-neutral forecasts in the right-hand side panel. The density forecasts are fundamental in risk management [Hull et al., 2014]. In Section 1.4.8, we illustrate the impact of the measure choice and the use of quantile forecasts for credit risk exposure calculation.

1.4.6 Out-of-Sample Forecasting Analysis

In this section we examine the out-of-sample forecasting performance of the Hull-White model. We use monthly observations of Euribor- and swap-based yields (the first business day of a month), and our sample spans over twenty years from January 2001 until January 2021. We use six years of data (72 time periods) to estimate the Hull-White model and generate one-month ahead forecasts. Thus, the first yield curve forecast is generated in January 2007 for February 2007. We then recursively reestimate the model by using six years of monthly data and generate one-month ahead forecasts until December 2020. In total, we obtain 168 (14 years) of one-month ahead yield curve forecasts.

We consider out-of-sample forecasts generated by the Hull-White model under the real-world measure ($HW^{\mathbb{P}}$), the Hull-White model under the risk-neutral measure ($HW^{\mathbb{Q}}$), and for a comparison also the random walk model (RW). The forecasting performance is evaluated by considering forecast errors. The forecast error is defined as the difference between the actual and forecasted value. For instance, for the $HW^{\mathbb{P}}$ model the time $t+h$ forecast error of a τ-period yield is given by $y(t+h, t+h+\tau) - \hat{y}_t^{\mathbb{P}}(t+h, t+h+\tau)$, where the model out-of-sample forecast generated at time t is given by (1.41) and where we set $h = 1/12$ as the forecast horizon is one month. The random walk forecast is $\hat{y}_t^{\text{RW}}(t+h, t+h+\tau) = y(t, t+\tau)$, that is, the forecast is naive as it always predicts no change of the yield.

τ	RMSE			Mean Error		
	$HW^{\mathbb{P}}$	$HW^{\mathbb{Q}}$	RW	$HW^{\mathbb{P}}$	$HW^{\mathbb{Q}}$	RW
1M	**0.154**	*0.163*	0.159	**0.017**	*0.064*	0.026
2M	**0.151**	*0.159*	0.155	**0.017**	*0.063*	0.026
3M	**0.148**	*0.156*	0.152	**0.017**	*0.063*	0.026
6M	**0.141**	*0.149*	0.145	**0.015**	*0.060*	0.026
1Y	0.140	*0.147*	**0.143**	**0.011**	*0.056*	0.027
2Y	0.156	*0.161*	**0.155**	**0.009**	*0.052*	0.027
3Y	0.163	*0.168*	**0.162**	**0.009**	*0.050*	0.027
4Y	0.163	*0.169*	**0.162**	**0.010**	*0.050*	0.027
5Y	0.162	*0.168*	**0.161**	**0.011**	*0.050*	0.027
6Y	0.162	*0.168*	**0.161**	**0.011**	*0.049*	0.027
7Y	**0.162**	*0.168*	**0.162**	**0.011**	*0.048*	0.027
10Y	**0.168**	*0.173*	**0.168**	**0.011**	*0.045*	0.026
15Y	**0.179**	*0.182*	0.180	**0.010**	*0.038*	0.026

Table 1.3: Out-of-sample one-month ahead forecasting results. In total forecasts for 168 months have been generated. The forecast precision is measured by the root mean squared error (RMSE), the forecast bias is summarised by the mean error (ME). Both metrics are measured in percent. Three models have been considered, the Hull-White model under the real-world measure ($HW^{\mathbb{P}}$), Hull-White model under the risk-neutral measure ($HW^{\mathbb{Q}}$), and the random walk model (RW).

The forecasting performance of the three models is summarised in Table 1.3. We report the root mean squared error (RMSE), and the mean error (ME) of the forecast errors for the 13 tenors that were used to form the likelihood function. The most favourable outcomes are in bold, the least favourable in italics. First note that the Hull-White model under the risk-neutral measure performs the worst for all tenors in terms of both the RMSE and ME. This is not surprising since the risk-neutral measure does not involve the market price of risk, and it is developed for arbitrage free pricing. The absence of the market price of risk is clearly manifested in the forecast ME which magnitude is 4.5 times higher for the $HW^{\mathbb{Q}}$ model compared to the $HW^{\mathbb{P}}$ model. This means that in average the yield curve expectations from the $HW^{\mathbb{Q}}$ understate the actual yields more than expectations from the $HW^{\mathbb{P}}$ model. This stresses the importance of estimating the model under the real-world measure for forecasting and risk management applications.

We next turn to the comparison of the $HW^{\mathbb{P}}$ model with the random walk model. The $HW^{\mathbb{P}}$ model dominates the RW model in terms of the forecast mean error, which magnitude is in average 2.3 times smaller for the $HW^{\mathbb{P}}$ model. In terms of the RMSE the two models perform comparably, the largest difference is 0.004% in favor of the $HW^{\mathbb{P}}$ model for tenors 1M, 2M, 3M, and 6M. We consider the forecasting results encouraging as it has been recognised that beating the RW yield curve forecasts is difficult, in particular on one-month ahead forecast horizon (Diebold and Li [2006], Nyholm and Vidova-Koleva [2012]). To the best of our knowledge forecasting performance of no-arbitrage models have not been studied in the literature, which may be attributed to the fact that these models have not been really considered under the real-world measure. It would be of interest to see how no-arbitrage models perform in this respect.

1.4.7 Comparison with Calibration to Swaptions

In this section we compare maximum likelihood parameter estimates to parameter values obtained by calibration methods. Calibration is the traditional way of bringing the Hull-White model to market data. Calibration methods are not based on historical time-series data, but rather consider yields and prices of vanilla interest rate derivatives that are quoted on the market at the calibration time t. Since calibration methods only use cross-sectional data from a single point of time, they only allows for identification of α and σ. In other words, the Hull-White model is only identified under the risk-neutral measure $\mathbb{Q}$ as the market price of risk λ cannot be retrieved. Thus, the typical use of a calibrated model is pricing and hedging of non-vanilla derivatives.

The current market practice is to calibrate the Hull-White model to prices of swaptions [Brigo and Mercurio, 2001, Section 3.14]. Swaption prices are available in the form of an implied volatility matrix with one dimension being the maturity of the swaption and the other dimension being the tenor of the underlying swap. A calibration method is based on numerical minimisation of relative differences between the market and model swaption prices. We use the typical objective

$$\arg\min_{\alpha,\sigma} \sum_{i=1}^{N} \left(\frac{\mathrm{Swpt}_i - \mathrm{Swpt}_i^M}{\mathrm{Swpt}_i^M} \right)^2, \tag{1.44}$$

where Swpt_i^M is the swaption price on the market, Swpt_i is the Hull-White model swaption price, and N is the number of calibrated swaptions (we use $N = 10$, $N = 24$, and $N = 170$ in what follows).

In contrast to a zero-coupon bond, the swaption pricing function is not known in a closed form for the Hull-White model. A popular approach to calculate a swaption price is based on Jamshidian's decomposition [Jamshidian, 1989], which requires numerical root finding. This can be particularly time consuming when the model is calibrated to market data and the model prices need to be recalculated many times in order to achieve the objective (1.44). Russo and Fabozzi [2016] suggest an alternative swaption pricing approach based on the concept of stochastic duration. They derive an approximative swaption pricing function in a semi-closed form (up to a straightforward numerical integration). Their pricing method is thus faster than using Jamshidian's decomposition.

Russo and Torri [2019] study calibration of the Hull-White model on the EUR interest rate market. They use both Russo and Fabozzi's and Jamshidian's approaches to swaption pricing. They calibrate the Hull-White model by using EUR swaption prices and Euribor- and swap-based yields from the last business day of 2011, 2012, 2013, 2014, 2015 and 2016 (they use data provided by Bloomberg). The calibrated parameter values are reported in Table 1.4 with α in the top panel and σ in the bottom panel. The second column (JA) reports calibration results based on the Jamshidian's approach, whereas the third column (RF) reports results based on the swaption pricing approach of Russo and Fabozzi [2016]. These results have been taken from Table 1[9] in Russo and Torri [2019], and are based

[9] Parameter values have been rounded to three decimal digits and σ values were multiplied

alpha

	JA	RF	RF170	RF24	ML I	ML II
2011	0.141	0.130	0.023	0.017	0.058	0.045
2012	0.080	0.117	0.001	0.012	0.040	0.038
2013	0.105	0.198	0.005	0.000	0.056	0.032
2014	0.032	0.004	0.000	0.000	0.051	0.036
2015	0.022	0.001	0.000	0.000	0.059	0.039
2016	0.030	0.001	0.000	0.004	0.056	0.027

sigma

	JA	RF	RF170	RF24	ML I	ML II
2011	1.810	1.550	0.873	0.862	0.475	0.606
2012	1.150	1.200	0.577	0.655	0.389	0.627
2013	1.420	1.850	0.654	0.605	0.367	0.585
2014	0.760	0.570	0.419	0.481	0.338	0.428
2015	0.800	0.650	0.467	0.453	0.304	0.394
2016	0.850	0.650	0.371	0.318	0.221	0.255

Table 1.4: Calibration and estimation results of the Hull-White model for the last business days of years 2011–2016. The calibration results are reported in columns two to five. The second (JA) and third column (RF) report calibration results based on 10-year co-maturity swaptions, the fourth column (RF170) reports results based on the whole swaption volatility matrix, and the fifth column (RF24) is based on a sparse version of the volatility matrix. The sixth (ML I) and seventh (ML II) column report our maximum likelihood estimation results. ML I provides estimates based on using both Euribor- and swap-based yields in the likelihood function, whereas ML II reports estimates based on using swap-based yields only.

on at-the-money swaptions with co-terminal maturity of 10 years. A co-terminal maturity is the sum of the swaption maturity and swap tenor. In the literature, it is common to study swaptions of a certain co-terminal maturity as they are related to pricing and hedging of Bermudan swaptions, which is a frequent use of no-arbitrage models. As noted by Russo and Torri [2019] using “co-terminal maturity may lead to sub-optimal calibration results due to inefficient use of data” (their choice of co-terminal maturity of 10 years amounts to use 10 out of 170 available swaption price quotes). We therefore calibrate the Hull-White model by using the Russo and Fabozzi [2016] pricing method to all 170 elements of the swaption volatility matrix. In addition, for robustness considerations, we calibrate to 24 swaption prices, which were selected by applying a sparse but evenly spaced grid on the volatility matrix. The calibrated parameter values based on all 170 swaption price quotes are reported in the fourth column (RF170) of Table 1.4, whereas the results obtained by using the sparse grid are given in the fifth column (RF24).

In order to compare our maximum likelihood approach to calibration, we esti-

by 100 to align with reporting used in this chapter, that is, parameter values rounded to three decimal digits and interest rates measured in percent.

mate the Hull-White model by using two years of daily data until the calibration time, that is, we follow the same procedure as in Section 1.4.2.[10] We report two sets of results. First, in the sixth column (ML I) of Table 1.4, we report α and σ estimates based on using both Euribor- and swap-based yields (1M, 2M, 3M, 6M, 1Y-7Y, 10Y, 15Y) in the likelihood function (as we did in Section 1.4.2). Second, in the seventh column (ML II), we report estimates based on using only the swap-based yields (1Y-7Y, 10Y, 15Y) in the likelihood function. The reason for using only swap-based yields is that a swaption is an option written on a swap, and hence its the swaption price is governed by the risk-neutral dynamics of the swap during the swaption maturity.

We can observe that calibrated values of α vary quite substantially across years, with respect to the swaption pricing method used in calibration (for instance, the ratio of JA to RF α values varies from 0.53 in 2013 to 60.80 in 2016), and also with respect to the selection of swaptions used for calibration (RF170 and RR24 α values are always smaller than α values based on co-terminal swaptions). The parameter α determines the speed of mean reversion to the forward curve, and thus can be problematic to estimate from cross-sectional data. Indeed, Russo and Torri [2019] point out that some authors propose to estimate α from historical data rather than calibrate it. The maximum likelihood estimates of α do not exhibit any abrupt changes across years. The ML II α estimates are always smaller than ML I α estimates.

The calibration of σ appears more stable than calibration of α (for example, the ratio of JA to RF σ values varies from 0.77 in 2013 to 1.33 in 2014). The stability is comforting as volatility is the key input to derivatives pricing. We can observe that the RF170 and RF24 σ values, which are based on the whole space of swaption prices are always smaller than the JA or RF σ values, which are based on swaption prices with co-terminal maturity of 10 years only. The ML II σ estimates are larger than ML I σ estimates for every date considered, and they match the calibrated RF170 or RF24 σ values reasonably well. Specifically, ML II σ estimates are in average just about 13% smaller than RF170 calibrated σ values (the ratio of ML II to RF170 σ values varies from 1.09 in 2012 to 0.69 in 2011 and 2016). These results suggest that our maximum likelihood estimation is a viable method for using the Hull-White model for interest rate derivatives pricing on markets where swaption price quotes are lacking.[11]

1.4.8 Application in Risk Management

In this section we use the estimated Hull-White model in the probability of default calculation, and illustrate the importance of the measure choice. We use the one-year yield density forecasts from Section 1.4.5 to calculate the probability of default of a loan. The creditor uses a proprietary logistic regression model and data to estimate at time t the one-year probability of default PD_t. For our illustrative purposes we consider the following simplified version of the creditor's

[10]We have compared our EUR yield curve data to yield curve data provided of Russo and Torri [2019]. The date sets are almost identical for all six trading dates considered. The largest total absolute difference between the yield curves is 0.53 basis points in 2013.

[11]Expert knowledge may be requested to tweak the estimated parameters, in particular σ, to reflect specifics of the priced derivative contract.

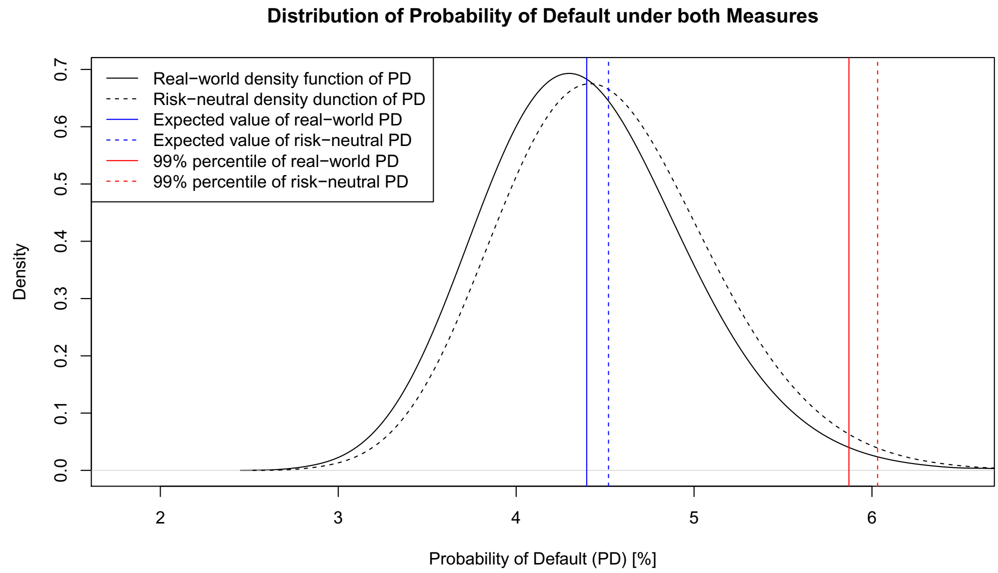

Figure 1.7: Distribution of the probability od default PD_t under both measures.

model:

$$\mathrm{PD}_t = \frac{\exp\{\hat{\beta}_0 + \hat{\beta}_1 \Delta y\}}{1 + \exp\{\hat{\beta}_0 + \hat{\beta}_1 \Delta y\}}, \qquad \hat{\beta}_0 = -3, \quad \hat{\beta}_1 = 1, \tag{1.45}$$

where $\Delta y := y(t+1, t+2) - y(t, t+1)$ is the change of one-year market yield over the one-year horizon and $\hat{\beta}_0, \hat{\beta}_1$ are regression coefficient estimates. Note that PD_t is a random variable as it is a function of Δy. The sensitivity of the probability of default to interest rates has been well established in credit risk modelling (Lando [2009]). The sensitivity of PD_t to other risk factors is in our simplified model averaged out by the β_0 coefficient. The β_0 estimate implies that $\mathrm{PD}_t = 4.7\%$ if the market yield did not change ($\Delta y = 0$).

We use the Hull-White model density forecasts presented in Figure 1.6 to evaluate the PD_t on 30 August 2019. The distribution of PD_t is captured in Figure 1.7. Since the probability of default is dependent on the real-world interest rate trajectory, risk calculations should be done under the real-world measure. We can observe that the risk-neutral distribution of PD_t overstates the default risk. Indeed, the real-world expected $\mathrm{PD}_t^{\mathbb{P}} = 4.39\%$, whereas the risk-neutral expected $\mathrm{PD}_t^{\mathbb{Q}} = 4.52\%$. The probability of default is given by a nonlinear transformation of the yield change, and therefore the difference between $\mathrm{PD}_t^{\mathbb{P}}$ and $\mathrm{PD}_t^{\mathbb{Q}}$ values is not constant. For instance, credit risk managers often consider the 99% percentile of the default probability; at this percentile we obtain that $\mathrm{PD}_t^{\mathbb{P}} = 5.87\%$ and $\mathrm{PD}_t^{\mathbb{Q}} = 6.03\%$. In our toy model the credit risk is overstated under the risk-neutral measure. However, in many risk-management applications the opposite applies, i.e., the risk is understated under the $\mathbb{Q}$ measure. For example, consider a portfolio with long position in bonds. As bond prices move in the opposite direction to yields, the Value at Risk of such a portfolio would be understated under the risk-neutral measure.

2. An Asset - Liability Management Stochastic Program with a Pricing Decision

The second part of this thesis discusses a decision-dependent randomness asset-liability management model. There, we formulate a model for a life-cycle of a loan which is provided by a company to an individual customer. Moreover, one of the decisions which the company has to make is to offer an interest rate to the customer. The customer's decision whether to accept or reject the loan directly depends on the offer from the company and hence it induces decision-dependent randomness to the program. The model formulation as well as the results and sensitivity analysis have been published in paper Kopa and Rusý [2021b].

First, we present the model formulation in detail in Section 2.1. We describe the general settings and the objective function in Section 2.1.1. Next, we continue in describing the individual stochastic components of the model in Section 2.1.2. In Section 2.1.3, we introduce the constraints and present the entire formulation of the model. Thereafter, in Section 2.2, we show the results of the model. We start in Section 2.2.1 where we show the results of the program for single parameters' setting to illustrate the optimal decisions and how the loan value would evolve depending on the company decisions. Furthermore, in Section 2.2.2 we discuss the effect of the customer's properties on the model solution, especially on the offered interest rate and the expected value of the loan. There, we also mention the losses which are incurred by the company if it does not behave in the optimal way.

2.1 Model Formulation

2.1.1 Model Environment and Objective

First, let us describe the situation and the time-frame on which the optimization model is built. We consider three economic agents:

- A credit institution operating in the interbank market,
- A company, or lending entity within this problem framework, that will borrow money in the interbank market and give loans to consumers
- Customers, or borrowing agents, who through the loan will be in condition to finalize the purchase of a good.

We illustrate the relationship of these three agents in Figure 2.1. The individual customer approaches the company and asks for a loan. On the other hand, the company borrows money from the financial market to obtain sufficient funds.

The primary objective of this chapter is to construct and analyse an optimization model (often called a program) which a company can use to determine its decisions. These decisions shall be taken so the company maximises the expected

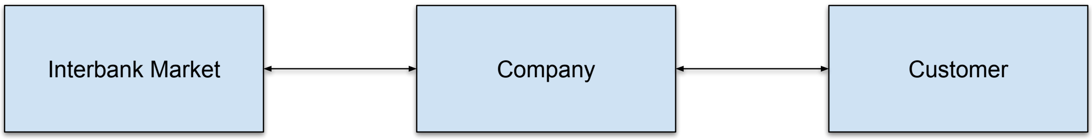

Figure 2.1: Economic agents which enter the optimization model.

value of the loan (the expected profit) at the agreed loan maturity. More specifically, at every point in time, we define the term *value of the loan* as the sum of cash and present value of assets and liabilities resulting from the loan contract with the customer (see, later, equation (2.14)). Consequently, all the loans from the market which are used to finance the consumer loan are also included. The value of the loan changes over time. This depends on the market interest rate evolution and the customer's behaviour (possible loan default or prepayment). These factors are, however, not known a-priori, so they are treated as random.

Next, let us highlight where the main difficulty in the asset-liability stochastic programming model arises. Usually, random effects are assumed to be exogenous. In other words, no decisions taken by the decision maker in the optimization model affect the uncertain elements entering the model. However, this is not the case in this situation. Here, the company's initial decision is to offer the customer a loan with a fixed interest rate. This decision directly affects the probability of the customer accepting the loan offer. This is a random event (from the company's point of view), and it is endogenous — dependent on the initial decision. The probability of the customer prepaying or defaulting also depends on the offered interest rate. The strength of this relationship and how this endogenous uncertainty is dealt with is discussed in detail in Section 2.1.2.

In such loan operations, a common practice is for the lending agencies to borrow in the market so that they can match asset and liability cash flows from the start of the loan. However, Rusý and Kopa [2018] showed that such a strategy is not optimal as there are alternative strategies which have better risk properties and higher profitability. For this reason, we give numerous options to the company on how to form and optimize its liability side. The fact that we jointly optimize the initial interest rate decision and the company's borrowings allows us to consider the different probabilities of cash flows that stem from the decision-dependent randomness and adjust the company's borrowing strategy accordingly. This joint optimization of the pricing problem and the consequent asset-liability problem is a key feature of the model and the main contribution of this chapter. It provides us with the optimal decisions corresponding to the genuine nature of the problem.

In this work, we restrict ourselves to the simplest loan type — a non-collateralized consumer loan with fixed maturity and fixed interest rate. However, this framework can easily be extended to a variety of other problems which combine pricing decision of some good together with subsequent action depending on the demand for the good. There, we have a decision-dependent randomness between the price and the demand. The framework of stochastic programming is then flexible enough to take into account numerous features which are connected to special properties of the good. From the financial perspective, other products which could be modelled are, for example, mortgages or pension/building savings.

There, the structure of the resulting actions and cash flows is more complex, but the general idea is the same.

We assume that a customer comes to the company and wants to borrow N_0 amount of money for a period of T months.[1] Afterwards, the company offers the customer an interest rate r for such a loan. We assume this is the only cost the company charges the customer. Next, the customer decides whether to accept or reject the offer. If the loan is agreed upon, we proceed further to model the life of the consumer loan, which is repaid regularly each month $t, t = 1, \ldots, T$ by equal instalments.

The multi-stage stochastic optimization program will, however, consist only from $K + 1 \leq T + 1, K \in \mathbb{N}$ stages, $0 = t_0 < \ldots < t_K = T$. At these times, scenarios of other random quantities (interest rate evolution, customer's loan default or prepayment) will be observed and the company will be able to make decisions. In other times, no decisions are made; only instalments are paid. The decisions will define cash flows between the company and the market and form the liability side of the company. We use time index $t, t \in \{0, \ldots, T\}$ to denote months within the duration of the loan; index $k, k \in \{0, \ldots, K\}$ and times t_k denote decision stages of the optimization program. We have $\{t_k, k = 0, \ldots, K\} \subset \{0, \ldots, T\}$. In some cases, indices i, j are also used to iterate over the set of decision times.

2.1.2 Random Elements, Scenarios and Decision-Dependent Randomness

In this section, we will describe four random elements which are part of the life of the loan in the model. They are represented by the following questions:

- Will the customer accept the loan?
- If yes, will he afterwards prepay the loan?
- Will he default on the loan?
- What will be the evolution of market interest rates?

The first three elements describe uncertain customer behaviour. They are all considered endogenous as they depend on the interest rate decision. The fourth element, which is considered exogenous, captures the evolution of prices in the financial market.

Probability of Accepting the Loan

Once the customer is offered a loan with a specific interest rate, there is the possibility that he will either accept or reject it. The probability that the customer accepts the loan offer and enters into a contract with interest rate r is denoted by $p(r)$. This function is customer-specific. Here, multiple customer-dependent factors can play a role, for example the customer's knowledge of market conditions

[1] Note that usually the principal N_0 is set by the company. However, as it enters the model only as a *scale* parameter (multiplier of the objective function), we treat it as fixed — for example, determined by a risk management unit of the company.

or alternative offers from other market participants. We assume that the relationship $p(r)$ is estimated by logistic regression, where one of the regressors is the interest rate offered to the customer. However, in general, any functional form which describes the desired relationship could be used. We employ a function:

$$p(r) = \frac{\exp\{b_1(b_0 - r)\}}{1 + \exp\{b_1(b_0 - r)\}} = \frac{1}{1 + \exp\{-b_1(b_0 - r)\}}, \tag{2.1}$$

where b_0 and b_1 are [customer-dependent] parameters. Parameter b_0 represents the rate at which the customer is indifferent to accepting or rejecting the loan. We will call this *midrate* and it holds that $p(b_0) = 0.5$. The second parameter, b_1, expresses the customer's sensitivity to interest rates as it captures the effect on the probability of accepting the loan when we deviate from midrate by a certain amount. This can be quantified exactly by the usual interpretation of logistic regression models. When $b_1 = 100$, an increase of 1% in the offered rate r implies a decrease in the odds ratio (accept/reject) by a factor of $e^{-b_1/100} = e^{-1}$.

For a given client, one can estimate the value of midrate and sensitivity by a simple logistic regression. Consider a dataset where the response is a successful or unsuccessful offer while regressors are the offered interest rate and customer's properties. Then, a model can be formulated, such that the first order terms of customer's properties define the midrate and covariate interactions with the interest rate specify the sensitivity relationship. The fitted model together with covariates of new customer would generate estimates of parameters b_0 and b_1.

This random event (accepting/rejecting the loan) is realized before issuing the loan. If the customer's decision is to reject, no cash-flows take place and the value of the loan is 0. On the other hand, if the decision is to accept, then the loan is issued, the company makes an initial decision on how to borrow at the market and cash-flows are exchanged.

The quantity $p(r)$ captures the essence of endogeneity of this random event. It links the company's decision to the random variable's distribution.

Interest Rate Evolution

Next, we introduce market interest rates which express the cost of money at the market. We denote y_t^τ to be the annualized, risk-free interest rate with time-to-maturity τ at time t. For $t > 0$, this quantity is random. It is considered exogenous because the company is not thought to be able to affect its evolution. We also denote m_t^τ to be the rates for which the company can borrow from a market participant — a bank. We define: $m_t^\tau = y_t^\tau + m(\tau)$.

Quantity $m(\tau)$ represents the spread between the risk-free rates and the rates which the bank charges the company. This is fixed over time and can be interpreted as the mark-up of the bank. In other words, we assume that the company has a contract with the bank regarding floating rate borrowing. In the numerical part, the values of mark-up $m(\tau)$ were set as follows. We defined $m(0) = 0.0048, m(2) = 0.0096$ and $m(5) = 0.0132$, where τ is in years. Values in between were obtained by linear interpolation.

The risk-free rates y_t^τ are modelled by the Hull-White model formulated by Hull and White [1990a], which belongs to the class of one-factor short-rate models introduced by Vasicek [1977]. It is defined by the following stochastic partial differential equation:

$$dr_t = (\theta(t) - \alpha r_t)dt + \sigma dW_t,$$

where r_t stands for short-rate and W_t is the standard Brownian motion. Parameter α stands for the mean reversion factor and σ summarizes the volatility of the short-rate. Finally, $\theta(t)$ is set such that the observed market prices are fitted perfectly, i.e:

$$\theta(t) = \frac{\partial f^M(0,u)}{\partial u}\bigg|_{u=t} + \alpha f^M(0,t) + \frac{\sigma^2}{2\alpha}\left(1 - e^{-2\alpha t}\right),$$

where $f^M(0,t)$ is the market instantaneous forward rate at time 0 for time t. We employ the usual starting condition $r_0 = f^M(0,0)$ and calculate the yields y_t^τ via the formulas for zero-coupon bond prices [Brigo and Mercurio, 2001]. Thanks to the fact that the Hull-White model uses exogenous information in the form of the observed market yield curve, predictions of yields based on this model are close to market expectations. The calibration of the model's parameters was inspired by Chen and Scott [1993] who estimated the Cox–Ingersoll–Ross model of Cox et al. [1985a] by the maximum likelihood method on observed yields. It is in detail described in Kladívko and Rusý [2021] or in Section 1.2 of this thesis. For this estimation, we used the daily PRIBOR rates observed on the Czech market from 28th June 2015 to 1st March 2018. The estimated values of the parameters were $\hat{\alpha} = 0.1346, \hat{\sigma} = 0.006427$. Such an estimation procedure is built on numerous model assumptions. The normality of short-rate movements and their autocorrelation require particular attention. From *acf* and *pacf* plots, we concluded that no autocorrelation is present, which is a consequence of the fact that the Hull-White model uses the observed market curve for future predictions. For testing normality, we used the Shapiro-Wilk test, which gave us a p-value of 0.00545. At the usual significance level, we would reject the normality hypothesis. This is mainly due to 4 outliers, which correspond to jumps that are a little larger than one would expect in normal distributions. However, because prediction based on this model looked reasonable, we decided to accept and use it in the scenario generation procedure.

In our model, we capture the interest rate evolution in the form of a regular scenario tree, often seen in multi-stage stochastic programming, such that in each stage every node has the same number of successors. Moreover, in such a tree, all scenarios are equiprobable. We will denote $S_k, k \in \{0, \dots, K\}$ as the set of nodes of the interest-rate tree in a decision stage t_k and $a_i(s_k)$ as the time t_i ancestor of a node $s_k \in S_k, 0 \leq i < k$. We will also denote $y_{t_k}^\tau(s_k)$ as the risk-free interest rate and $m_{t_k}^\tau(s_k) = y_{t_k}^\tau(s_k) + m(\tau)$ as the rates for the company at time t_k with time to maturity τ in the scenario node $s_k \in S_k$.

The scenarios of the short-rate were chosen to be the quantiles of the model-implied distribution (conditioned on the observed value of the short-rate in the ancestor node). We derived the term structure $y_{t_k}^\tau(s_k)$ and discount factors $P(t_k, t_k + \tau; s_k), \tau > 0$ from the short-rate based on well-known formulas for zero-coupon bond prices implied by the Hull-White model (see, for example, Brigo and Mercurio [2001] for more details).

Probability of Default and Prepayment

The final two types of randomness which enter the model are loan prepayments and customer defaults. We treat these effects as endogenous, as they are closely connected to the offered rate of the loan. This, however, introduces another decision-dependent randomness into the model, which is present in all stages of the multi-stage program. First, let us describe why we need to take these effects into account in the model.

Customer defaults are generally considered to be the biggest risk factor affecting the profitability of a loan. This is simply because it can happen that the customer becomes unable to fulfil his commitments. Under such an event, the company loses not only the interest rate charged to the customer but also a part of its principal. Such a proportion is called *loss given default* and we include it in the model as a fixed parameter, denoted as lgd and set to a value of 0.5. If the customer defaults at time t_k, the company is modelled to receive $(1 - \text{lgd})$ times the remaining principal as the recovered part of the loan. Scenarios of customer defaults will be added to the model. Thereafter, we assign them probabilities so they match our initial assumption about hazard rates $h(t_k, r)$ — the probability of default at time t_k given that the loan with interest rate r survived up to time t_{k-1}.

Prepayment means that the customer repays more money than it was scheduled in the original contract. Usually, such a prepayment comes from one of the two following reasons:

- The customer has spare money which he can afford to use for loan prepayment.
- The customer finds a cheaper loan and he refinances it.

The first reason is unconnected to the decision variables or random quantities in the problem. On the other hand, the second reason is closely related to the interest rate of the loan. Naturally, if the price of the loan is too high, the customer is more likely to look for cheaper options at the market and, thus, more likely to repay the loan earlier. Therefore, this random effect is also endogenous. Similar to customer defaults, we add scenarios of loan prepayments into the model and assign them probabilities to match our assumptions about the hazard rate $g(t_k, r)$ of prepayment at time t_k of the loan with interest rate r. By loan prepayment, we mean only full prepayment of the loan, so when it occurs, the remaining principal is repaid to the company.

Next, let us describe how we determine the hazard rates for default $h(t_k, r)$ and prepayment $g(t_k, r)$. We formulate both the default and the prepayment model on the following ideas:

- The probability depends on the interest rate, time to maturity of the loan and the initial rating of the customer.
- There might be an interaction between the interest rate and the rating of the customer.

From there, we formulate a logistic regression model:

$$\text{prob}(t_k, r, \rho) = \frac{\exp(\eta)}{1 + \exp(\eta)}, \qquad \eta = \beta_0 + \beta_1 r + \beta_2 \rho + \beta_3 t_k + \beta_4 \rho \cdot r, \tag{2.2}$$

where $\rho = \{1, 2, 3, 4\}$ denotes rating of the customer[2] and $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4$ are parameters. Symbol $\text{prob}(t_k, r, \rho)$ stands for the probability that a loan with interest rate r of a customer with rating ρ will be defaulted/prepaid by the time t_k, given that it survived till time t_{k-1}.

We fitted the model to real market data of a Czech company operating in the industry and tested whether models in (2.2) could be used to capture the relationship. The data consisted of all [17 554] consumer loans of the company with maturity between 5 and 6 years active at one point in time. We had the initial rating of the customer (which was valid when the loan was agreed) and we observed whether the loans were repaid or defaulted in the following year. The observed and fitted default and prepayment probabilities are summarized graphically in Figure 2.2, while the estimated parameters are shown in Table 2.1. For better readability and interpretation, values of interest rate r are thought to be in percent. To obtain values of $g(t_k, r)$ and $h(t_k, r)$, one needs to use a given value of rating ρ. The corresponding intercept and coefficients for interest rate and time will be denoted $\beta_0^g, \beta_1^g, \beta_3^g$ and $\beta_0^h, \beta_1^h, \beta_3^h$ respectively.

[2]We should note that we considered four different ratings of a customer, from the best rating (1) up to the worst rating (4). A few customers with a rating worse than 4 on the usual scale $1 - 8$ were assigned rating 4 for this analysis.

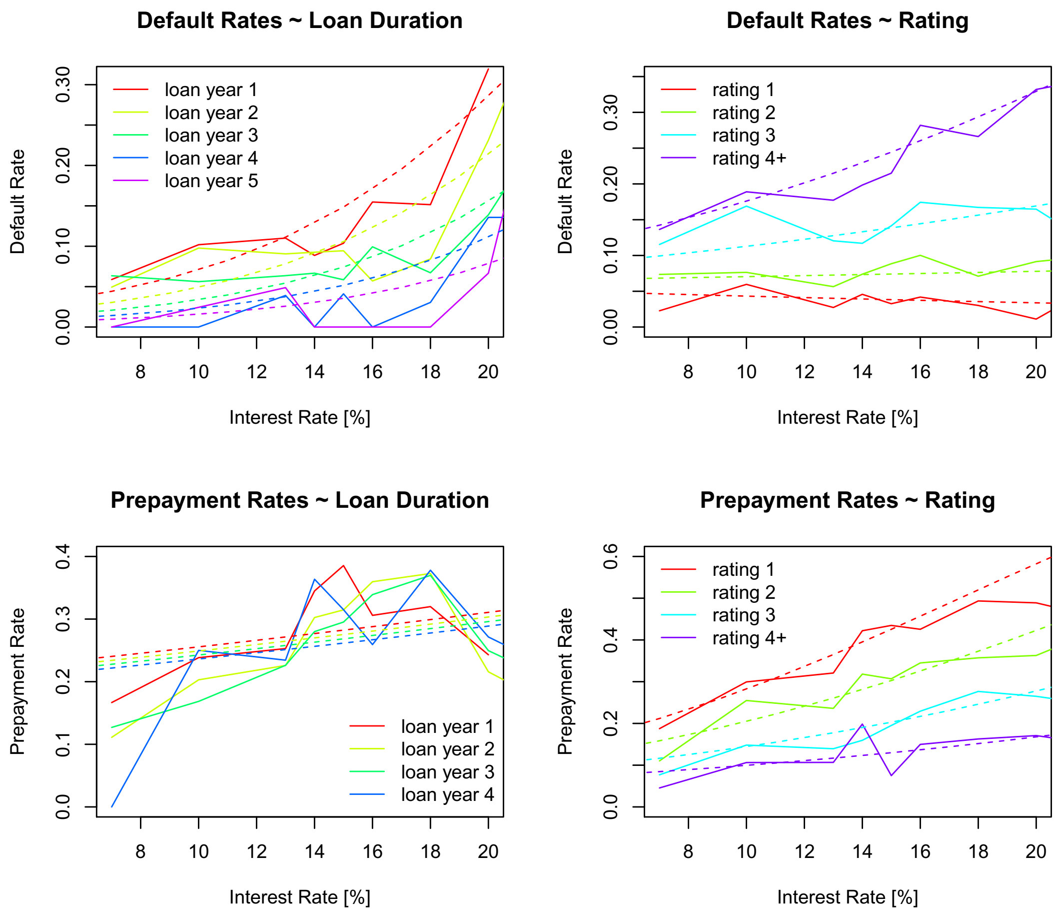

Figure 2.2: Prepayment and default rates and fitted probabilities in analysed dataset. Dashed lines show the fitted logistic model.

Let us briefly comment on the interpretation of this model. The coefficient estimates are very much in-line with our expectations. The parameter estimates show that defaults occur in early stages of the loan duration and that there is a strong relationship between interest rate and default for customers with bad rating. On the other hand, for prepayment, we can see that customers with good rating have high prepayment rate for high interest rate loans. This interaction also has reasonable interpretation. In some of the plots, the fitted values seem to be far from the observed quantities. This is partly due to few observations in these areas (for example, low- and high-interest rate). Finally, we would like to stress that this analysis was performed in order to get a reasonable (real) estimate of the relationship between interest rates and default/prepayment probabilities. We could have described the econometric relationship in more detail. However, this will increase model complexity and we might lose its computational tractability.

Now, let us discuss how we implement the scenarios of loan prepayments and customer defaults into the program. First, we already have $|S_K|$ interest rate scenarios from the initial decision period to the maturity of the loan. Take one scenario as fixed and on every node/stage, loan prepayment or customer default can occur. Note that during the life of the loan, only one customer default or loan prepayment can take place. Hence, each program scenario can be defined as a pairing of the interest rate scenario and the event specification. Events are formulated to be loan prepayments or customer defaults at any stage $t_k \neq t_0$. For example, default at stage t_2 is considered to be one event. This implies, that we have in total $2 \cdot K$ events. We should also mention that loan prepayment at time $t_K = T$ corresponds to the loan being repaid at the initially agreed time.

Mathematically speaking, we define set E as a set of all possible events. For event $e \in E$, we define $t(e)$ to be the time when the event occurs, $d(e)$ an indicator function which is 1 when the event is customer default and 0 when the event is full prepayment. So, for example, for the event that the customer defaults at stage t_2, we have $t(e) = t_2, d(e) = 1$. These functions are important for determining scenario probabilities as in (2.3) (calculated from conditional probabilities of default and prepayment) and for the specification of the budget equation given in (2.11). The scenario of a program is then uniquely defined by a pair $(s, e), s \in S_K, e \in E$.

What is important to realize under this parametrisation is that at time t_k, we cannot distinguish between two program scenarios $(s, e_1), (s, e_2), s \in S_K, e_1, e_2 \in E$ such that $t_k < \min\{t(e_1), t(e_2)\}$. This is simply because by time t_k, we do not observe the nature of event e. Such a property will lead to the inclusion of non-anticipativity constraints into the program.

Model	symbol	R^2	β_0	β_1	β_2	β_3	β_4
Prepayment	$g(t_k, r)$	0.053	-1.93***	0.18***	-0.17*	-0.21***	-0.028***
Default	$h(t_k, r)$	0.106	-2.93***	-0.033	0.20$^{\cdot}$	-0.22 ***	0.031***

Table 2.1: Estimated parameters of models for probability of default and prepayment together with McFadden R^2. Asterisks denote a statistical significance of the coefficient: *** denotes a p-value smaller than 0.001, * smaller than 0.05, and $^{\cdot}$ smaller than 0.1.

The Probabilities of Scenarios

In the final part of this section, we will comment on how we calculate probabilities of scenarios $p(s, e, r), s \in S_K, e \in E$ where r is the initial interest rate decision. These are obtained iteratively by the multiplication of hazard rates in each node — a probability that one reached this node multiplied by the probability that from this node, one moves to its child. For event $e \in E$ and time $t_k \leq t(e)$, we use hazard rate $h(t_k, r)$ if default occurs: $\mathbb{I}_{[t(e)=t_k]} = 1, d(e) = 1$, hazard rate $g(t_k, r)$ if prepayment occurs: $\mathbb{I}_{[t(e)=t_k]} = 1, d(e) = 0$ and finally hazard rate $1 - h(t_k, r) - g(t_k, r)$ if the event is not yet observed: $\mathbb{I}_{[t(e)>t_k]} = 1$. Otherwise, when $\mathbb{I}_{[t(e)<t_k]} = 1$, probabilities are distributed equally depending only on the branching of the interest rate tree given by $|S_{k-1}|/|S_k|$. The formula is as follows:

$$p(s, e, r) = \prod_{k=1}^{K} \Big(\mathbb{I}_{[t(e)=t_k]} d(e) h(t_k, r) + \mathbb{I}_{[t(e)=t_k]} \Big(1 - d(e)\Big) g\Big(t_k, r\Big) \qquad (2.3)$$

$$+ \mathbb{I}_{[t_k<t(e)]} \Big(1 - h(t_k, r) - g(t_k, r)\Big) + \mathbb{I}_{[t_k>t(e)]} \Big) \Big(|S_{k-1}|/|S_k|\Big),$$

$$s \in S_K, e \in E.$$

We require that $\forall k \in 1, \ldots, K$ it holds that $h(t_k, r) \geq 0, g(t_k, r) \geq 0$ and also that $\Big(1 - h(t_k, r) - g(t_k, r)\Big) \geq 0$. Moreover, in the last stage, we must have $g(t_K, r) = 1 - h(t_K, r), \forall s \in S_K$. Under such conditions, the probabilities $p(s, e, r)$ of scenarios are non-negative and sum up to one for all values of interest rate decision r.

The equation (2.3) provides another link between the decision variable r and scenario probability. First, we have a model measuring the effect of the interest rate decision on the probability of moving from each node to its successor as in (2.2). Then, we calculate the probability of each scenario by multiplication of the hazard rates. Here we use Bayesian conditional probabilities and the Bayes' theorem. This captures the effect of the decision-dependent randomness in defaults and prepayments into the multi-stage program.

2.1.3 Stochastic Programming Model Formulation

In this section, we will complete the formulation of the asset-liability stochastic programming model. We have already introduced the interest rate decision r and the probability of accepting the loan $p(r)$. We described the time structure of the model — we have months $t, t \in \{0, \ldots, T\}$, where T is the maturity of the loan, and decision times $t_k, k \in \{0, \ldots, K\}$. See Section 2.1.1 for more details.

Scenarios are defined as a pair $(s, e), s \in S_K, e \in E$, where s captures the interest rate evolution and e the event which realizes on the side of the customer. For a scenario (s, e) and offered interest rate r we have the scenario's probability $p(s, e, r)$, see Section 2.1.2.

In the next part, we will describe the evolution of the principal of the loan (Section 2.1.3) and define how the company can borrow and lend money in the financial market (Section 2.1.3) at the prevailing interest rates. Apart from defining which decisions can be made by the company, we will also derive quantities (such as income at given time and scenario) corresponding to the cash-flows between the company and the market.

In Section 2.1.3, we introduce the current account equation (2.11), which links the company's cash-flows with both the customer and the market. Other constraints, such as the non-anticipativity constraints, are also introduced. All the decisions, quantities and equations are then summarized in Section 2.1.3, where the entire model is presented in the compact form.

Loan Instalment and Principal

First, we calculate the value of a single instalment π. This is constant for the duration of the loan. It depends on the interest rate decision r and it is calculated as in equation (2.4). Next, we denote $N_t, t = 1, \dots, T$, the principal which stays on the account after payment of the instalment in month t. The interest credited from month $t-1$ to t is equal to $N_{t-1} \cdot (r/12)$, while the amortization is $\pi - N_{t-1} \cdot (r/12)$. We can also determine the value of N_1 and other principal amounts $N_t, t = 1, \dots, T$.

$$N_0 = \sum_{t=1}^{T} \frac{\pi}{(1 + r/12)^t} \quad \Rightarrow \quad \pi = N_0 (r/12) \Big(1 - (1 + r/12)^{-T} \Big)^{-1}, \tag{2.4}$$

$$N_t = N_0 \left(1 - \frac{(1 + r/12)^t - 1}{(1 + r/12)^T - 1} \right), \qquad t = 1, \dots, T. \tag{2.5}$$

The principal behaves as we calculated only in the case when the loan is repaid in the way as agreed at the beginning of the contract. In case of a default of the customer or a full prepayment of the loan, the evolution is different. If default occurs at time t_k, the company is thought to receive $\text{lgd} \cdot N_{t_{k-1}}$ at time t_k, while no cash-flows are exchanged between the consumer and the company in months between decision stages t_{k-1} and t_k. When prepayment occurs, the company receives the remaining principal N_{t_k}.

Cost of Financing the Loan

Another aspect which needs to be considered is the cost of financing such a loan. To obtain sufficient funds, the company could enter exactly the same contract with the market, a practice often seen in the industry. Rusý and Kopa [2018] showed that such a simple approach is not efficient. Hence, we go beyond it and give the company many possibilities to form its liability side, so the company can find the optimal financing strategy. Consider now two time instances t_i and $t_j, t_i < t_j \leq T_K$ of the program. At time t_i, in each node of every scenario $(s_i, e), s_i \in S_i, e \in E, i = 0, \dots, K-1$, the company will have two possibilities of borrowing money from the market. It could borrow from t_i to t_j and repay the money monthly with regular instalments (typically funded by the loan with matched dates) or reimburse all costs at the expiry. We will denote such amounts as $u_{t_i,t_j}(s_i, e)$ and $v_{t_i,t_j}(s_i, e)$ respectively. We can calculate the amount $u_{t_i,t_j}(t; s_i, e)$ repaid at time $t, t_i < t \leq t_j$ from a loan $u_{t_i,t_j}(s_i, e)$ as

$$u_{t_i,t_j}(t; s_i, e) = \frac{u_{t_i,t_j}(s_i, e)}{\sum_{\tau=1}^{t_j - t_i} \left(1 + m_{t_i}^{\tau}(s_i)/12\right)^{-\tau}}, \qquad t = t_i + 1, \dots, t_j. \tag{2.6}$$

On the contrary, at time t_j, the company pays back $v_{t_i,t_j}(t_j; s_i, e)$ such that

$$v_{t_i,t_j}(t_j; s_i, e) = v_{t_i,t_j}(s_i, e) \left(1 + m_{t_i}^{t_j - t_i}(s_i)/12\right)^{t_j - t_i}. \tag{2.7}$$

We also make it possible for the company to invest spare money and gain interest. Such an opportunity may arise, for example, when the client unexpectedly prepays the loan. We denote $w_{t_i,t_j}(s_i,e)$ as the amount of money lent to others for the market risk-free yield $y_{t_i}^{t_j-t_i}(s_i)$. This money will be repaid at time t_j, as the company would receive amount $w_{t_i,t_j}(t_j;s_i,e)$:

$$w_{t_i,t_j}(t_j;s_i,e) = w_{t_i,t_j}\left(1+y_{t_i}^{t_j-t_i}(s_i)/12\right)^{t_j-t_i}. \tag{2.8}$$

These decisions allow the company to freely initiate numerous contracts. However, we still require the company to meet all obligations from all previous decisions — i.e. there is no option for prepayment on the company's side. This is summarized in the following equations, which express both the amount of the company's income (I) and expenditure (J) at time instance $t_k, k=0,\dots,K$ and node $(s_k,e), s_k \in S_k, e \in E$ from the financial market. These are:

$$I_{t_k}(s_k,e) = \sum_{t_i:t_i<t_k} w_{t_i,t_k}(t_k;a_i(s_k),e) + \sum_{t_j:t_k<t_j} u_{t_k,t_j}(s_k,e) + \sum_{t_j:t_k<t_j} v_{t_k,t_j}(s_k,e), \tag{2.9}$$

$$J_{t_k}(s_k,e) = \sum_{t_j:t_k<t_j} w_{t_k,t_j}(s_k,e) + \sum_{\substack{t,t_i,t_j:\\ t_i\le t_{k-1}<t\le t_k\le t_j}} u_{t_i,t_j}(t;a_i(s_k),e) + \sum_{t_i:t_i<t_k} v_{t_i,t_k}(t_k;a_i(s_k),e). \tag{2.10}$$

In the income equation (2.9), we sum the money returned from loans provided to other institutions operating in the interbank market maturing at time t_k with the inflows from loans provided to the company by the market in that scenario. In the expenditures equation (2.10), the company needs to pay instalments from loans provided by the market in previous times and also pay the money lent to the market in the given scenario.

The mismatch between the assets and liabilities could cause duration gaps in the optimal solution. Such a portfolio composition may be considered risky and volatile. This property can be controlled by introducing various risk constraints restricting the space of the decision vector in the model. This is, however, a well-studied area of stochastic programming and it goes beyond the aims and objectives of this thesis.

Constraints and Objective Function

We continue with the specification of the remaining constraints implemented in the model. Let us denote $B_{t_k}(s_k,e)$ as the amount of money the company has in its account immediately after time t_k in scenario (s_k,e) and C_{t_k} as the company's operating costs of the loan from time t_k to t_{k+1}. We have:

$$\begin{aligned} B_{t_k}(s_k,e) = {} & B_{t_{k-1}}(a_{k-1}(s_k),e) - C_{t_k} + I_{t_k}(s_k,e) - J_{t_k}(s_k,e) - \mathbb{I}_{[k=0]}N_0 \\ & + \mathbb{I}_{[t_k<t(e)]}(t_k-t_{k-1})\pi + \mathbb{I}_{[t_k=t(e)]}d(e)\cdot \mathrm{lgd}\cdot N_{t_{k-1}} \\ & + \mathbb{I}_{[t_k=t(e)]}(1-d(e))\left((t_k-t_{k-1})\pi + N_{t_k}\right), \end{aligned} \tag{2.11}$$

for $k=0,\dots,K$, where $B_{t_{-1}}=0$, and also $C_{t_K}=0$. The relationship on the first line of (2.11) expresses the initial exchange of the principal and the cash-flows between the bank and the company. On the second and the third line, the amount of funds the company receives from the customer in different stages under the scenario $e \in E$ is described. The indicator functions mean the same as described in Section 2.1.2.

The definition of the company's cash account brings us to a very natural survival condition such that the company's cash account must not be lower than 0. We require this only in stages from $0, \dots, K-1$ as, in the last stage, the loan is concluded and we look at the final balance, its performance through its life and asses its profitability. We require:

$$B_{t_k}(s_k, e) \geq 0, \qquad e \in E, s_k \in S_k, \;\; k = 0, \dots, K-1. \tag{2.12}$$

Next, we move to the cash-flows which take place between decision stages. The company has to make sure it has enough money to cover its expenditures up to the next decision stage. Such a liquidity constraint can be implemented only by checking whether the company has enough funds in the month before the next decision stage as all the cash-flows in between the decision stages are the same every month. The constraint is as follows:

$$0 \leq B_{t_k}(s_k, e) - \sum_{\substack{t, t_i, t_j: \\ t_i \leq t_k < t < t_{k+1} \leq t_j}} u_{t_i,t_j}(t; a_i(s_k), e) + \mathbb{I}_{[t_k < t(e)]}(t_{k+1} - t_k - 1)\pi, \tag{2.13}$$

$$k = 0, \dots, K-1.$$

The next step is to express the value of the loan in each node. Such a value is calculated as the sum of discounted cash-flows from loans running at the given time. Let us denote $P(t_k, t_l; s_k)$ as the discount factor from time t_k to time t_l at a node $s_k \in S_k$. For easier formulation, we divide payments between assets $A_{t_k}(s_k, e)$ and liabilities $L_{t_k}(s_k, e)$. These can be calculated as follows:

$$A_{t_k}(s_k, e) = \sum_{\substack{t_i, t_j: \\ t_i \leq t_k < t_j}} P(t_k, t_j; s_k) w_{t_i,t_j}(t_j; a_i(s_k), e) + \mathbb{I}_{[t_k < t(e)]} \sum_{t: t_k < t \leq T} P(t_k, t; s_k)\pi,$$

$$L_{t_k}(s_k, e) = \sum_{\substack{t, t_i, t_j: \\ t_i \leq t_k < t \leq t_j}} P(t_k, t; s_k) u_{t_i,t_j}(t; a_i(s_k), e) + \sum_{\substack{t_i, t_j: \\ t_i \leq t_k < t_j}} P(t_k, t_j; s_k) v_{t_i,t_j}(t_j; a_i(s_k), e).$$

Assets are calculated as the sum of discounted cash-flows stemming from loans provided to the market and payments the customer is yet to make — before the event e is observed. For liabilities, we total all instalments which the company is yet to make. The difference between values of assets and liabilities (the asset-liability gap) together with the amount of money in the current account $B_{t_k}(s_k, e)$ gives us the value of the portfolio $V_{t_k}(s_k, e)$ at node (s_k, e). This reads as:

$$V_{t_k}(s_k, e) = B_{t_k}(s_k, e) + A_{t_k}(s_k, e) - L_{t_k}(s_k, e). \tag{2.14}$$

In the final stage, when we have no running contracts, this turns into the net income — the total loan profit. This leads us to the formulation of the objective function $f(r, u, v, w)$, which expresses the value of the loan at the final time horizon. We have:

$$f(r, u, v, w) = p(r) \cdot \sum_{s_K \in S_K, e \in E} p(s_K, e, r) V_{t_K}(s_K, e), \tag{2.15}$$

where variables u, v, w symbolically stand for the sets of decision variables as defined above. In (2.15), we weigh each scenario according to its probability

$p(s_K, e, r)$ and we multiply the entire sum by the probability $p(r)$ that the loan is agreed to.

To complete the model formulation, we need to specify the final set of constraints. This will consist of the already-mentioned non-anticipativity constraints, as we need to make sure that the decisions of the company in times and scenarios where the event has not yet been observed are the same. We impose:

$$\begin{gathered} u_{t_i,t_j}(s_i, e_1) = u_{t_i,t_j}(s_i, e_2), \quad v_{t_i,t_j}(s_i, e_1) = v_{t_i,t_j}(s_i, e_2), \\ w_{t_i,t_j}(s_i, e_1) = w_{t_i,t_j}(s_i, e_2), \\ \forall s_i \in S_i, e_1, e_2 \in E : t_i < \min\{t(e_1), t(e_2)\}, t_j > t_i, i = 0, \dots, K. \end{gathered} \tag{2.16}$$

Finally, we would like to comment on model limitations. The model is specifically designed to capture the relationship between an individual person and the lending company. For other borrower-lender relationships, the model would need to be slightly adjusted. For example, for peer-to-peer lending, the liability side would need to be formulated differently, as individual people usually do not have the borrowing options that are specified in the model. On the other hand, the borrowing of small companies also has its specifics. Furthermore, individual approaches and a more detailed analysis of the relevant data are necessary, especially for larger loans.

Stochastic Programming Model Formulation

Let us now summarize the decision variables and present the complete program. We have months $t, t \in \{0, \dots, T\}$ and decision times $t_k, k \in \{0, \dots, K\}$. In the first decision stage t_0, the company offers a loan to the customer. If it is accepted, the company borrows money at the market at observed (known) interest rates. In the next decision stages $(t_1, \dots t_{K-1})$, we model the evolution of interest rates by the interest rate tree, where at decision time t_k we have a set of nodes S_k. Finally, in the last stage t_K, no decisions are made as we only evaluate the final position at the end of the loan contract. Time t_i ancestor of a node $s_k, i < k$, is denoted as $a_i(s_k)$. Set S_K denotes all final-stage nodes and, hence, also denotes interest rate scenarios. In these future stages $(t_k, k > 0)$, the rate of the consumer loan does not change. We only evaluate all contracts of the company, which can also initiate new contracts at a price determined by the interest rates in the given scenario. Set E describes all events which can happen to the customer. We have functions $d(e)$ identifying whether it is default or prepayment and $t(e)$ specifying at which decision stage it is observed. These are important mainly for calculation of scenario probabilities and specification of cash flows between the company and the customer in different stages. Finally, the pair $(s_K, e), s_K \in S_K, e \in E$ denotes the program scenario. We have introduced four types of decisions over which we optimize:

- r — the interest rate decision

For times $t_i, t_j, i < j \in 0, \dots, K$ and each node $(s_i, e), s_i \in S_i, e \in E$:

- $u_{t_i,t_j}(s_i, e)$ — amount borrowed at t_i, repaid monthly with maturity t_j
- $v_{t_i,t_j}(s_i, e)$ — amount borrowed at t_i, repaid in total at time t_j

- $w_{t_i,t_j}(s_i, e)$ — amount lent at t_i, repaid in total at time t_j

We also defined the following quantities:

- b_0, b_1 — midrate and interest rate sensitivity of the customer
- $p(r)$ — probability of accepting the loan by the customer
- $h(t_k, r)$ — hazard rate for default of the customer at t_k with rate r
- $\beta_0^h, \beta_1^h, \beta_3^h$ — coefficients for logistic regression model for $h(t_k, r)$
- $g(t_k, r)$ — hazard rate for loan prepayment at t_k with rate r
- $\beta_0^g, \beta_1^g, \beta_3^g$ — coefficients for logistic regression model for $g(t_k, r)$
- $p(s_K, e, r)$ — probability of scenario (s_K, e), if loan with interest r is agreed
- π — instalment of the consumer loan
- N_t — the principal remaining at time t from the consumer loan
- y_t^τ, m_t^τ — risk-free rates and rates the company pays for loans at the market at time t with time-to-maturity τ
- $u_{t_i,t_j}(t; s_i, e)$ — amount repaid at time t from loan $u_{t_i,t_j}(s_i, e)$
- $v_{t_i,t_j}(t_j; s_i, e)$ — amount repaid at t_j from loan $v_{t_i,t_j}(s_i, e)$
- $w_{t_i,t_j}(t_j; s_i, e)$ — amount repaid at t_j from loan $w_{t_i,t_j}(s_i, e)$
- $I_{t_k}(s_k, e)$ — income on the market side at t_k, node (s_k, e)
- $J_{t_k}(s_k, e)$ — expenditures on the market side at t_k, node (s_k, e)
- C_{t_k} — operating costs of the loan from time t_k to time t_{k+1}
- $B_{t_k}(s_k, e)$ — amount on current account at t_k, node (s_k, e)
- $A_{t_k}(s_k, e)$ — value of assets at t_k, node (s_k, e)
- $L_{t_k}(s_k, e)$ — value of liabilities at t_k, node (s_k, e)
- $V_{t_k}(s_k, e)$ — value of the loan at t_k, node (s_k, e)

From here, we formulate the asset-liability multi-stage stochastic program as:

$$\min_{r,u,v,w} \quad -p(r) \cdot \sum_{s_K \in S_K, e \in E} p(s_K, e, r) V_{t_K}(s_K, e), \tag{2.17}$$

$$\text{s.t. } p(r) = \frac{1}{1 + \exp\{-b_1(b_0 - r)\}},$$

$$h(t_k, r) = \frac{1}{1 + \exp\{-(\beta_0^h + \beta_1^h r + \beta_3^h t_k)\}},$$

$$g(t_k, r) = \frac{1}{1 + \exp\{-(\beta_0^g + \beta_1^g r + \beta_3^g t_k)\}}, \qquad k = 1, \dots, K,$$

$$
\begin{aligned}
p(s_K, e, r) = \prod_{k=1}^{K} \Big(& \mathbb{I}_{[t(e)=t_k]} d(e) h(t_k, r) + \mathbb{I}_{[t(e)=t_k]} \Big(1 - d(e)\Big) g\Big(t_k, r\Big) \\
& + \mathbb{I}_{[t_k<t(e)]} \Big(1 - h(t_k, r) - g(t_k, r)\Big) + \mathbb{I}_{[t_k>t(e)]} \Big) \Big(|S_{k-1}|/|S_k|\Big), \\
& \qquad\qquad\qquad s_K \in S_K, e \in E, \\
\pi &= N_0(r/12)\Big(1 - (1 + r/12)^{-T}\Big)^{-1}, \\
N_t &= N_0 \left(1 - \frac{(1 + r/12)^t - 1}{(1 + r/12)^T - 1}\right), \qquad t = 1, \dots, T, \\
u_{t_i,t_j}(t; s_i, e) &= \frac{u_{t_i,t_j}(s_i, e)}{\sum_{\tau=1}^{t_j-t_i} (1 + m_{t_i}^{\tau}(s_i)/12)^{-\tau}}, \qquad t = t_i + 1, \dots, t_j, \\
v_{t_i,t_j}(t_j; s_i, e) &= v_{t_i,t_j}(s_i, e) \left(1 + m_{t_i}^{t_j-t_i}(s_i)/12\right)^{t_j-t_i}, \\
w_{t_i,t_j}(t_j; s_i, e) &= w_{t_i,t_j} \left(1 + y_{t_i}^{t_j-t_i}(s_i)/12\right)^{t_j-t_i}, \qquad i, j = 0, \dots, K, i < j, \\
I_{t_k}(s_k, e) &= \sum_{t_i: t_i<t_k} w_{t_i,t_k}(t_k; a_i(s_k), e) + \sum_{t_j: t_k<t_j} u_{t_k,t_j}(s_k, e) + \sum_{t_j: t_k<t_j} v_{t_k,t_j}(s_k, e), \\
J_{t_k}(s_k, e) &= \sum_{t_j: t_k<t_j} w_{t_k,t_j}(s_k, e) + \sum_{\substack{t, t_i, t_j: \\ t_i \le t_{k-1} < t \le t_k \le t_j}} u_{t_i,t_j}(t; a_i(s_k), e) + \sum_{t_i: t_i<t_k} v_{t_i,t_k}(t_k; a_i(s_k), e), \\
B_{t_k}(s_k, e) &= B_{t_{k-1}}(a_{k-1}(s_k), e) - C_{t_k} + I_{t_k}(s_k, e) - J_{t_k}(s_k, e) - \mathbb{I}_{[k=0]} N_0 \\
& \quad + \mathbb{I}_{[t_k<t(e)]}(t_k - t_{k-1})\pi + \mathbb{I}_{[t_k=t(e)]} d(e) \cdot \mathrm{lgd} \cdot N_{t_{k-1}} \\
& \quad + \mathbb{I}_{[t_k=t(e)]}(1 - d(e)) \left((t_k - t_{k-1})\pi + N_{t_k}\right), \\
A_{t_k}(s_k, e) &= \sum_{\substack{t_i, t_j: \\ t_i \le t_k < t_j}} P(t_k, t_j; s_k) w_{t_i,t_j}(t_j; a_i(s_k), e) + \mathbb{I}_{[t_k<t(e)]} \sum_{t: t_k<t\le T} P(t_k, t; s_k)\pi, \\
L_{t_k}(s_k, e) &= \sum_{\substack{t, t_i, t_j: \\ t_i \le t_k < t \le t_j}} P(t_k, t; s_k) u_{t_i,t_j}(t; a_i(s_k), e) + \sum_{\substack{t_i, t_j: \\ t_i \le t_k < t_j}} P(t_k, t_j; s_k) v_{t_i,t_j}(t_j; a_i(s_k), e), \\
V_{t_k}(s_k, e) &= B_{t_k}(s_k, e) + A_{t_k}(s_k, e) - L_{t_k}(s_k, e), \quad k = 0, \dots, K, s_k \in S_k, e \in E, \\
B_{t_k}(s_k, e) &\ge \sum_{\substack{t, t_i, t_j: \\ t_i \le t_k < t < t_{k+1} \le t_j}} u_{t_i,t_j}(t; a_i(s_k), e) - \mathbb{I}_{[t_k<t(e)]}(t_{k+1} - t_k - 1)\pi, \\
B_{t_k}(s_k, e) &\ge 0, \qquad k = 0, \dots, K - 1, s_k \in S_k, e \in E, \\
u_{t_i,t_j}(s_i, e_1) &= u_{t_i,t_j}(s_i, e_2), \quad v_{t_i,t_j}(s_i, e_1) = v_{t_i,t_j}(s_i, e_2), \\
w_{t_i,t_j}(s_i, e_1) &= w_{t_i,t_j}(s_i, e_2), \\
& i, j = 0, \dots, K, i < j, s_i \in S_i, e_1, e_2 \in E : t_i < \min\{t(e_1), t(e_2)\}
\end{aligned}
$$

Note especially the first four equations in the model definition, which capture the endogeneity in the random variables induced by the initial interest rate decision r. There is another non-linearity in the program in the equations for π and N_t. The other equations form a usual asset-liability multi-stage stochastic program.

2.2 Numerical Results

In this section, we present the results of the model. We focus on how decision-dependent uncertainty and the parameters associated with it affect the model solution, especially the interest rate decision. We also discuss the losses connected to offering a non-optimal interest rate for the loan.

For this model, we set the notional to be $N_0 = 50000$ CZK with maturity $T = 5$ years. We set decision stages to be at the end of each year ($K = 5$). The branching of the interest rate tree was chosen to be $5-4-3-2-1$, leading to $|S_K| = 120$ interest rate scenarios. There is no branching to the final stage, as no decision is made there and we only evaluate the final value of a loan. Given that we have 5 "future" decision stages and that, in each stage, we can have default or prepayment, we have a total of $|E| = 2 \cdot K = 10$ events leading up to $|S_K| \cdot |E| = 120 \cdot 10 = 1200$ scenarios.

The program was written in GAMS and solved by CONOPT3 on a standard laptop (Intel Core i5 2.60 GHz, 8GB RAM). Scenario generation and results' analysis were performed in R. The model itself had 58690 variables and 45619 constraints with 322783 Jacobian elements, 4989 of which were non-linear. The Hessian of the Lagrangian had 1 element on the diagonal, 4250 elements below the diagonal and 3052 non-linear variables.

2.2.1 Model Solution

Here, we will give a detailed description of the solution and its properties for a single model. For this purpose, we chose parameter values as midrate $b_0 = 0.14$, interest rate sensitivity $b_1 = 100$ and rating $\rho = 2$. The optimal solution of the program was to offer the customer a rate $r = 12.24\%$, with the probability of accepting the loan as $p(r) = 0.853$. The optimal borrowing and lending strategy of the company was to close only one-year loans. If spare money is available, then the company should lend to the market for the longest period possible (until the final time horizon). That is due to interest rate tree having relatively constant expected future rates and also because the shorter the loan, the cheaper. If the rates were, for example, increasing, the company would tend to close longer loans with the market. The optimal value of the program showed that expected profit from the loan was 7392 CZK on the considered 50000 CZK loan. This approximates to an annual gain of 0.028 CZK per 1 CZK borrowed. However, note that this depends greatly on the characteristics of the customer.

To analyse the results, we investigated loan performance in different interest rate scenarios for different events. First, we divided interest rate scenarios into five groups — *purple, blue, green, yellow* and *red* — depending on their first-stage node (see top-left figure in Figure 2.3). This will help us to illustrate the dependency of the optimal loan value on interest rate evolution. We looked into performance when the customer complied with the original terms of the loan. Because of the one-year borrowing strategy, the company profits on an interest rate decrease and it loses money when the interest rate increases. However, it is also able to use a *high* (purple) interest rate environment to compensate for the initial loss by lending money earned from the loan in order to earn high interest in latter stages of the loan. The bottom figures in Figure 2.3 show the performance of a loan when it is fully repaid/defaulted at the end of the first year. One can see that a high interest rate environment is preferred for loan prepayment because the company can then reinvest money for higher interest. It is the complete opposite for the case of customer defaults. Then, the company is required to borrow additional money to finance its liabilities and this costs more in an environment with higher interest rates.

We also shortly investigated sensitivity on the mark-up $m(\tau)$, which the market charges the company. Our hypothesis was that increasing the mark-up increases the costs of the loan for the company and hence, "not accepting" the loan is relatively less expensive. This should lead to higher interest rate decision. However, the company would not increase the interest rate by the same margin as the client would be discouraged from entering the loan. This was confirmed by the program, as when we set the mark-up twice the analysed value, the objective value of the program decreased to 6725 CZK and the optimal offered interest rate was 12.34% with $p(r) = 0.840$.

A question arises about what the added value of the inclusion of decision-dependent randomness is. First, note that it is not possible to formulate the model without decision-dependent randomness in the probability of accepting the loan, as this would not really make sense. Therefore, we have solved a model without decision-dependent randomness in default and prepayment. There, the results depend heavily on the strength of the relationship between the offered interest rate and the default by the customer. For our particular case, the objective value of the non-decision-dependent randomness model was 7348 CZK, which means a difference of 44 CZK. However, for ratings 3 and 4, the difference was 2209 and 5635 CZK respectively. From this, we can conclude that it is advantageous to reflect the relationship, especially in cases where the initial decision has a greater

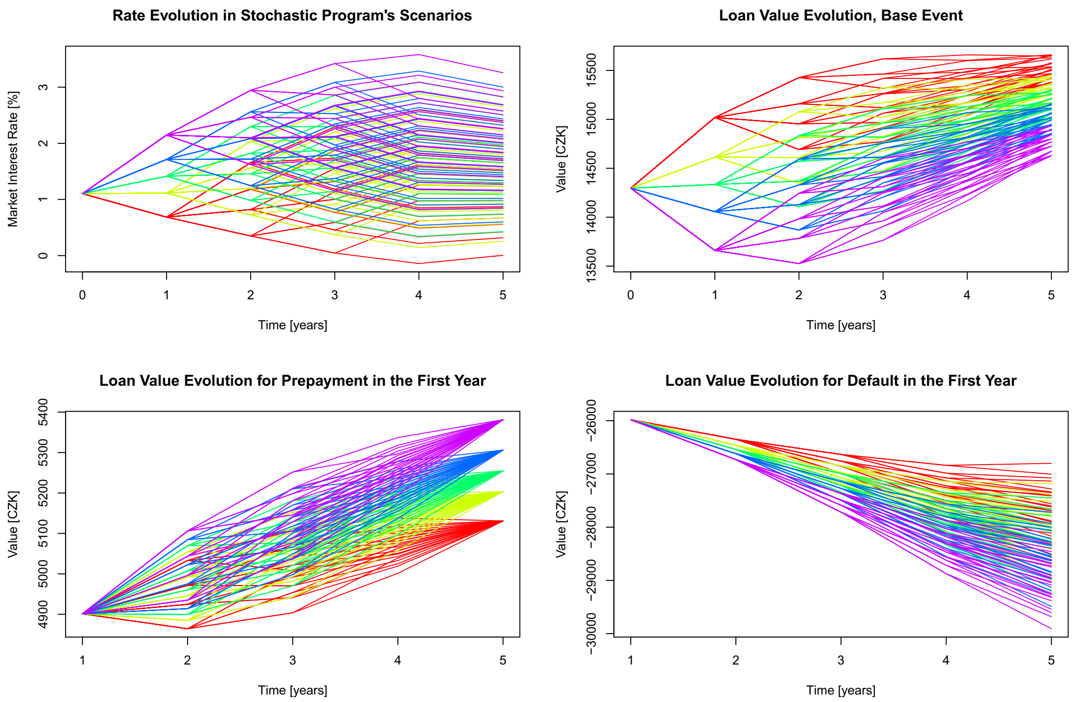

Figure 2.3: Interest rate and the optimal loan value evolution in time.

impact on the random variable.

The framework we have introduced allows us to take into account all the elements of the life of a consumer loan (customer properties, event probabilities, interest rate evolution, etc.), assess their costs depending on the company's decisions and then select the optimal ones. Division of this joint optimization into sub-problems is a simplification which does not produce accurate results. For example, if the company would be forced to replicate the customer's loan at the market, then, it would pay more interest than it would be required for the case of customer's prepayment. This higher cost of prepayment would make the company conservative and cause them to offer rate lower than the optimal one.

2.2.2 Sensitivity Analysis

In this section, we will investigate the behaviour of the optimal decisions and optimal value of the model when we modify the characteristics of the customer. Parameters ρ and b_1 are of the main interest, as they capture the decision-dependent randomness in the model. We will look into how the offered interest rate depends on the probability distribution of accepting the loan. Moreover, we will analyse how much the company loses when it makes a wrong decision regarding the offered interest rate. This will be studied for all possible values of the rating.

We consider rating $\rho = \{1, 2, 3, 4\}$, which defines the probability of default and prepayment as given in (2.2) and (2.3). Then, we specify the probability of accepting the loan by parameters midrate b_0 and interest rate sensitivity b_1, as described in Section 2.1.2. Interest rate sensitivity is the parameter which captures the strength of decision-dependent randomness in the probability of accepting the loan. We see that a higher value of b_1 implies that the customer is more sensitive — he has good information about current market conditions and any deviation from the midrate means either a large increase or a large decrease in the probability of accepting the loan. Midrates are considered to be from a sequence $\{0.1, 0.12, 0.14, 0.16, 0.18\}$, while sensitivities will take on values of $\{25, 35, 50, 75, 100, 125, 150, 175, 200\}$.

In Figure 2.4 and Figures 2.7, 2.9, 2.11 and 2.13 in Section **??**, we show the results of the model for each midrate, each interest rate sensitivity and each rating listed above. The figures report the results of runs of the program with a common midrate and consist of four plots. First, in the top-left one, we show logistic curves which are generated by the pair b_0 and b_1. There, one can note that all the curves intersect at one point — the common midrate with 50% probability. Second, in the top-right plot, we show how the optimal interest rate varies for different ratings across all sensitivity values. Finally, in the bottom two figures, the loan probability (probability that the customer accepts the offered interest rate) and the objective value of the program as given by the optimal solution are shown.

From there, we can observe several properties of the optimization problem:

- It is always more profitable to have a consumer with a better rating.
- A consumer with a better rating gets better rates from the company.
- The higher the interest rate sensitivity b_1, the higher the loan probability.

This is simply because it is less costly to ensure the customer has a greater probability of accepting the loan.

- The objective function is not monotone in b_1. For some settings (see Figure 2.4 for midrate 0.1), the company makes money on the fact that the customer is willing to accept higher rates.

We can see that all these findings have real application which is required for any practical use of the model. It is clear that the shape of the "probability of accepting the loan" curve is absolutely crucial for the model to produce the most realistic results. We believe that the set of curves we chose approximate most of the options which can practically happen. In the end, only local properties of the curve around the "almost optimal rate" are what matters most.

We also investigated how important it is for the company to set the interest rate correctly. In other words, how much the company loses when it offers a non-optimal interest rate to the customer. To answer this question, we present Figure 2.5, which depicts the dependence of the objective value of the program when fixing the interest rate r and also the interest rate sensitivity b_1 on certain values. The figure is given for a customer with midrate $b_0 = 0.1$ and rating 2. Figures for other midrates are presented in Figures 2.8, 2.10, 2.12, 2.14 and they can be also viewed in the interactive mode available at *https://plot.ly/∼rusy/*.

The conclusion which we obtain from Figure 2.5 is that the difference between objective values of the optimal solution and a solution with fixed r depends greatly on the interest rate sensitivity of the customer. It is apparent that when the customer is sensitive (large b_1) it is extremely important to "hit" the optimal rate

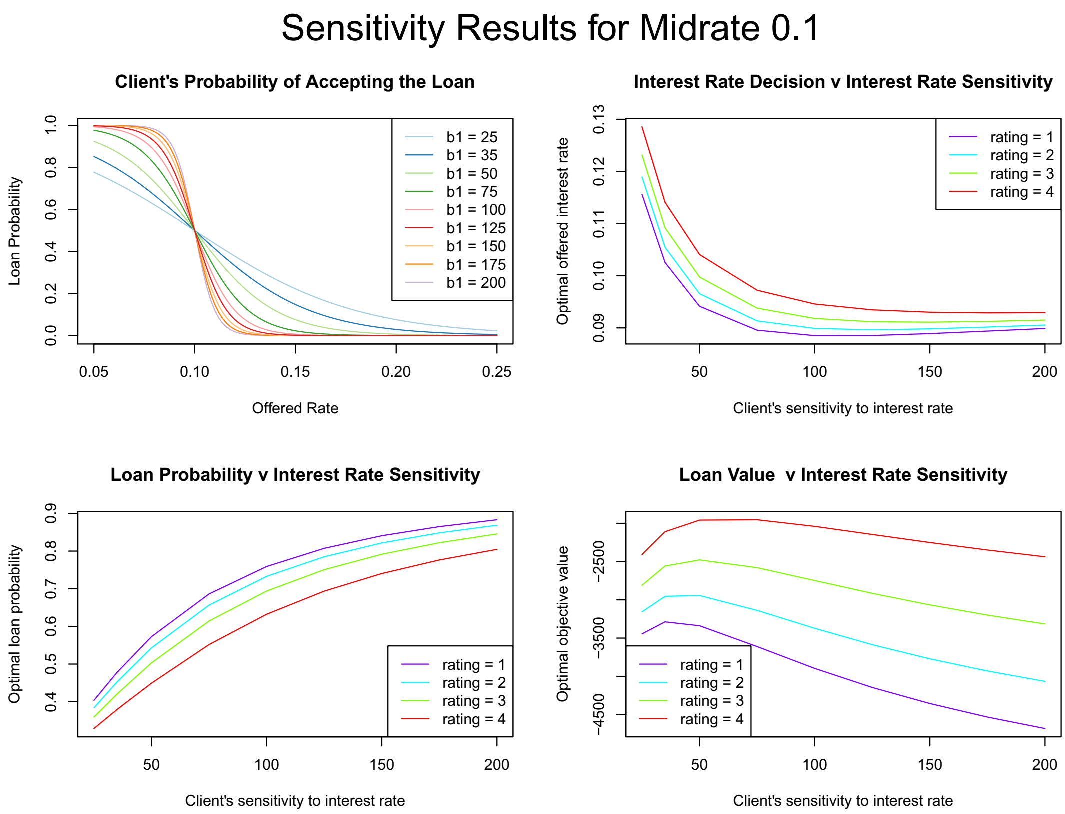

Figure 2.4: Sensitivity analysis results for midrate 0.1.

with the offer, otherwise the company loses a significant amount of money. What can also be seen from Figure 2.5 is that it is more costly to offer a higher interest rate than a lower interest rate compared to the optimal rate. The potentially missed opportunity on a loan has greater impact on the objective than smaller revenue from a loan with a lower interest rate.

In order to quantify the loss which is incurred by offering a non-optimal interest rate, we present Table 2.2 (and Tables 2.3, 2.4, 2.5 and 2.6). Here, we show for each considered customer, the loss incurred by missing the optimal interest rate by $\pm 1\%$ - i.e. by ± 100bps. From an initial glance, we can learn that the losses increase with the value of interest rate sensitivity. This means the more sensitive a customer is, the more careful the company should be with its offer. We can also see the effect of the rating. Here, absolute costs are greater with a better rating. On the contrary, relative costs increase with lower ratings. This is due to the smaller objective value of a loan for worse customers. The tables also confirm that it is generally better to offer lower rates than higher rates, which is something we explored in the previous paragraph. We also compare losses across midrates. Here, we see that they become larger with increasing midrate in absolute values, but the opposite is true in relative terms. This imbalance is due to the absolute change of 1% which is applied. This has a stronger relative effect for lower midrates. However, the fact that we “play” for more money for higher midrates implies a greater absolute effect there.

The relationship between the loss incurred to the company and the distance of the offered interest rate from the optimal value can be exploited in more detail in Figure 2.6. Here, one can see the effect of rating, interest rate sensitivity, midrate and the distance of the offered rate from the optimal decision. We can draw a similar conclusion as that from the numbers in Table 2.2; it is more costly to offer

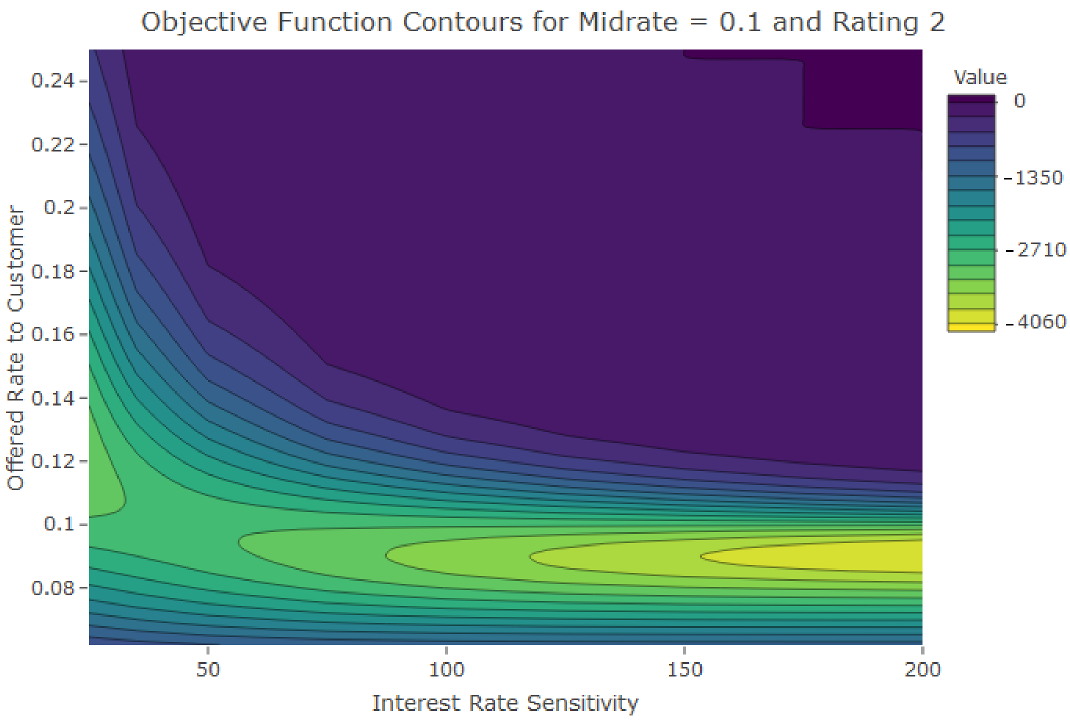

Figure 2.5: Contour plot constructed from objective values of the program when fixing offered interest rate r for different values of b_1.

a higher interest rate than a lower one and the loss depends largely on the interest rate sensitivity of the customer. Moreover, we can see that the loss appears to be concave in distance from the optimal interest rate decision, meaning the loss increases with increasing rate when moving away from the optimal decision.

In summary, we investigated the effect of customer's properties such as the expected offered rate, interest rate sensitivity and credit quality on a loan provided by a company. These properties are essential as they capture the decision-dependent uncertainty which is present in the optimization model. We focused on calculation of losses caused by not offering optimal interest rate. We saw that especially for the more interest rate-sensitive customers, the losses can be very high and any kind of simplification of the joint optimization model can lead to wrong decisions. That implies that it is important to consider the decision-dependent nature of the model in its entirety and ignoring it, even only in parts, can lead to a significant reduction of profits for the company.

	$b_1 \downarrow$	Absolute [CZK]				Relative [%]			
Rating $\rightarrow$		1	2	3	4	1	2	3	4
-1%	25	66	62	57	52	1.9	2	2	2.2
	35	107	102	95	86	3.2	3.4	3.7	4.1
	50	177	169	158	142	5.3	5.8	6.4	7.3
	75	296	286	272	250	8.2	9.1	10.5	12.8
	100	401	391	377	353	10.3	11.6	13.7	17.3
	125	489	481	467	445	11.8	13.4	16	20.7
	150	564	557	544	522	13	14.8	17.7	23.2
	175	628	621	608	588	13.9	15.8	19	25
	200	683	676	663	643	14.6	16.6	20	26.4
+1%	25	57	54	50	44	1.7	1.7	1.8	1.8
	35	95	89	81	70	2.9	3	3.2	3.3
	50	163	151	136	117	4.9	5.1	5.5	6
	75	304	282	252	211	8.4	9	9.8	10.8
	100	472	439	392	327	12.1	13	14.3	16.1
	125	666	619	552	460	16.1	17.3	19	21.4
	150	881	819	730	605	20.2	21.7	23.8	26.9
	175	1117	1037	923	761	24.6	26.4	28.9	32.4
	200	1372	1270	1128	927	29.3	31.2	34	38

Table 2.2: Differences in objective function for ±1% change in offered rate against the optimal value for midrate 10%.

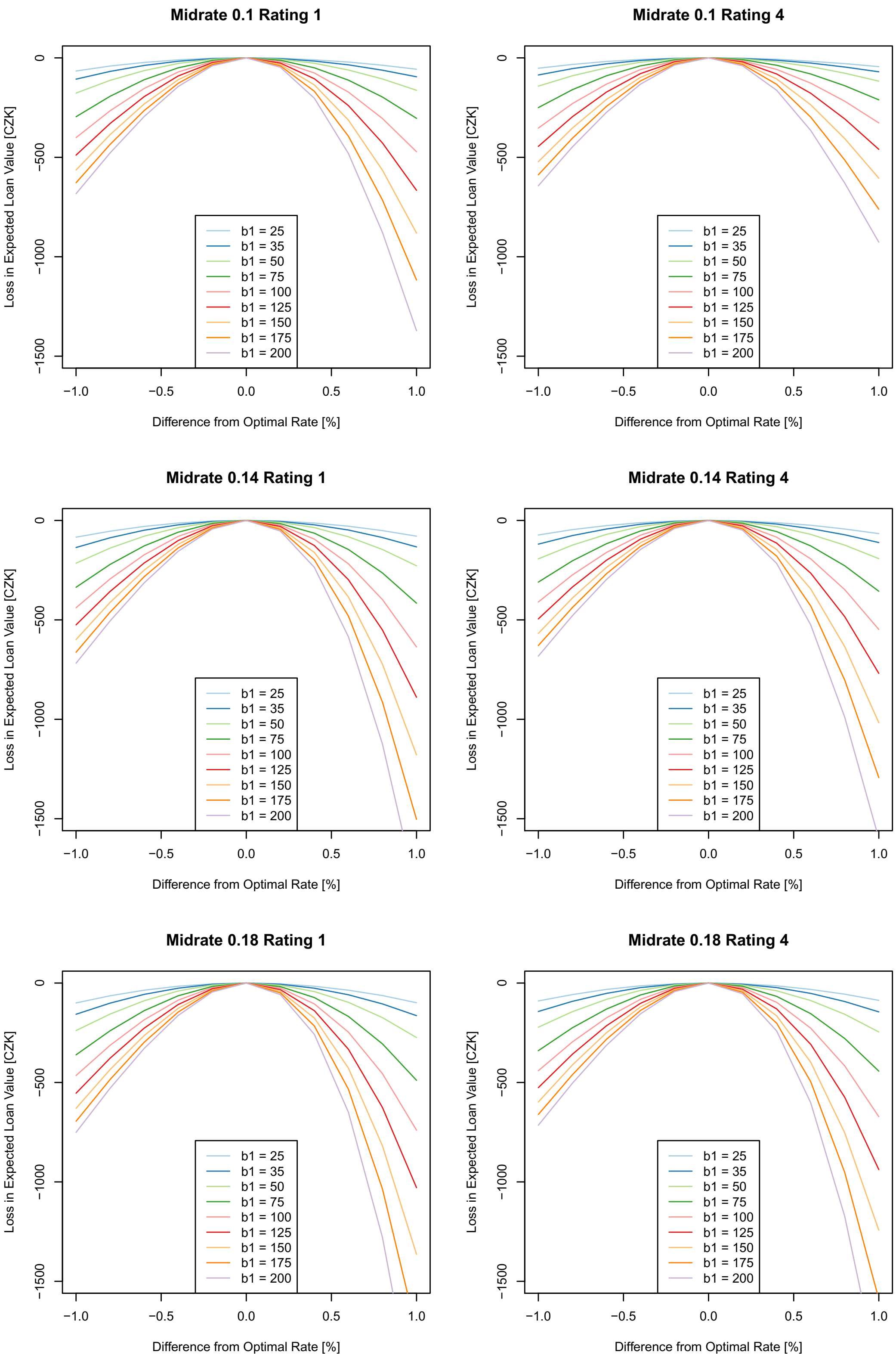

Figure 2.6: Comparison of losses to the company caused by not offering optimal interest rate to the customer.

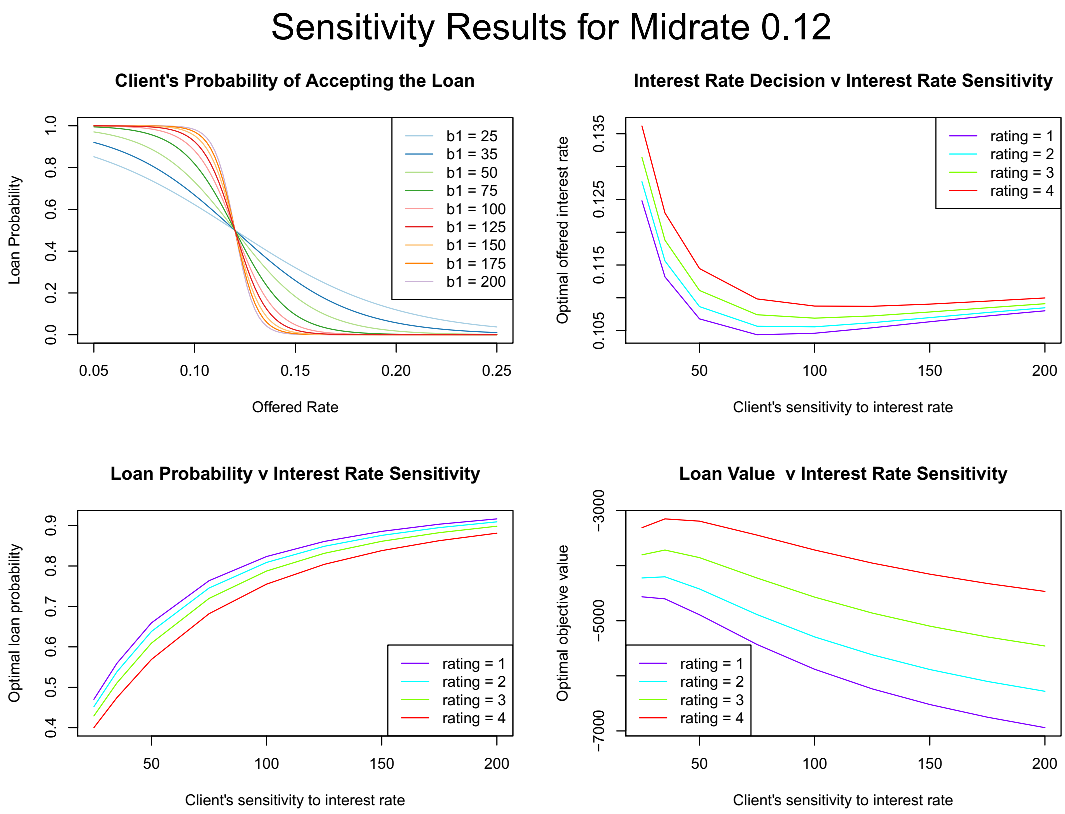

Figure 2.7: Sensitivity analysis results for midrate 0.12.

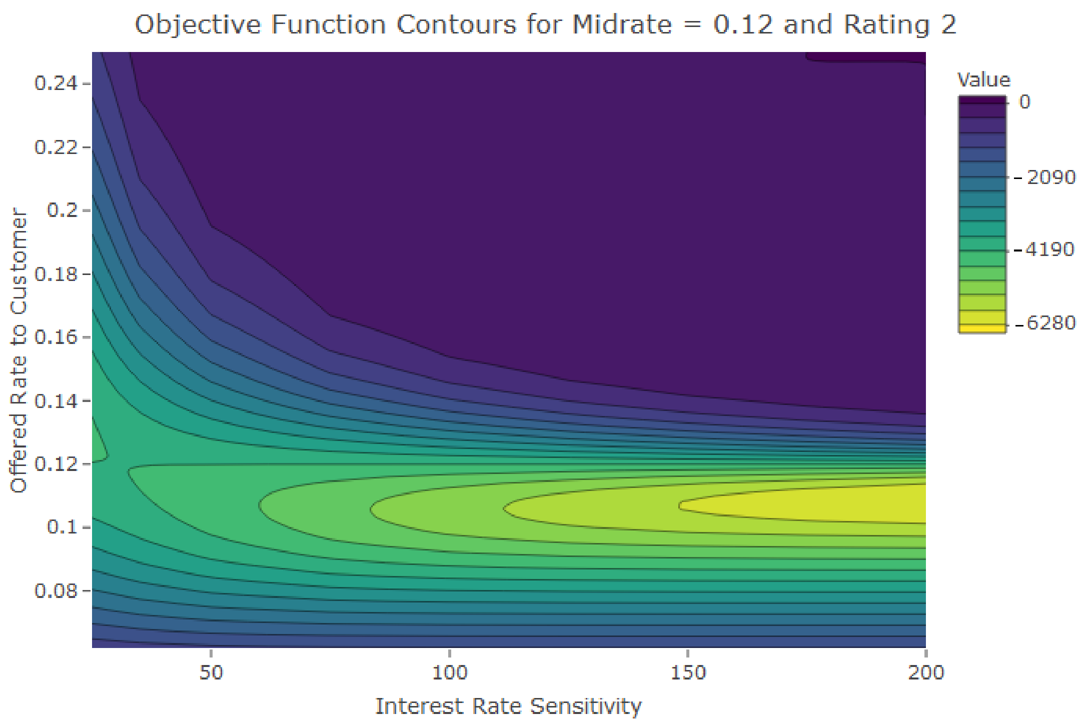

Figure 2.8: Contour plot constructed from objective values of the program when fixing offered interest rate r for different values of interest rate sensitivity b_1.

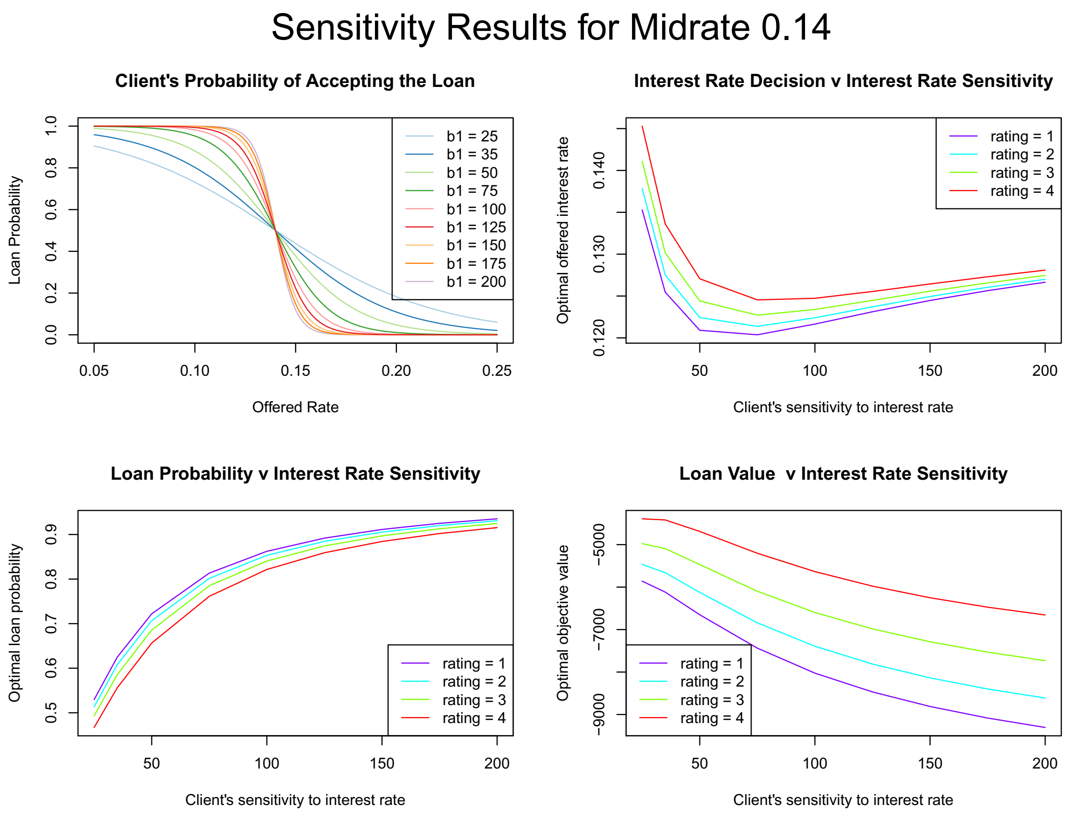

Figure 2.9: Sensitivity analysis results for midrate 0.14.

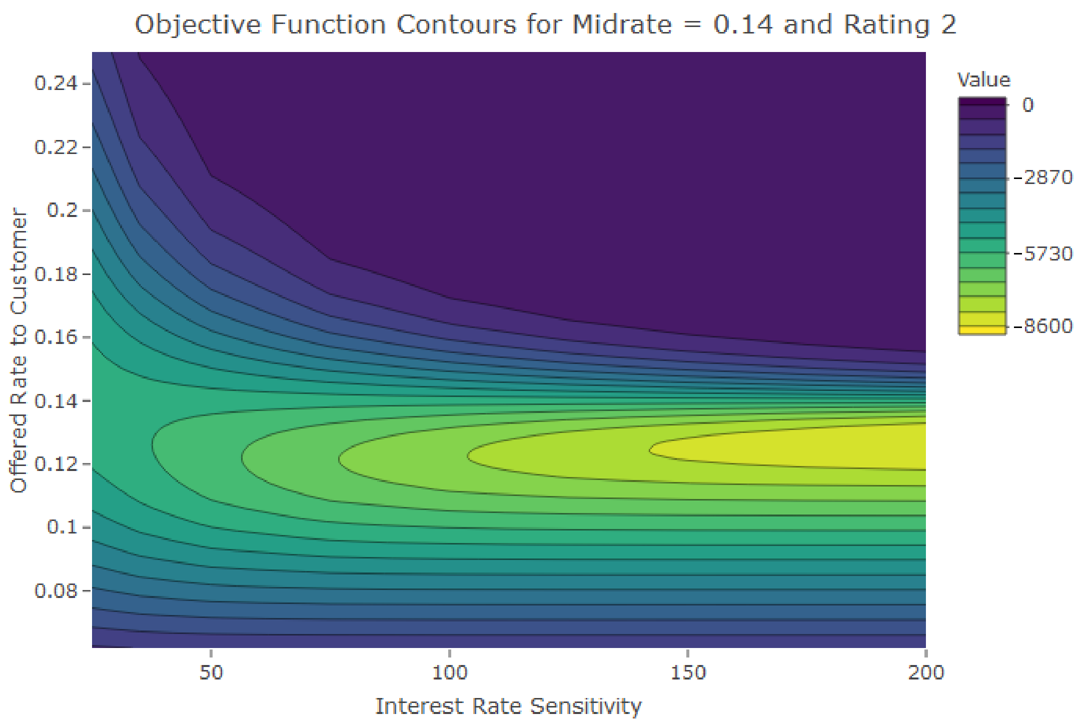

Figure 2.10: Contour plot constructed from objective values of the program when fixing offered interest rate r for different values of interest rate sensitivity b_1.

Rating →	b_1 ↓	Absolute [CZK]				Relative [%]			
		1	2	3	4	1	2	3	4
-1%	25	75	72	68	63	1.6	1.7	1.8	1.9
	35	125	120	112	103	2.7	2.8	3	3.3
	50	200	195	186	173	4.1	4.4	4.8	5.4
	75	319	314	304	289	5.9	6.4	7.2	8.4
	100	423	417	406	391	7.2	7.9	8.9	10.5
	125	511	504	494	478	8.2	9	10.2	12.1
	150	586	580	568	552	9	9.8	11.1	13.3
	175	649	642	631	614	9.6	10.5	11.9	14.2
	200	703	697	685	668	10.1	11.1	12.6	14.9
+1%	25	69	65	61	55	1.5	1.5	1.6	1.7
	35	115	109	101	91	2.5	2.6	2.7	2.9
	50	200	190	176	157	4.1	4.3	4.6	4.9
	75	370	353	328	294	6.8	7.2	7.8	8.5
	100	569	544	509	458	9.7	10.3	11.1	12.3
	125	797	764	713	643	12.8	13.6	14.7	16.3
	150	1053	1008	943	849	16.2	17.1	18.5	20.4
	175	1341	1282	1196	1076	19.9	21	22.6	24.9
	200	1655	1581	1472	1321	23.9	25.2	27	29.6

Table 2.3: Differences in objective function for ±1% change in offered rate against the optimal value for midrate 12%.

Rating →	b_1 ↓	Absolute [CZK]				Relative [%]			
		1	2	3	4	1	2	3	4
-1%	25	84	82	78	73	1.4	1.5	1.6	1.7
	35	136	132	127	119	2.2	2.3	2.5	2.7
	50	215	210	203	193	3.2	3.4	3.7	4.1
	75	336	331	322	310	4.5	4.8	5.3	5.9
	100	439	433	424	410	5.5	5.9	6.4	7.3
	125	525	519	510	495	6.2	6.6	7.3	8.3
	150	599	594	583	568	6.8	7.3	8	9.1
	175	663	657	646	629	7.3	7.8	8.6	9.7
	200	718	711	699	682	7.7	8.3	9	10.2
+1%	25	79	76	72	66	1.4	1.4	1.4	1.5
	35	133	128	121	111	2.2	2.3	2.4	2.5
	50	228	220	208	192	3.4	3.6	3.8	4.1
	75	416	403	383	356	5.6	5.9	6.3	6.8
	100	636	617	589	548	7.9	8.4	8.9	9.7
	125	889	864	824	769	10.5	11.1	11.8	12.9
	150	1179	1142	1092	1017	13.4	14	15	16.3
	175	1503	1457	1388	1293	16.6	17.3	18.4	20
	200	1866	1808	1721	1600	20.1	21	22.3	24

Table 2.4: Differences in objective function for ±1% change in offered rate against the optimal value for midrate 14%.

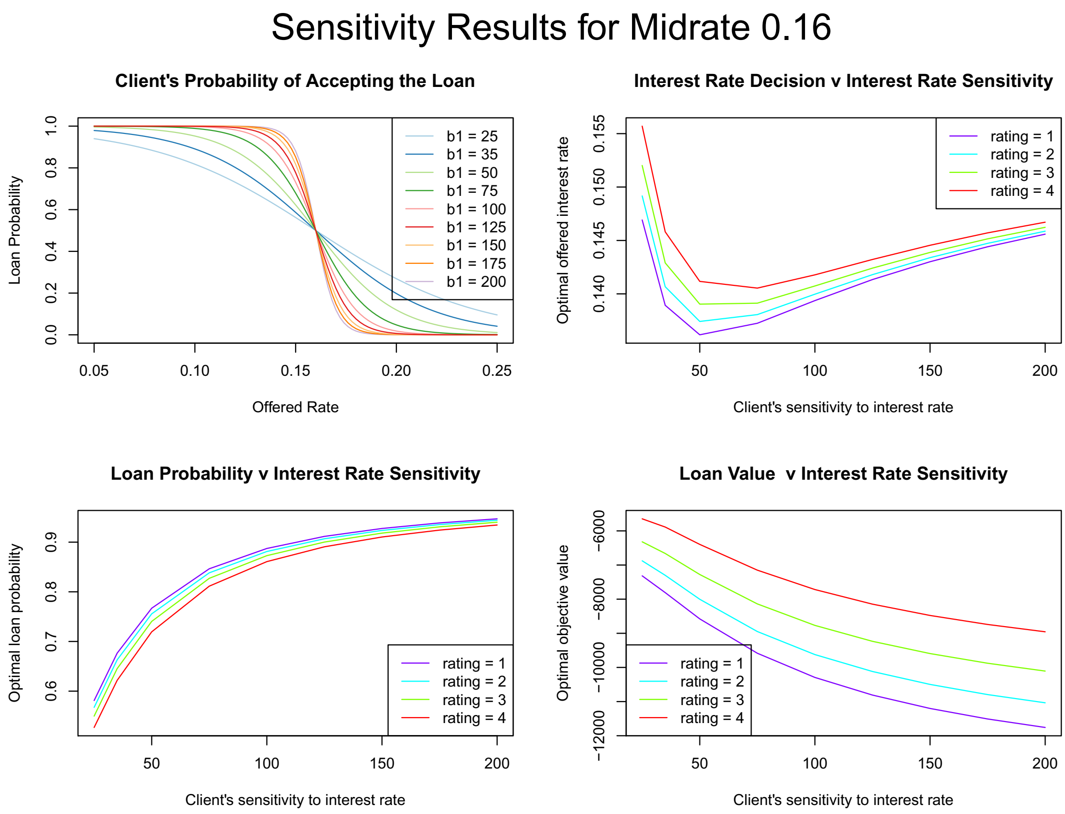

Figure 2.11: Sensitivity analysis results for midrate 0.16.

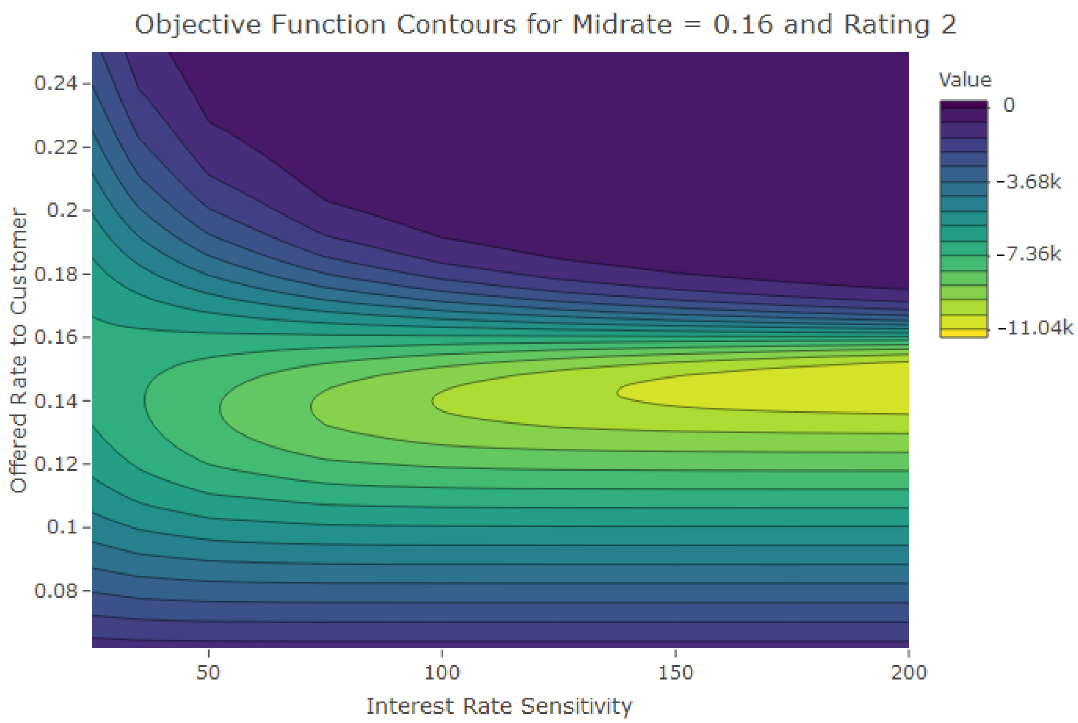

Figure 2.12: Contour plot constructed from objective values of the program when fixing offered interest rate r for different values of interest rate sensitivity b_1.

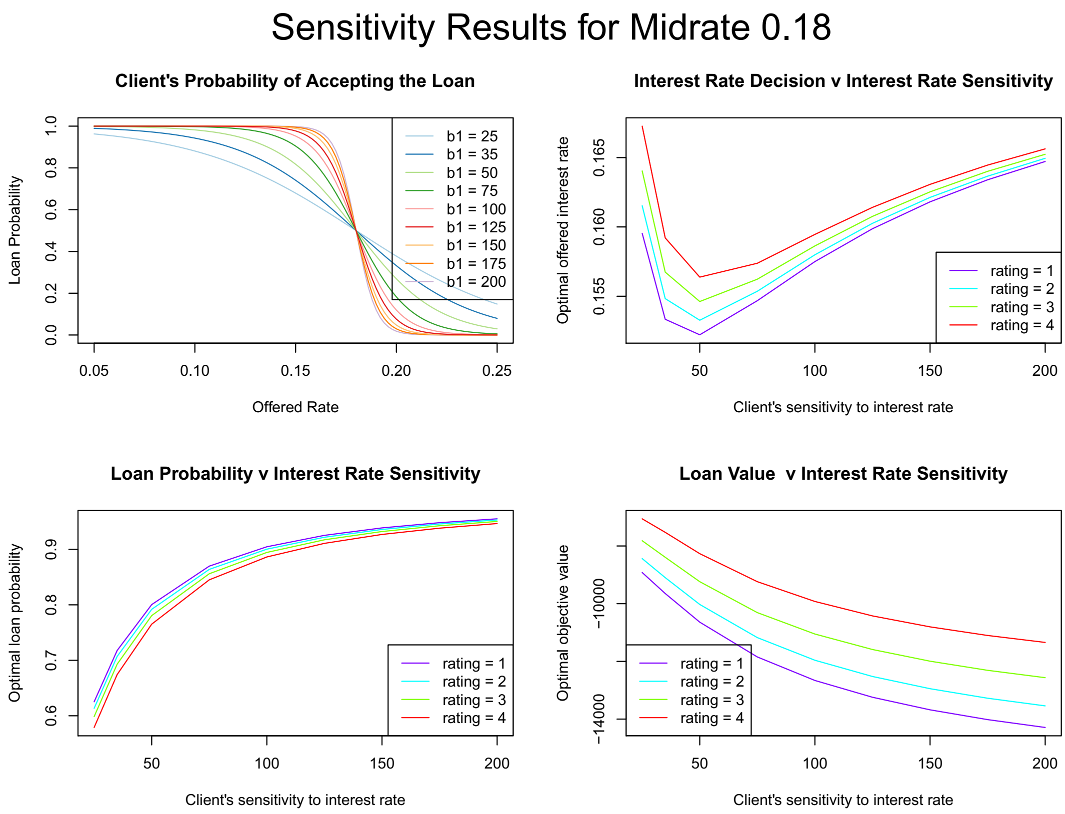

Figure 2.13: Sensitivity analysis results for midrate 0.18.

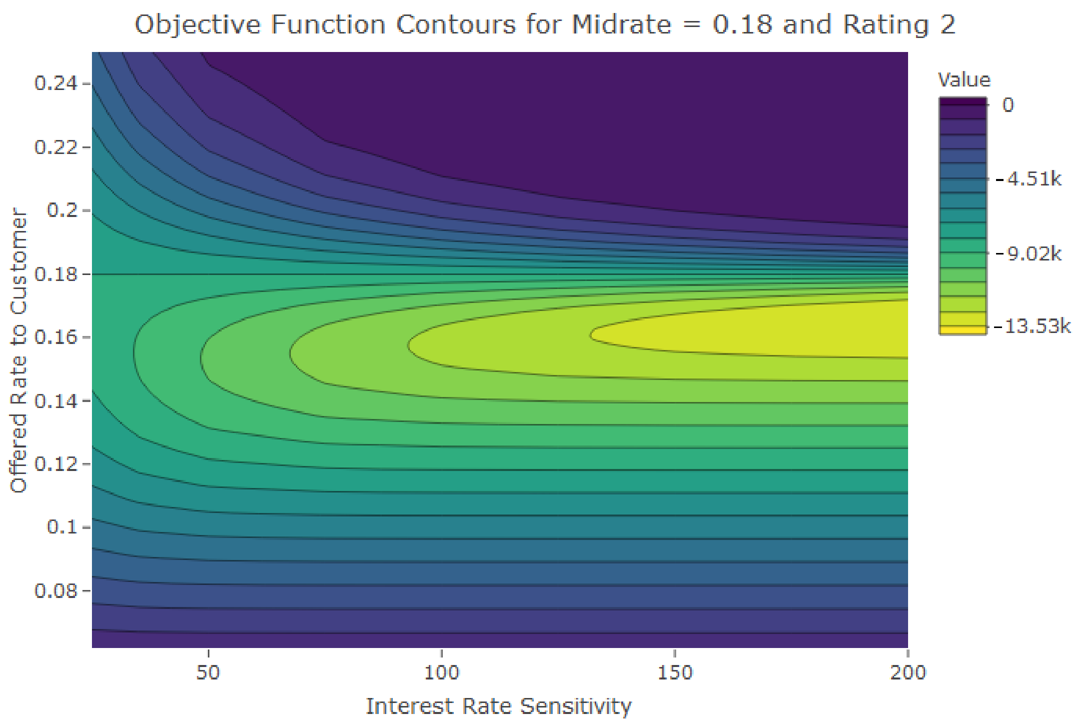

Figure 2.14: Contour plot constructed from objective values of the program when fixing offered interest rate r for different values of interest rate sensitivity b_1.

	$b_1 \downarrow$	Absolute [CZK]				Relative [%]			
Rating $\rightarrow$		1	2	3	4	1	2	3	4
	25	93	90	87	82	1.3	1.3	1.4	1.4
	35	147	144	139	132	1.9	2	2.1	2.2
	50	228	224	218	209	2.7	2.8	3	3.3
	75	349	345	337	327	3.6	3.9	4.1	4.6
-1%	100	452	447	440	427	4.4	4.6	5	5.5
	125	540	535	526	512	5	5.3	5.7	6.3
	150	615	609	599	583	5.5	5.8	6.2	6.9
	175	679	673	662	645	5.9	6.2	6.7	7.4
	200	734	728	716	699	6.2	6.6	7.1	7.8
	25	89	87	82	77	1.2	1.3	1.3	1.4
	35	149	145	138	129	1.9	2	2.1	2.2
	50	253	246	236	221	3	3.1	3.2	3.5
	75	456	444	428	404	4.8	5	5.3	5.6
+1%	100	693	677	652	618	6.7	7	7.4	8
	125	967	944	911	863	8.9	9.3	9.9	10.6
	150	1281	1251	1205	1144	11.4	11.9	12.6	13.5
	175	1637	1599	1541	1459	14.2	14.8	15.6	16.7
	200	2040	1992	1916	1811	17.3	18	19	20.2

Table 2.5: Differences in objective function for ±1% change in offered rate against the optimal value for midrate 16%.

Rating →	$b_1 \downarrow$	Absolute [CZK] 1	2	3	4	Relative [%] 1	2	3	4
-1%	25	100	98	95	90	1.1	1.2	1.2	1.3
	35	157	154	149	143	1.6	1.7	1.8	1.9
	50	239	236	230	222	2.2	2.4	2.5	2.7
	75	361	357	351	340	3.1	3.2	3.4	3.7
	100	465	461	453	441	3.7	3.9	4.1	4.4
	125	554	549	539	526	4.2	4.4	4.7	5
	150	630	624	614	598	4.6	4.8	5.1	5.5
	175	695	690	678	661	5	5.2	5.5	6
	200	751	745	733	715	5.3	5.5	5.8	6.3
+1%	25	99	96	92	87	1.1	1.1	1.2	1.2
	35	164	159	153	145	1.7	1.8	1.8	1.9
	50	274	268	259	246	2.6	2.7	2.8	3
	75	489	479	464	443	4.1	4.3	4.5	4.8
	100	740	725	704	672	5.8	6.1	6.4	6.8
	125	1029	1011	982	938	7.8	8.1	8.5	9
	150	1364	1340	1298	1243	10	10.4	10.8	11.5
	175	1748	1713	1664	1589	12.5	12.9	13.5	14.3
	200	2186	2143	2073	1979	15.3	15.8	16.5	17.4

Table 2.6: Differences in objective function for ±1% change in offered rate against the optimal value for midrate 18%.

3. Stress Testing via Contamination

The final chapter of this thesis discusses an extension of a contamination approach of Dupačová [1986, 1996] for decision-dependent randomness stochastic programs. We believe that this methodology can have great applications in financial industry, as it is suitable for stress-testing. That is something which is frequently required by market regulators as it helps to assess the ability of companies to cope with extreme events. The results which we present in this chapter have been submitted for publication in Kopa and Rusý [2021a].

This chapter is structured as follows. In Section 3.1, we review current work on contamination, which is focused on the stochastic programs with exogenous randomness. Specifically, in Section 3.1.1, we look at the results for the case when the set of feasible decisions is not dependent on the probability distribution and in Section 3.1.2 for cases with probability-dependent set of feasible decisions. There, we also show that the lower bound developed by Dupačová and Kopa [2012] can be improved. The Section 3.2 is devoted to the programs with decision-dependent randomness, where we define the meaning of contamination for this class of models. Further in Section 3.2.1 and Section 3.2.2 we formulate lower and upper bounds for models without and with decision-dependent randomness feasibility set, respectively. There, we specify all the assumptions under which the bounds can be applied. Thereafter, in Section 3.3, we apply the developed bounds on the asset-liability management stochastic program introduced in the second chapter of the thesis. In Section 3.3.1, We contaminate a "good" customer with a "bad" customer and show that contamination can significantly help with assessing the effect of a quality of the customer on the profit of the company. Finally, in Section 3.3.2, we extend the initial model for a CVaR constraint to show the functionality of the bounds for the decision-dependent randomness feasibility set as well.

3.1 Stochastic Programs with Exogenous Randomness

In this section, we briefly recall the main results developed in the field of contamination of stochastic programs. We consider a model

$$\begin{aligned} \min_{\mathbf{x}\in\mathcal{X}} \ & F_0(\mathbf{x}, P) \\ \text{s.t.} \ & F_j(\mathbf{x}, P) \leq 0, \qquad j = 1, \ldots, J, \end{aligned} \tag{3.1}$$

where $\mathcal{X} \subset \mathbb{R}^n$ is a closed non-empty convex set, P is a probability distribution of a random vector ω and $F_j\Big(\cdot, \cdot\Big) : \mathcal{X} \times \mathcal{P} \to \mathbb{R}, j = 0, \ldots, J$, are functions. The set $\mathcal{P}$ is some convex set of probability measures, $P \in \mathcal{P}$. We also assume that a solution of (3.1) exists. Note that we consider all the constraints which do not depend on P to be included in the definition of $\mathcal{X}$

When considering contamination, we perturb the original distribution P of a random vector ω by some contamination distribution $Q \in \mathcal{P}$. This technique, introduced by Dupačová [1986, 1996], became largely popular for its interpretation. Usually, the distribution of random vector ω is approximated. Hence it is important to study the stability of the solution and optimal objective value with respect to changes in P. Contamination can be also viewed as a stress-testing technique, where Q plays the role of the stress-testing distribution.

We introduce parameter $t, t \in [0,1]$ and define $P_t = (1-t)P + tQ$, $F_{Q,j}(\mathbf{x}, t) = F_j(\mathbf{x}, P_t), j = 0, \dots, J$ and the contaminated stochastic program as

$$\begin{aligned} \min_{\mathbf{x} \in \mathcal{X}} \ & F_{Q,0}(\mathbf{x}, t) \\ \text{s.t. } & F_{Q,j}(\mathbf{x}, t) \leq 0, \qquad j = 1, \dots, J. \end{aligned} \tag{3.2}$$

We denote $\mathcal{X}_Q(t), \mathcal{X}^*_Q(t)$ and $\varphi_Q(t)$ the set of feasible and optimal solutions and the optimal value function of a stochastic program (3.2) with a contamination parameter t, respectively. The main question we aim to answer is if we can bound the value function $\varphi_Q(t)$, so we can make conclusions on the values of the contaminated stochastic programs without solving them for all t.

3.1.1 Fixed Set of Feasible Decisions

First, let us consider the simplest case of (3.1), when no constraints depend on P and consider $t_1, t_2, \lambda \in [0,1]$. Provided that F_0 is concave in P (so $F_{Q,0}$ is concave in t), we obtain

$$\begin{aligned} \varphi_Q(\lambda t_1 + (1-\lambda)t_2) &\geq \min_{\mathbf{x} \in \mathcal{X}} \lambda F_{Q,0}(\mathbf{x}, t_1) + (1-\lambda) F_{Q,0}(\mathbf{x}, t_2) \\ &\geq \lambda \varphi_Q(t_1) + (1-\lambda)\varphi_Q(t_2), \end{aligned}$$

which means that the optimal value function $\varphi_Q(t)$ of (3.2) is concave in t. This provides the main argument under which Dupačová [1996] constructs global lower and upper bounds for the optimal value function $\varphi_Q(t)$:

$$(1-t)\varphi_Q(0) + t\varphi_Q(1) \leq \varphi_Q(t) \tag{3.3}$$

$$\varphi_Q(t) \leq \min\Big\{\varphi_Q(0) + t\varphi'_Q(0^+), \varphi_Q(1) - (1-t)\varphi'_Q(1^-)\Big\}. \tag{3.4}$$

The inequality (3.3) follows directly from the concavity of $\varphi_Q(t)$; on the contrary (3.4) relies on the fact that directional derivatives of concave functions bound the function from above. Under additional assumptions (most notably the convexity of $F_{Q,0}(\mathbf{x}, t)$ in $\mathbf{x}$ for all t), the directional derivatives can be calculated using the formula from Theorem 17 of Gol'shtein [1970] as

$$\varphi'_Q(0^+) = \min_{\mathbf{x} \in \mathcal{X}^*_Q(0)} \frac{d}{dt} F_{Q,0}(\mathbf{x}, 0^+).$$

A useful simplification of the above formula can be obtained in a case when $F_{Q,0}$ is linear in t. Then it holds

$$\frac{d}{dt} F_{Q,0}(\mathbf{x}, 0^+) = F_{Q,0}(\mathbf{x}, 1) - F_{Q,0}(\mathbf{x}, 0). \tag{3.5}$$

Even though the restriction on the set of feasible solutions is rather strict, these results have been applied in number of cases. For example, in investigating the resistance of scenario-based stochastic programs to changes in scenarios [Dupačová, 1996] or for stress-testing of [multi-stage] risk-averse stochastic programs [Dupačová and Polívka, 2007, Dupačová and Kozmík, 2015a,b].

3.1.2 Probability Distribution Dependent Set of Feasible Decisions

Next, let us consider the model (3.1) where the constraints depend on the probability distribution P. Results for this particular case were developed in Dupačová and Kopa [2012].

Let us denote

$$F_Q(\mathbf{x}, t) = \max_{j=1,\dots,J} \Big\{ F_{Q,j}(\mathbf{x}, t) \Big\}.$$

If $F_Q(\mathbf{x}, t)$ is concave in t and $F_{Q,0}(\mathbf{x}, t)$ does not depend on t (e.g. objective function is independent of the probability distribution), then it can be shown [Dupačová and Kopa, 2012] that $\varphi_Q(t)$ is quasi-concave on $[0, 1]$. From there, Dupačová and Kopa [2012] develop a lower bound

$$\varphi_Q(t) \geq \min\{\varphi_Q(0), \varphi_Q(1)\}.$$

The above bound is valid also when we substitute the condition that $F_Q(\mathbf{x}, t)$ is concave in t by showing that the inclusion $\mathcal{X}_Q(t) \subset \mathcal{X}_Q(0) \cup \mathcal{X}_Q(1)$ holds for each t. In case when $F_{Q,0}(\mathbf{x}, t)$ depends on t and it is concave in t, the above formula can be extended. Dupačová and Kopa [2012] show

$$\begin{aligned} \varphi_Q(t) &\geq \min_{\mathcal{X}_Q(t)} (1-t) F_{Q,0}(\mathbf{x}, 0) + t F_{Q,0}(\mathbf{x}, 1) \\ &\geq (1-t) \min \Big\{ \varphi_Q(0), \min_{\mathcal{X}_Q(1)} F_{Q,0}(\mathbf{x}, 0) \Big\} \\ &\quad + t \min \Big\{ \varphi_Q(1), \min_{\mathcal{X}_Q(0)} F_{Q,0}(\mathbf{x}, 1) \Big\}. \end{aligned} \tag{3.6}$$

Such a bound is linear, and to compute it, one needs to evaluate four stochastic programs. However, the assumptions on concavity of $F_Q(\mathbf{x}, t)$ seems very difficult to fulfil, especially when considering cases where $J > 1$.

Next, we show that this result can be improved in two different ways. First, let us formulate and prove the following lemma.

Lemma 1. *Let $F_Q(\mathbf{x}, t)$ be a quasi-concave function in t, then it holds $\mathcal{X}_Q(t) \subset \mathcal{X}_Q(0) \cup \mathcal{X}_Q(1)$.*

Proof. Using that $F_Q(\mathbf{x}, t) \geq \min\{F_Q(\mathbf{x}, 0), F_Q(\mathbf{x}, 1)\}$, we directly obtain

$$\begin{aligned} \mathcal{X}_Q(t) &= \{x : F_Q(\mathbf{x}, t) \leq 0\} \\ &\subset \{x : \min\{F_Q(\mathbf{x}, 0), F_Q(\mathbf{x}, 1)\} \leq 0\} \\ &= \mathcal{X}_Q(0) \cup \mathcal{X}_Q(1). \end{aligned}$$

□

As all concave functions are quasi-concave, we can relax the concavity assumption to requiring $F_Q(\mathbf{x}, t)$ to be quasi-concave in t only. Moreover, due to the definition of $F_Q(\mathbf{x}, t)$, this relaxation can be crucial in being able to show this assumption and use the bound. We note that each monotone function is quasi-concave, and the *max* function of monotone functions of the same direction (either all non-increasing or non-decreasing) is quasi-concave.

Apart from relaxing one of the main assumptions of Dupačová and Kopa [2012], Theorem 1, we can also improve the lower bound. First, let us formulate the following lemma.

Lemma 2. *Let $a, b, c, d \in \mathbb{R}$ and $t \in [0, 1]$. Then it holds*

$$\min\Big\{(1-t)a + tb, (1-t)c + td\Big\} \geq (1-t)\min(a, c) + t\min(b, d).$$

Proof. Using inequalities of type $a \geq \min(a, c)$ we obtain

$$\begin{aligned}\min\Big\{(1-t)a + tb, (1-t)c + td\Big\} \geq \Big\{&(1-t)\min(a, c) + t\min(b, d),\\ &(1-t)\min(a, c) + t\min(b, d)\Big\}\\ = &(1-t)\min(a, c) + t\min(b, d).\end{aligned}$$

□

In the following theorem, we prove a tighter lower bound.

Theorem 3. *For model* (3.2)*, let the max function $F_Q(\mathbf{x}, t)$ be quasi-concave in t and let $F_{Q,0}(\mathbf{x}, t)$ be concave in t. Then it holds that*

$$\begin{aligned}\varphi_Q(t) \geq \min\Big\{&(1-t)\varphi_Q(0) + t\min_{\mathcal{X}_Q(0)} F_{Q,0}(\mathbf{x}, 1),\\ &(1-t)\min_{\mathcal{X}_Q(1)} F_{Q,0}(\mathbf{x}, 0) + t\varphi_Q(1)\Big\}\end{aligned} \tag{3.7}$$

and hence (3.7) *is a lower bound for the optimal value function $\varphi_Q(t)$. Moreover, this lower bound* (3.7) *is tighter than lower bound* (3.6).

Proof. We proceed as follows

$$\begin{aligned}\varphi_Q(t) &\geq \min_{\mathcal{X}_Q(0)\cup\mathcal{X}_Q(1)} F_{Q,0}(\mathbf{x}, t)\\ &= \min\Big\{\min_{\mathcal{X}_Q(0)} F_{Q,0}(\mathbf{x}, t), \min_{\mathcal{X}_Q(1)} F_{Q,0}(\mathbf{x}, t)\Big\}\\ &\geq \min\Big\{(1-t)\varphi_Q(0) + t\min_{\mathcal{X}_Q(0)} F_{Q,0}(\mathbf{x}, 1),\\ &\qquad (1-t)\min_{\mathcal{X}_Q(1)} F_{Q,0}(\mathbf{x}, 0) + t\varphi_Q(1)\Big\}.\end{aligned}$$

The first inequality stems from the fact that $\mathcal{X}_Q(t) \subset \mathcal{X}_Q(0) \cup \mathcal{X}_Q(1)$, which is shown in Lemma 1. The second one is a basic operation with mathematical programs and finally, the third one is a consequence of applying lower bound (3.3) to contaminated stochastic programs with fixed decision set. Hence, we conclude that (3.7) is a lower bound.

Moreover, denoting $a = \varphi_Q(0), b = \min_{\mathcal{X}_Q(0)} F_{Q,0}(\mathbf{x}, 1), c = \min_{\mathcal{X}_Q(1)} F_{Q,0}(\mathbf{x}, 0)$ and $d = \varphi_Q(1)$ and using Lemma 2, we get that lower bound (3.7) is never lower than the lower bound (3.6). $\square$

The computational complexity of the new lower bound (3.7) is the same as (3.6), as it also requires calculation of four stochastic programs. The resulting bound is then piece-wise linear and is equal to (3.6) when $a \leq c$ and $b \leq d$ or when both $a \geq c$ and $b \geq d$ (here we use notation from the proof of Theorem 3).

An alternative lower bound can be developed in case of $F_{Q,0}(\mathbf{x}, t)$ being quasi-concave in t instead of concave in t. Under this setting, we can show the following corollaries.

Corollary. For model (3.2), let $J = 0$ (feasibility set is independent of P) and let $F_{Q,0}(\mathbf{x}, t)$ be quasi-concave in t. Then $\varphi_Q(t)$ is quasi-concave in t, so it holds that

$$\varphi_Q(t) \geq \min\Big\{\varphi_Q(0), \varphi_Q(1)\Big\}. \tag{3.8}$$

Proof. Consider $t_1, t_2, \lambda \in [0, 1]$, provided that F_0 is quasi-concave in P, we obtain

$$\begin{aligned}\varphi_Q(\lambda t_1 + (1-\lambda)t_2) &\geq \min_{\mathbf{x}\in\mathcal{X}} \Big\{ \min\Big\{F_{Q,0}(\mathbf{x}, t_1), F_{Q,0}(\mathbf{x}, t_2)\Big\}\Big\} \\ &= \min\{\varphi_Q(t_1), \varphi_Q(t_2)\},\end{aligned}$$

hence $\varphi_Q(t)$ is quasi-concave and (3.8) follows. $\square$

Corollary. For model (3.2), let the *max* function $F_Q(\mathbf{x}, t)$ be quasi-concave in t and let $F_{Q,0}(\mathbf{x}, t)$ be quasi-concave in t. Then it holds that

$$\varphi_Q(t) \geq \min\Big\{\varphi_Q(0), \min_{\mathcal{X}_Q(0)} F_{Q,0}(\mathbf{x}, 1), \min_{\mathcal{X}_Q(1)} F_{Q,0}(\mathbf{x}, 0), \varphi_Q(1)\Big\}. \tag{3.9}$$

Proof. By applying similar steps as in Theorem 3 and using the consequence of quasi-concavity of $F_{Q,0}(\mathbf{x}, t)$ in t given in (3.8), we prove (3.9). $\square$

Such bounds are obviously less strict then the ones developed with the assumption of concavity of $F_{Q,0}(\mathbf{x}, t)$, however, they are applicable to a wider class of stochastic programs.

The case of upper bound is more complicated because the function $\varphi_Q(t)$ is no longer concave. This allows to get a global upper bound only under special assumptions such as linearity of $F_{Q,j}(\mathbf{x}, t)$ in t for all j, and convexity $F_{Q,j}(\mathbf{x}, t)$ in $\mathbf{x}$. See Dupačová and Kopa [2012] for more details. On the other hand, Dupačová and Kopa [2012] show that under some second order differentiability conditions, it is possible to prove the existence of a local upper bound.

Theorem 4. *Let the model* (3.2) *be a twice differentiable program,* $\mathbf{x}_0^* \in \mathcal{X}_Q^*(0)$ *its optimal solution and* $\varphi_Q(0)$ *its optimal value. Assume that at* $\mathbf{x}_0^*$ *linear independence, the strict complementary conditions and the second-order sufficient conditions are satisfied. Then, there exists* $t_0 > 0$*, such that* $\forall t \in [0, t_0]$*, the optimal value function* $\varphi_Q(t)$ *is concave and the local upper contamination bound is given by*

$$\varphi_Q(t) \leq \varphi_Q(0) + t\varphi_Q'(0^+), \quad \forall t \in [0, t_0]. \tag{3.10}$$

In cases with convex problems, directional derivative $\varphi'_Q(0^+)$ can be again calculated from the Theorem 17 of Gol'shtein [1970].

3.2 Stochastic Programs with Endogenous Randomness

Next, we move from the stochastic programs with exogenous uncertainty to the class of stochastic programs with decision-dependent randomness. That is defined by the fact that the underlying probability distribution depends on the decision vector. In general, we formulate a model

$$\begin{aligned} \min_{\mathbf{x}\in\mathcal{X}} \quad & F_0\Big(\mathbf{x}, P(\mathbf{x})\Big) \\ \text{s.t.} \quad & F_j\Big(\mathbf{x}, P(\mathbf{x})\Big) \leq 0, \qquad j = 1, \dots, J, \end{aligned} \tag{3.11}$$

where $\mathcal{X} \subset \mathbb{R}^n$ is a closed non-empty convex set and $P(\mathbf{x})$ is a probability distribution of a random vector $\omega(\mathbf{x})$. We assume that $\omega(\mathbf{x})$ and $P(\mathbf{x})$ are uniquely assigned to each $\mathbf{x} \in \mathcal{X}$. The set $\mathcal{P}$ is a set of probability measures, $\{P(\mathbf{x}), \mathbf{x} \in \mathcal{X}\} \subset \mathcal{P}$ and $F_j\Big(\cdot, \cdot\Big) : \mathcal{X} \times \mathcal{P} \to \mathbb{R}, j = 0, \dots, J$, are functions. We assume that a solution of (3.11) exists.

Let us compare the formulations in (3.1) and (3.11). The main difference is in the probability distribution; in (3.1), we have the same distribution for all decisions. On the contrary, in (3.11), the underlying distribution of the stochastic program can be different for each decision vector $\mathbf{x}$. This is important to realize when we define contamination for this class of problems. Most notably, we will not have a contamination distribution Q, but a contamination mapping $Q(\mathbf{x})$, which assigns the contamination distribution for each $\mathbf{x}$. The set $\{Q(\mathbf{x}), \mathbf{x} \in X\}$ then forms a contamination family of distributions.

We consider the contamination mapping

$$Q(\mathbf{x}) : \mathcal{X} \to \mathcal{P},$$

and define

$$P_t(\mathbf{x}) = (1 - t) \cdot P(\mathbf{x}) + t \cdot Q(\mathbf{x}),$$

which leads to a contaminated stochastic program

$$\begin{aligned} \min_{\mathbf{x}\in\mathcal{X}} \quad & F_0\Big(\mathbf{x}, P_t(\mathbf{x})\Big) \\ \text{s.t.} \quad & F_j\Big(\mathbf{x}, P_t(\mathbf{x})\Big) \leq 0, \qquad j = 1, \dots, J. \end{aligned} \tag{3.12}$$

We again denote $\mathcal{X}_Q(t), \mathcal{X}^*_Q(t)$ and $\varphi_Q(t)$ as the set of feasible and optimal solutions and the optimal value function of a program (3.12) with a parameter t. We implicitly assume that $P_t(\mathbf{x})$ is a probability distribution $\forall t \in [0, 1]$ and $\forall \mathbf{x} \in \mathcal{X}$.

The motivation behind the formulation of (3.12) stems from stress-testing routines frequently used in financial management. Practitioners often formulate their "base case" and their "stress-testing" case, and they optimize for best base case result subject to the stress-test result meeting some criteria. Our approach provides them which much more detailed description (sensitivity) of their problem to the stress-testing case, as we deal with continuous contamination of the base case, yet, it preserves the desired simplicity of a "single bad guy".

The problem of constructing the lower and upper bounds for the objective value $\varphi_Q(t)$ has now increased in complexity. For this reason, we will focus only on some tractable cases, for which these bounds can be proved. We split the problem into two separate sub-classes based on whether the feasibility set is fixed or with decision-dependent randomness.

3.2.1 Fixed Set of Feasible Decisions

Let us first consider the simplest case of (3.11), when no constraints depend on $P(\mathbf{x})$.

Theorem 5. *Let $\mathcal{X}$ be a fixed, non-empty, convex set and $\mathcal{P}$ be a set of probability measures. Let $F_0(\mathbf{x}, P) : \mathcal{X} \times \mathcal{P} \to \mathbb{R}$ be a concave function in P for all $\mathbf{x} \in \mathcal{X}$. Let $P(\mathbf{x}) : \mathcal{X} \to \mathcal{P}$ and $Q(\mathbf{x}) : \mathcal{X} \to \mathcal{P}$ be mappings and $P_t(\mathbf{x}) = (1-t)P(\mathbf{x}) + Q(\mathbf{x}), t \in [0,1]$. Then the optimal value function*

$$\varphi_Q(t) = \min_{\mathbf{x}\in\mathcal{X}} \; F_0\Big(\mathbf{x}, P_t(\mathbf{x})\Big), \quad t \in [0,1]$$

is concave with respect to t and hence lower bound (3.3) *and upper bound* (3.4) *apply.*

Proof. Let us consider $t_1, t_2, \lambda \in [0,1]$. We have

$$\begin{aligned}
\varphi_Q(\lambda t_1 + (1-\lambda)t_2) &= \min_{\mathbf{x}\in\mathcal{X}} F_0(\mathbf{x}, \lambda P_{t_1}(\mathbf{x}) + (1-\lambda)P_{t_2}(\mathbf{x})) \\
&\geq \min_{\mathbf{x}\in\mathcal{X}} \lambda F_0(\mathbf{x}, P_{t_1}(\mathbf{x})) + (1-\lambda)F_0(\mathbf{x}, P_{t_2}(\mathbf{x})) \\
&\geq \lambda\varphi_Q(t_1) + (1-\lambda)\varphi_Q(t_2).
\end{aligned}$$

The first inequality is a consequence of F_0 being concave in the second argument. Hence, we have that $\varphi_Q(t)$ is concave and lower bound (3.3) and upper bound (3.4) can be constructed. □

The complication in this case is that parameters $\mathbf{x}$ and t in F_0 are no longer separable and in order to be able to use the formulas of Gol'shtein [1970] for the directional derivative, we need to show the convexity of $F_0(\mathbf{x}, P_t(\mathbf{x}))$ in $\mathbf{x}$ for all t.

An alternative approach, which does not require convexity of F_0 in $\mathbf{x}$ is to approximate the derivative from below numerically, which would also result into lower bound. It holds that

$$\varphi'_Q(0^+) \geq \frac{\varphi_Q(t_0) - \varphi_Q(0)}{t_0}. \tag{3.13}$$

Choosing an arbitraryt_0, one can obtain an estimate of the directional derivative, which will ensure a valid upper bound for all $t \geq t_0$. An upper bound valid for $t \in [0, t_0]$ can be obtained for example as

$$\varphi_Q(t) \leq \varphi_Q(t_0) - (t_0 - t)\frac{\varphi_Q(t_1) - \varphi_Q(t_0)}{t_1 - t_0}, \qquad t \in [0, t_0], t_1 > t_0.$$

The classes of problems which meet requirements of Theorem 5 are mainly expectation-type objective functions and some other probability functionals, such as CVaR or mean absolute deviation from the median. Some of such functionals are introduced and discussed in Postek et al. [2016]. Consider for example

$$F_0(\mathbf{x}, P(\mathbf{x})) = \mathbb{E}_{P(\mathbf{x})} f_0(\mathbf{x}, \omega),$$

for some value function f_0. Such a function F_0 is then linear in P and hence concave. Assume that the distribution $P(\mathbf{x})$ is defined by a density $p(\mathbf{x}, \omega)$ with respect to some common probability measure $\mu(\omega)$. Then, we write

$$F_0(\mathbf{x}, P(\mathbf{x})) = \mathbb{E}_{P(\mathbf{x})} f_0(\mathbf{x}, \omega) = \int f_0(\mathbf{x}, \omega) \cdot p(\mathbf{x}, \omega) \mu(d\omega).$$

We see again that F_0 is linear in P. Contamination mapping $Q(\mathbf{x})$ is then defined by densities $q(\mathbf{x}, \omega)$, and density of $P_t(\mathbf{x})$ with respect to $\mu(\omega)$ is given by $p_t(\mathbf{x}, \omega) = (1-t)p(\mathbf{x}, \omega) + tq(\mathbf{x}, \omega)$ and as a consequence $\varphi_Q(t)$ is concave in t.

Another important case occurs when considering only finite discrete distributions, or in other words, when $\mu(\omega)$ has finite support. In that case, we consider scenarios $\Omega = \{\omega_1, \dots, \omega_K\}$; the distribution $P(\mathbf{x})$ is defined by probabilities $p(\mathbf{x}, \omega_k) \geq 0, k = 1, \dots, K, \mathbf{x} \in \mathcal{X}$ of taking scenario ω_k and the objective function is given as

$$F_0(\mathbf{x}, P(\mathbf{x})) = \sum_{k=1}^{K} p(\mathbf{x}, \omega_k) f_0(\mathbf{x}, \omega_k).$$

As with the contamination mapping, we consider $Q(\mathbf{x})$ to be defined by $q(\mathbf{x}, \omega_k) \geq 0, k = 1, \dots, K, \mathbf{x} \in \mathcal{X}$. Note here, that we require $Q(\mathbf{x})$ to also be defined on the set Ω. However, we allow the probabilities to be zero and hence the distributions $P(\mathbf{x})$ and $Q(\mathbf{x})$ can have completely different (but finite) support. This is important especially for stress-testing. Consequently $P_t(\mathbf{x})$ has probabilities

$$p_t(\mathbf{x}, \omega_k) = (1-t)p(\mathbf{x}, \omega_k) + tq(\mathbf{x}, \omega_k), \quad t \in [0, 1].$$

Hence, we face a contaminated stochastic program

$$\min_{\mathbf{x} \in \mathcal{X}} \quad \sum_{k=1}^{K} p_t(\mathbf{x}, \omega_k) f_0(\mathbf{x}, \omega_k), \tag{3.14}$$

where again the objective function is linear in t and hence the optimal value function is concave in t. We also see that even when $f_0(\mathbf{x}, \omega_k)$ is convex with respect to $\mathbf{x}$ for all ω_k, the product $p_t(\mathbf{x}, \omega_k) f_0(\mathbf{x}, \omega_k)$ does not have this property. Consequently, the Gol'shtein formulas for directional derivatives are not applicable.

However, in the case of $F_0(\mathbf{x}, P)$ being linear in P, as in the both abovementioned examples, an alternative upper bound can be constructed.

Theorem 6. *Let all the assumptions from Theorem 5 hold and let $F_0(\mathbf{x}, P)$ be linear in P. The following upper bound for the optimal value function then applies:*

$$\begin{aligned} \varphi_Q(t) \leq \min \Big\{ & \varphi_Q(0) + t\Big(F_0\Big(\mathbf{x}_0^*, Q(\mathbf{x}_0^*)\Big) - \varphi_Q(0)\Big), \\ & \varphi_Q(1) - (1-t)\Big(F_0\Big(\mathbf{x}_1^*, P(\mathbf{x}_1^*)\Big) - \varphi_Q(1)\Big)\Big\}, \end{aligned} \tag{3.15}$$

where $\mathbf{x}_0^* \in \mathcal{X}_Q^*(0), \mathbf{x}_1^* \in \mathcal{X}_Q^*(1)$ are arbitrary optimal solutions of the original and the fully contaminated problems.

Proof. Let us choose an arbitrary $\mathbf{x}_0^* \in \mathcal{X}_Q^*(0)$. Then we have

$$\begin{aligned}\varphi_Q(t) &= \min_{\mathbf{x}\in\mathcal{X}} F_0(\mathbf{x}, P_t(\mathbf{x}))\\ &\le F_0(\mathbf{x}_0^*, P_t(\mathbf{x}_0^*)) = F_0\Big(\mathbf{x}_0^*, (1-t)P(\mathbf{x}_0^*) + tQ(\mathbf{x}_0^*)\Big)\\ &= (1-t)F_0\Big(\mathbf{x}_0^*, P(\mathbf{x}_0^*)\Big) + tF_0\Big(\mathbf{x}_0^*, Q(\mathbf{x}_0^*)\Big)\\ &= (1-t)\varphi_Q(0) + tF_0\Big(\mathbf{x}_0^*, Q(\mathbf{x}_0^*)\Big)\\ &= \varphi_Q(0) + t\Big(F_0\Big(\mathbf{x}_0^*, Q(\mathbf{x}_0^*)\Big) - \varphi_Q(0)\Big).\end{aligned}$$

Comparing this to (3.5), one can see that the term $F_0\Big(\mathbf{x}_0^*, Q(\mathbf{x}_0^*)\Big) - \varphi_Q(0)$ is an upper approximation of the directional derivative, which meets the true value in the case of no decision-dependent randomness. We can also perform analogical manipulations from the fully contaminated problem, then we have that

$$\varphi_Q(t) \le \varphi_Q(1) - (1-t)\Big(F_0\Big(\mathbf{x}_1^*, P(\mathbf{x}_1^*)\Big) - \varphi_Q(1)\Big), \quad \mathbf{x}_1^* \in \mathcal{X}_Q^*(1).$$

Combining the two results, we obtain bound for the optimal value function $\varphi_Q(t)$ as in (3.15). To obtain such a bound, one needs to solve two programs and twice evaluate the program in the selected solutions. □

In the final part of this section, we will shortly describe the case when the cardinality of the set $\mathcal{P}$ is finite.

Corollary. Consider model (3.11) with $J = 0, F_0(\mathbf{x}, P)$ be concave in P and $\mathcal{P} = \{P_1, \dots, P_n\}, n \in \mathbb{N}$ and let there be a partition of $\mathcal{X} : \mathcal{X}_1, \dots, \mathcal{X}_n$ such that

1. $\mathcal{X}_i \cap \mathcal{X}_k = \emptyset, \ \forall i \neq k$
2. $\bigcup_{i=1}^n \mathcal{X}_i = \mathcal{X}$
3. $P(\mathbf{x}) = P_i, \ \forall \mathbf{x} \in \mathcal{X}_i$
4. $F_0(\mathbf{x}, P_i)$ is continuous in $\mathbf{x}$ on closure (clo) of $\mathcal{X}_i, \ \forall i$.

Let us also assume that the contamination distribution has the same structure as P, i.e. most notably that $Q(\mathbf{x}) = Q_i, \forall \mathbf{x} \in \mathcal{X}_i$ and that the optimal solution of (3.12) exists for each t. Then we can partition the contaminated stochastic program into n sub-programs and obtain

$$\varphi_Q(t) = \min_{\mathbf{x}\in\mathcal{X}} \ F_0\Big(\mathbf{x}, P_t(\mathbf{x})\Big) = \min_{i=1,\dots,n} \min_{\mathbf{x}\in\mathrm{clo}(\mathcal{X}_i)} \ F_0\Big(\mathbf{x}, P_{t,i}\Big),$$

where $P_{t,i} = (1-t)P_i + tQ_i$. For all i, there is a lower bound $L_i(t)$ and an upper bound $U_i(t)$ such that

$$L_i(t) \le \min_{\mathbf{x}\in\mathrm{clo}(\mathcal{X}_i)} \ F_0\Big(\mathbf{x}, P_{t,i}\Big) \le U_i(t)$$

and it holds that

$$\min_{i=1,\dots,n} L_i(t) \le \varphi_Q(t) \le \min_{i=1,\dots,n} U_i(t).$$

Proof. The fact that we can partition the program into n sub-programs stems from the assumption (iv) and from the fact that the optimal solution of the original problem exists. Since $F_0(\mathbf{x}, P)$ is concave in P, then for each of the n sub-programs, we can obtain the lower bound $L_i(t)$ and upper bound $U_i(t)$ as in (3.3) and (3.4). Taking minimum over all sub-programs on both sides, we obtain the final statement of the corollary. □

3.2.2 Decision Dependent Randomness Feasibility Set

We continue to allow the feasibility set depend on the probability distribution.

Theorem 7. *Let $\mathcal{X}$ be a fixed, non-empty, convex set and $\mathcal{P}$ be a set of probability measures. Let $P(\mathbf{x}) : \mathcal{X} \to \mathcal{P}$ and $Q(\mathbf{x}) : \mathcal{X} \to \mathcal{P}$ be mappings and $P_t(\mathbf{x}) = (1-t)P(\mathbf{x}) + tQ(\mathbf{x}), t \in [0,1]$. Moreover, let $F_j(\mathbf{x}, P) : \mathcal{X} \times \mathcal{P} \to \mathbb{R}, j = 0, \ldots, J$ be functions and $F(\mathbf{x}, P) = \max_{j=1,\ldots,J} \left\{F_j(\mathbf{x}, P)\right\}$. If $F(\mathbf{x}, P)$ is quasi-concave in P for each $\mathbf{x} \in \mathcal{X}$, then for the feasibility set $\mathcal{X}_Q(t) = \{\mathbf{x} \in \mathcal{X} : F(\mathbf{x}, P_t(\mathbf{x})) \leq 0\}$ of the contaminated problem* (3.12) *holds that*

$$\mathcal{X}_Q(t) \subset \mathcal{X}_Q(0) \cup \mathcal{X}_Q(1). \tag{3.16}$$

Proof. Let us choose arbitrary $t \in [0,1]$. Then we obtain

$$\begin{aligned}
\mathcal{X}_Q(t) &= \{\mathbf{x} \in \mathcal{X} : F(\mathbf{x}, P_t(\mathbf{x})) \leq 0\} \\
&= \{\mathbf{x} \in \mathcal{X} : F(\mathbf{x}, (1-t)P(\mathbf{x}) + tQ(\mathbf{x})) \leq 0\} \\
&\subset \{\mathbf{x} \in \mathcal{X} : \min\{F(\mathbf{x}, P(\mathbf{x})), F(\mathbf{x}, Q(\mathbf{x}))\} \leq 0\} \\
&= \{\mathbf{x} \in \mathcal{X} : F(\mathbf{x}, P(\mathbf{x})) \leq 0\} \cup \{\mathbf{x} \in \mathcal{X} : F(\mathbf{x}, Q(\mathbf{x})) \leq 0\} \\
&= \mathcal{X}_Q(0) \cup \mathcal{X}_Q(1).
\end{aligned}$$

□

Theorem 8. *Let all the assumptions from Theorem 7 hold and, moreover, let $F_0(\mathbf{x}, P) : \mathcal{X} \times \mathcal{P} \to \mathbb{R}$ be a concave function of P for all $\mathbf{x} \in \mathcal{X}$. Then the lower bound* (3.7) *for the objective value function $\varphi_Q(t), t \in [0,1]$ of the contaminated problem*

$$\begin{aligned}
\varphi_Q(t) = \min_{\mathbf{x} \in \mathcal{X}} \quad & F_0\Big(\mathbf{x}, P_t(\mathbf{x})\Big) \\
\text{s.t.} \quad & F_j\Big(\mathbf{x}, P_t(\mathbf{x})\Big) \leq 0, \qquad j = 1, \ldots, J.
\end{aligned}$$

applies.

Proof. The proof is identical to proof of Theorem 3 with the use of Theorem 7. □

Similar to the case with exogenous randomness, we rely on the fact that $F(\mathbf{x}, P)$ is quasi-concave in P. We recall the discussion below Lemma 1, where cases when this happens are discussed. The condition is fulfilled, for example, when we have only one risk-averse type or expectation-type constraint in the model.

If we assume a model where the constraints are of type

$$F_j(\mathbf{x}, P(\mathbf{x})) = \mathbb{E}_{P(\mathbf{x})} f_j(\mathbf{x}, \omega),$$

i.e. expectation-type constraints, and $Q(\mathbf{x})$ such that it holds for all $\mathbf{x} \in \mathcal{X}$ that

$$F_j(\mathbf{x}, P(\mathbf{x})) \leq F_j(\mathbf{x}, Q(\mathbf{x})) \ \forall j \quad \text{or} \quad F_j(\mathbf{x}, P(\mathbf{x})) \geq F_j(\mathbf{x}, Q(\mathbf{x})) \ \forall j,$$

then $F_{Q,j}(\mathbf{x}, t)$ are all linear in t and also for each $\mathbf{x}$ functions $F_{Q,j}(\mathbf{x}, t)$ are all either non-increasing or non-decreasing. Hence the max function

$$F_Q(\mathbf{x}, t) = \max_{j=1,\dots,J} F_{Q,j}(\mathbf{x}, t),$$

is quasi-concave in t. Consequently, we can apply Theorem 7 as well as the lower bound (3.7).

Finally, we also comment on how the bounds would look for the decision-dependent randomness where the set $\mathcal{P}$ has finite cardinality. There, we would be able to break the model into n smaller sub-problems, solve them individually, obtain global lower and upper bounds as from the case of contamination with exogenous randomness, and get global lower and upper bound for the master problem the same way as we did for fixed feasibility set for decision-dependent randomness stochastic programs.

The upper bound is again more complicated, as $\varphi_Q(t)$ is not concave. Therefore, we are left with two possibilities:

- construct local or global upper bound based on feasibility of some solution,
- or construct local bound based on directional derivative, similar to one discussed in Theorem 4.

For the first case, we need to add an additional assumption that $F(\mathbf{x}, P)$ is quasi-convex in P and $F_0(\mathbf{x}, P)$ is linear in P. We choose an arbitrary $\mathbf{x}_0^* \in \mathcal{X}_Q^*(0)$ such that, $\mathbf{x}_0^* \in \mathcal{X}_Q(t_0), t_0 > 0$. Then $\forall t \in [0, t_0]$, we have

$$F(\mathbf{x}_0^*, P_t(\mathbf{x}_0^*)) \leq \max\Big\{F(\mathbf{x}_0^*, P(\mathbf{x}_0^*)), F(\mathbf{x}_0^*, P_{t_0}(\mathbf{x}_0^*))\Big\} \leq 0$$

and hence $\mathbf{x}_0^*$ is feasible $\forall t \in [0, t_0]$. We can further continue similar to calculations of (3.15) to obtain

$$\begin{aligned} \varphi_Q(t) &\leq (1-t)\varphi_Q(0) + tF_0\Big(\mathbf{x}_0^*, P_{t_0}(\mathbf{x}_0^*)\Big) \\ &= \varphi_Q(0) + t\Big(F_0\Big(\mathbf{x}_0^*, P_{t_0}(\mathbf{x}_0^*)\Big) - \varphi_Q(0)\Big), \quad t \in [0, t_0]. \end{aligned} \tag{3.17}$$

This bound becomes global if $t_0 = 1$. It is also possible to use some feasible solution $\mathbf{x}_0 \in \mathcal{X}_Q(0)$ instead of the optimal solution $\mathbf{x}_0^* \in \mathcal{X}_Q^*(0)$. The only difference would be that such a bound would be possibly less tight for small t. On the contrary, it could perform better for larger values of t.

For the local bound based on the directional derivative, we need to switch to stability results for non-linear parametric programming. There, we would proceed in the same manner as Dupačová and Kopa [2012]. Under various assumptions on differentiability, strict complementarity conditions and uniqueness of optimal solution, we could conclude that the optimal value function $\varphi_Q(t)$ is concave on the interval $[0, t_0]$ for some $t_0 > 0$. Hence, a local upper bound (3.10) is valid. However, the verifiability of these assumptions is questionable in the complexity of our problem; moreover, for calculation of the directional derivatives, we would require $F_j(\mathbf{x}, P(\mathbf{x}))$ to be convex in $\mathbf{x}$. Therefore, such a result appears difficult to use.

3.3 Application on the Asset-Liability Management Model

Next, we move to a numerical example where we illustrate the use of contamination. This will be performed on a multi-stage decision-dependent randomness stochastic program, where a company is asked to provide a loan to a customer. That program consists of the following:

- a company offering to a customer a loan with interest rate r,
- a customer accepting the loan with probability $p(r)$.
- In case of the acceptance, the company enters the financial market and optimize its liability side of the loan, while observing
  - possible default or prepayment of the client,
  - evolution of market interest rates.

The relationship between the interest rate decision and the probability of accepting the loan is given by the equation

$$p(r) = \frac{\exp\{b_1(b_0 - r)\}}{1 + \exp\{b_1(b_0 - r)\}},$$

where we call parameter b_0 *midrate* and b_1 interest rate sensitivity. For midrate, it holds that $p(b_0) = 0.5$, so it describes the rate at which the customer is indifferent to accepting or rejecting the loan.

Furthermore, we have that customer's rating affects the probability of default and prepayment. This event can happen at any year of the loan duration. A customer with the best rating one is estimated to have low default probabilities, almost independent of r, and high prepayment probabilities. These become even higher when he is given higher interest rate r, as the customer becomes more likely to refinance such a loan in future. On the other hand, a customer with bad rating has higher default probabilities, which increase with increasing r, but low prepayment probabilities, which only slightly depend on r.

More details on the stochastic program and relationship between the decisions and random quantities are described in 2. There, the model is introduced step by step as well as in its concise formulation together with all quantities which enter the model. The only modification in this part is, that we consider the model formulation as it is a minimization problem, to be consistent with the results which were derived in this chapter.

There, note especially the first four equations in the model definition (2.17), which capture the endogeneity in the random variables induced by the initial interest rate decision r. The other equations form a usual asset-liability multi-stage stochastic program. With 6 stages ($K = 6$) and 100 interest rate scenarios, the non-linear program had 45 619 equations and 58 690 variables. It was written in GAMS and solved by CONOPT. Solution of one model took approximately minutes (5-10) on a standard laptop (Intel core i5 2.60 GHz, 8GB RAM).

3.3.1 Contamination

In the multi-stage model, we considered b_0 to be from a range of values $b_0 = \{0.1, 0.12, 0.14, 0.16, 0.18\}$. The interest rate sensitivity parameter then measures how much this probability changes when the offer deviates from midrate. There we tested values $b_1 = \{25, 35, 50, 75, 100, 150, 200\}$. Rating ρ was set to take on values $\rho = \{1, 2, 3, 4\}$, where 4 is the worst. The loan was set to take 5 years, with monthly repayments and decision stages once a year. Moreover, we denote $P(r; b_0, b_1, \rho)$ the probability distribution of the underlying multi-stage model for a client with parameters b_0, b_1, ρ and the interest-rate decision r.

The general idea of contamination is to approximate how the customer's characteristics influence the optimal value of the loan without really solving the program for all considered types of customers. We saw that there are three characteristics customers possess which affect the uncertainty in the program. We investigate these characteristics both one by one and jointly. We will denote the parameters of the base case customer as (b_0, b_1, ρ) and the parameters of the "contaminating" customer as $(\bar{b}_0, \bar{b}_1, \bar{\rho})$. The settings employed in our contamination study is summarized in Table 3.1. The dot in the corresponding row means that all analysed values of this parameter have been tested. For example, in the first case, when we contaminated sensitivity, we had a customer with midrate 0.16 and a certain rating ρ (one of $1, 2, 3, 4$). Then as the base case we selected interest rate sensitivity 150 and as the contaminating case we chose sensitivity 50. Hence, we had $(b_0, b_1, \rho) = (0.16, 150, \rho)$ and $(\bar{b}_0, \bar{b}_1, \bar{\rho}) = (0.16, 50, \rho), \rho \in \{1, 2, 3, 4\}$. The values are set so they represent "realistically good" and "realistically bad" client.

As we face a decision-dependent randomness stochastic program, where the feasibility set is independent of the probability distribution, the lower bound (3.3) and the upper bound (3.15) were employed.

To describe the results which we achieved, we enclose seven figures. In the first three pairs (Figures 3.1, 3.2, 3.3, 3.4, 3.5 and 3.6), we show the bounds computed for contamination in interest rate sensitivity ($b_1 = 150, \bar{b}_1 = 50$), midrate ($b_0 = 0.16, \bar{b}_0 = 0.12$) and rating ($\rho = 1, \bar{\rho} = 4$). In case of contamination in sensitivity and rating, from Figures 3.1 and 3.5, we can note that the optimal value function $\varphi_Q(t)$ is almost linear and that the bounds are able to capture such behaviour and create a tight space in which the optimal value function lies. From Figure 3.2, one can learn that the maximum difference between lower and upper bound is around 80, which is pretty good approximation given that the true value of the program is in thousands. A similar conclusion can be drawn from Figure 3.6,

Table 3.1: Parameters of original and contaminated distribution employed in the empirical study.

	$P(r; b_0, b_1, \rho)$			$Q(r; \bar{b}_0, \bar{b}_1, \bar{\rho})$		
case/param	b_0	b_1	ρ	$\bar{b}_0$	$\bar{b}_1$	$\bar{\rho}$
Sensitivity	0.16	150	·	0.16	50	·
Midrate	0.16	·	2	0.12	·	2
Rating	0.16	·	1	0.16	·	4
All	0.16	150	1	0.12	50	4

where the maximum difference is about 40. The surprising thing in this figure is that we have the highest uncertainty for the stochastic program with $b_1 = 25$ and $b_1 = 35$, which are also the stochastic programs with the lowest optimal values and the lowest difference between the original problem ($t = 0$) and the fully contaminated problem ($t = 1$).

We see a completely different story for the optimal value function when we contaminate midrate, as depicted in Figures 3.3 and 3.4. There, the function is concave, but highly non-linear. This is again captured by the bounds quite effec-

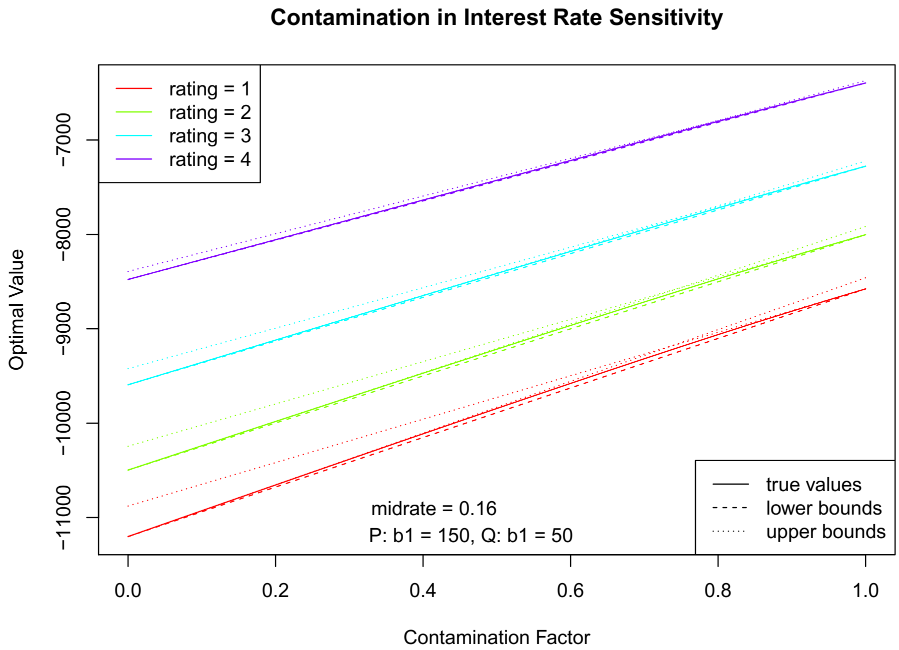

Figure 3.1: Lower and upper bounds for contamination of interest rate sensitivity at midrate 0.16.

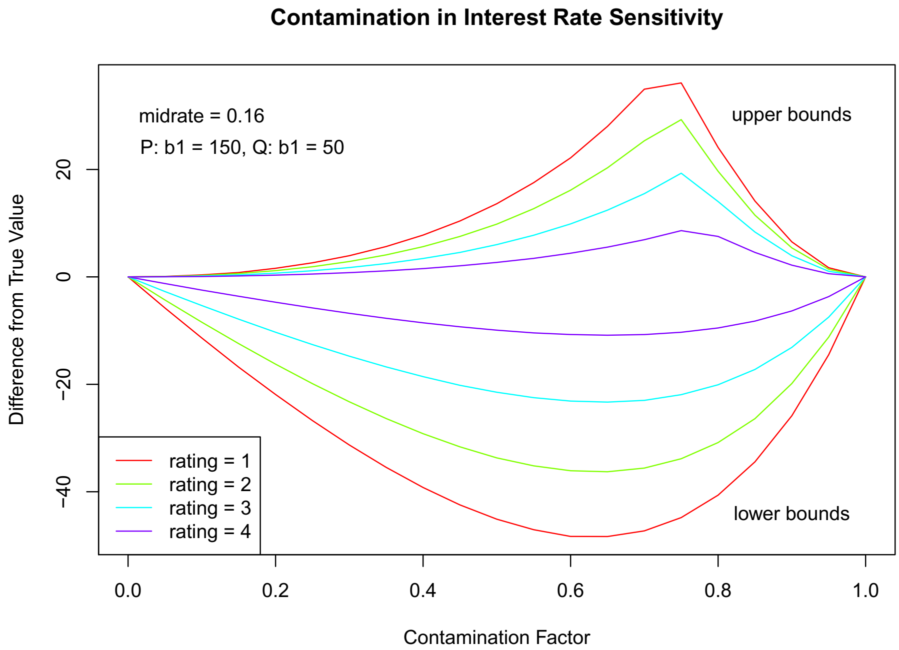

Figure 3.2: Differences of lower and upper bounds from true value for contamination of interest rate sensitivity at midrate 0.16.

tively, as the upper bounds are very tight and close to the optimal value function. What underlines the usefulness of this methodology is that if we needed tighter bounds, we could calculate the optimal solution at a point with highest difference between lower and upper bound and construct bounds from that solution too. It is quite obvious that if we had adopted this approach, we would obtain much tighter bounds in this case as well.

Finally, we looked at how the true values and lower and upper bounds differ when we modify all three characteristics of the customer jointly. This study is summarized in Figure 3.7. There, one can see a mixture of what we saw in

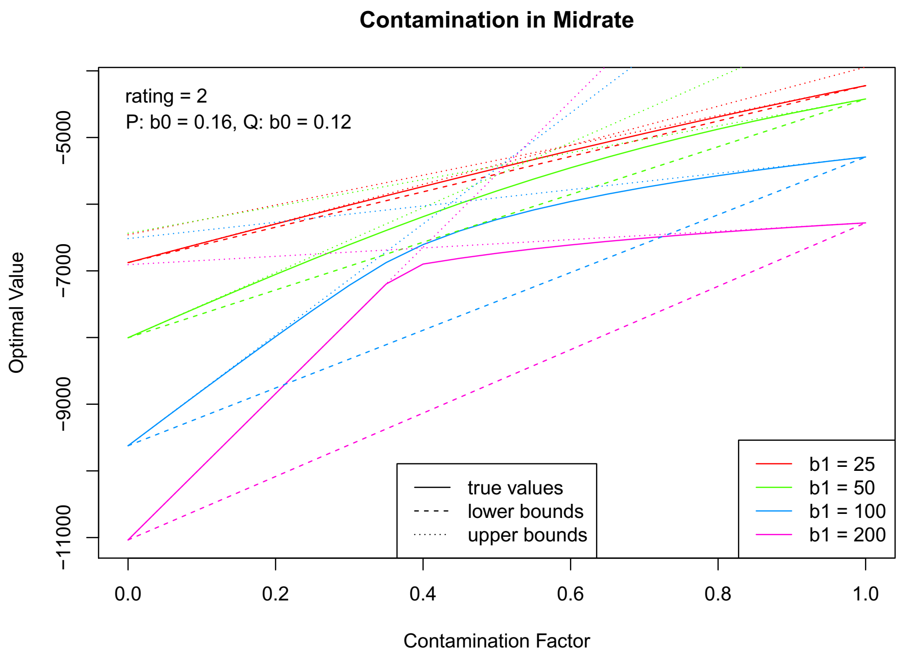

Figure 3.3: Lower and upper bounds for contamination of midrate at rating 2.

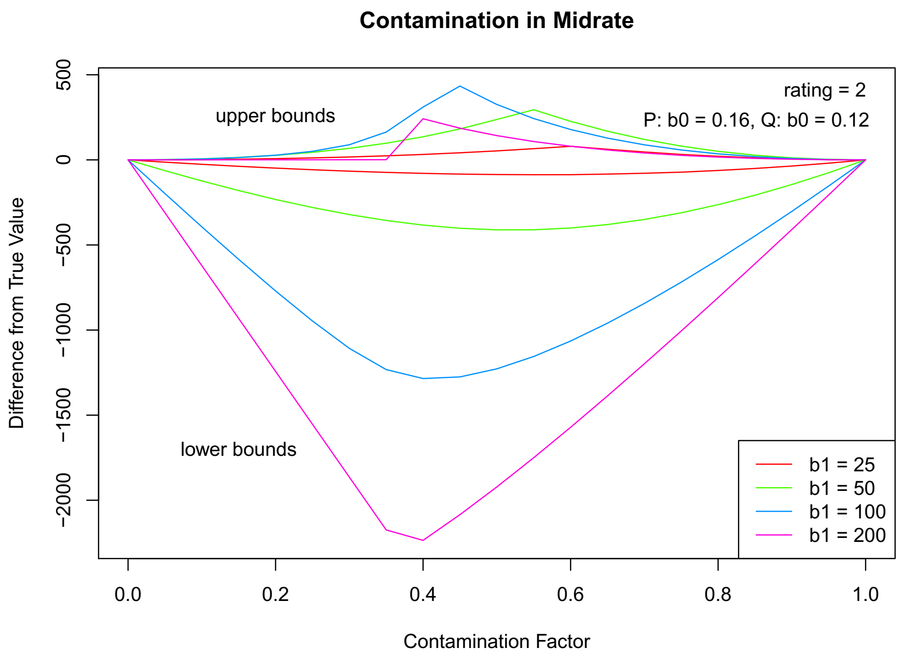

Figure 3.4: Differences of lower and upper bounds from true value for contamination of midrate at rating 2.

previous figures and fairly tight approximation of the optimal value function, especially on the edges of the definition set $[0, 1]$ of the contamination parameter t.

3.3.2 Decision-Dependent Randomness Feasibility Set

Further, we can extend the model (2.17) to illustrate the functionality of the contamination technique in the case with decision-dependent randomness feasibility set. To the initial model, we add a CVaR constraint, such that for given

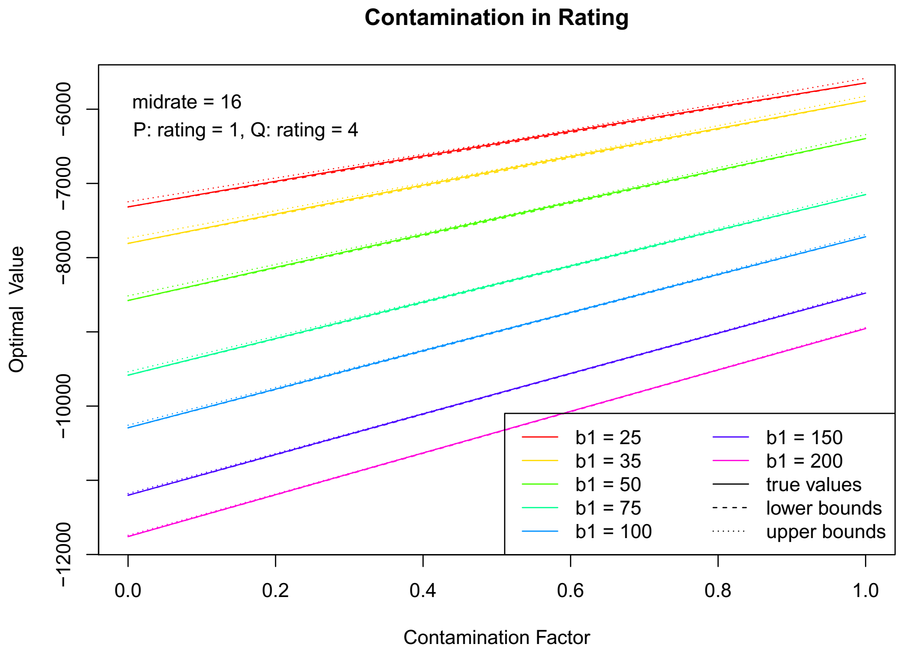

Figure 3.5: Lower and upper bounds for contamination of rating at midrate 0.16.

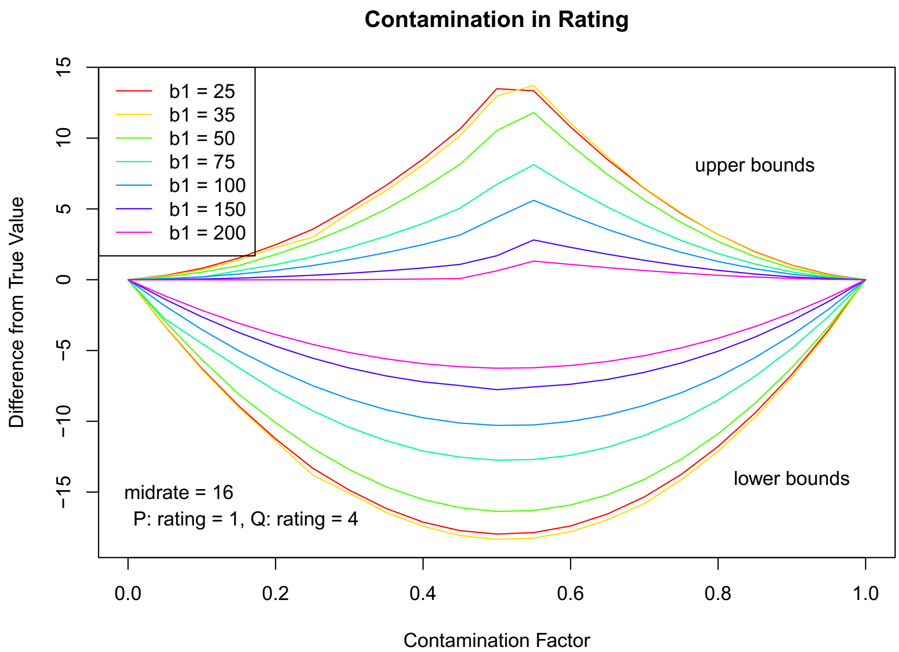

Figure 3.6: Differences of lower and upper bounds from true value for contamination of rating at midrate 0.16.

confidence level $0 < \alpha < 1$, we require

$$\mathrm{CVaR}_{\alpha}\Big(-V_{t_K}\Big) \leq 0, \tag{3.18}$$

where V_{t_K} denotes the value of the loan at the final stage of the program (see formulation in the appendix for the exact definition). The constraint means that the average loss in the $(1-\alpha)\%$ worst cases is lower than or equal to 0. Note that CVaR is concave with respect to the probability distribution [Dupačová and Polívka, 2007] and hence it meets the assumptions imposed in Theorem 7. The constraint is implemented into the model via Rockafellar and Uryasev [2000] as

$$\begin{aligned} z(s_K, e) \geq -V_{t_K}(s_K, e) - a, \quad z(s_K, e) \geq 0, & \qquad s_K \in S_K, e \in E, \\ a + \frac{1}{1-\alpha} \sum_{s_K \in S_K, e \in E} p(r) \cdot p(s_K, e, r) z(s_K, e) \leq 0, & \qquad a \in \mathbb{R}. \end{aligned}$$

For the analysis, we set the parameter $\alpha = 0.9$, so we limit the average return of the 10% worst scenarios. We apply the same contamination distribution as described in Table 3.1 in the All case, i.e. we contaminate all parameters at once. First, we show in Figure 3.8 the objective value function constructed from the solution of the contaminated models on a dense grid in $t \in [0, 1]$. There, with the help of the dashed line, which connects the first and the last point on the curve, one can notice that the optimal value function is not concave.

Next, we solve the models required to calculate the bound (3.7). We obtain the following values as the solution of the models:

$$\varphi_Q(0) = -4774.84,$$

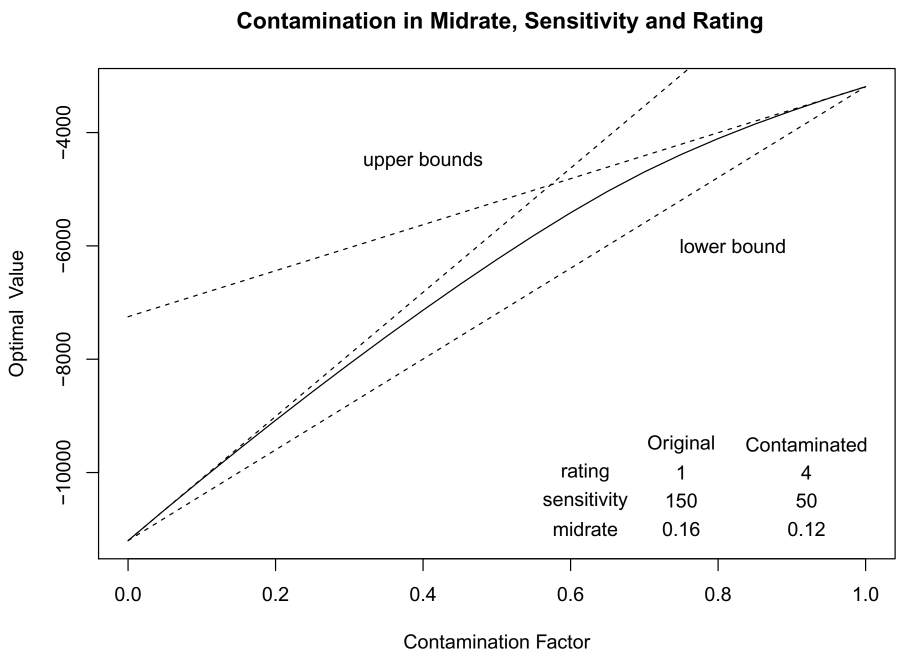

Figure 3.7: Lower and upper bounds for contamination of a customer in midrate, interest rate sensitivity and rating.

$$\min_{\mathcal{X}_Q(0)} F_{Q,0}(\mathbf{x}, 1) = -1007.01,$$

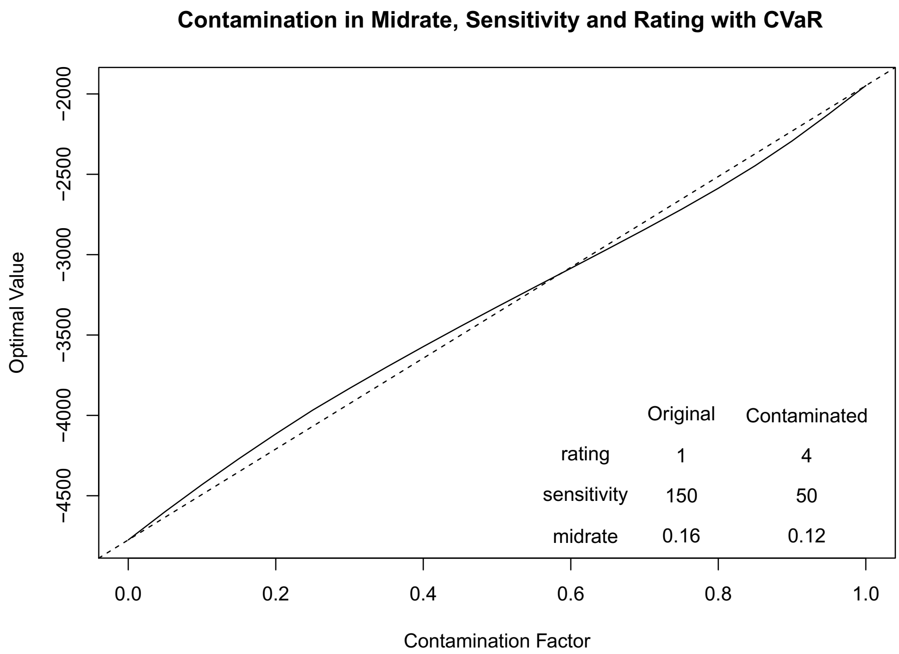

Figure 3.8: Optimal values of the program with CVaR for different levels of contamination. Dashed line is a segment connecting the first and the last point of the curve.

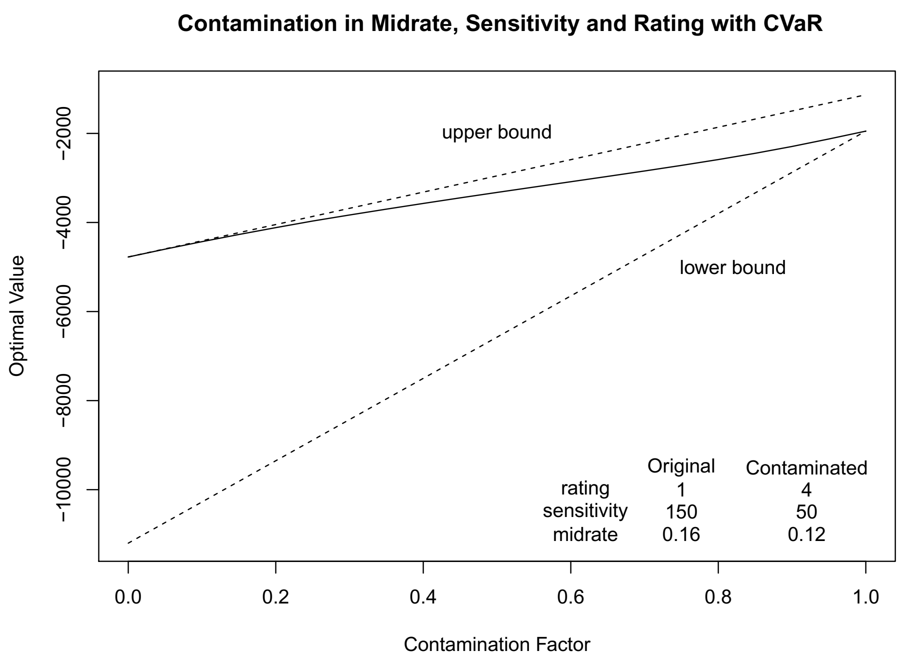

Figure 3.9: Upper bound for contamination of a client at midrate, interest rate sensitivity and rating of the CVaR model.

$$\min_{\mathcal{X}_Q(1)} F_{Q,0}(\mathbf{x}, 0) = -11201.66,$$
$$\varphi_Q(1) = -1948.34,$$

which, given that $\varphi_Q(0) \geq \min_{\mathcal{X}_Q(1)} F_{Q,0}(\mathbf{x}, 0)$ and $\min_{\mathcal{X}_Q(0)} F_{Q,0}(\mathbf{x}, 1) \geq \varphi_Q(1)$ gives us that

$$\varphi_Q(t) \geq -11201.66(1-t) - 1948.34t.$$

Finally, we also calculate the trivial upper bound (3.17), where the optimal solution of the non-contaminated problem was used as an estimate for the solution of the fully contaminated problem. However, note that we cannot theoretically prove the quasi-convexity property of the constraint in our example so the upper bound is only informative. Yet, numerical calculations show that the bound is a valid upper bound.

The results are showed graphically in Figure 3.9. We see that the lower bound is very loose, especially in the area around $t = 0$. That is down to the CVaR constraint for $t = 1$ not being active when optimizing the model with objective function for $t = 0$. Hence the optimal solution is the same as in the previous section when the CVaR constraint was not employed for the case of no contamination.

Conclusion

The main objective of this thesis was to extend the scientific knowledge in the area of stochastic optimization problems in finance. We have focused on three different aspects in the area. First, we developed a new method for scenario generation of interest rates. Second, we formulated, solved and analysed a decision-dependent asset-liability management stochastic program and third, we derived new results for contamination in decision-dependent randomness stochastic programs.

In this thesis we have suggested the maximum likelihood method for calibration of the one-factor short-rate models, arguably one of the most popular classes of yield curve models. In order to estimate the parameters we have introduced postcalibration errors and proposed to maximise the likelihood of yields in periods subsequent to the calibration time. Our estimation method uses a time series of yields to identify the model under both the real-world and risk-neutral measure. The identification of both measures facilitates the use of the model for risk and portfolio management as well as for derivatives pricing on illiquid markets. To demonstrate the method we have performed an extensive empirical analysis on the Hull-White model and EUR interest rate data. We have analysed the estimation results both in-sample and out-of-sample. We have compared the suggested estimation method to calibration methods, and illustrated the impact of the measure choice on the probability of default.

Certainly, there are natural next steps which further improve this method. These may include extensions of the suggested estimation method to multi-factor models, time-varying market prices of risk, or a heavy-tailed distribution of the postcalibration errors.

In the second part, we have considered an asset-liability management problem of a consumer loan, where, due to the possibility of the customer not accepting the loan and, upon acceptance, prepaying or defaulting on the loan, we formulate the problem as a non-linear multistage stochastic program with endogenous source of uncertainty. There, two groups of decisions appear: first, the initial fixed rate decision and second, the decisions associated with the asset-liability management policy. The fixed rate decision on the loan affects the future (uncertain) loan evolution and hence its value. The presented optimization problem allows the determination of such interest rate and optimal loan management taking all the contingencies into account through a set of conditional, decision-dependent scenario probabilities. We focused on the formulation of the problem in the theoretical part, where all the features of this problem, especially the decision-dependent randomness and its implementation into the program, were described.

The practical part was then devoted to the solutions of the program. First, we have investigated the performance of a unique optimal solution within the stochastic program's scenarios. The optimal strategy was to borrow only for the shortest time period as these loans are, in general, the cheapest. Such a strategy also allows great flexibility, which is beneficial, for example, when the loan is prepaid. As a drawback, it increases *interest rate risk.* The exposure to interest rate risk could possibly be addressed by the implementation of various risk constraints (e.g. a chance constraint [Telser, 1955], a Value-at-Risk constraint [J.P. Morgan Risk Metrics, 1995], a conditional Value-at-Risk constraint [Rockafellar and Uryasev, 2000, 2002] a second-order stochastic dominance constraint [Hadar and Russell, 1969, Dentcheva and Ruszczynski, 2003]) or application of robustness or contamination approaches [Dupačová and Kopa, 2012, 2014].

Moreover, we have implemented and solved the stochastic program under several parameter settings, capturing different customer's properties to determine what the implications of such a model would be on decisions made by the company. There, we noticed different actions on customers with different strengths in decision-dependent uncertainty. This was the case for both the changes in interest rate sensitivity and the changes in rating. This demonstrates that decision-dependent randomness needs to be considered in this problem and that the model which takes it into account produces strong results.

We were interested in the incurred losses which are caused by the company not offering the optimal interest rate. We found that such costs depend greatly on interest rate sensitivity. Moreover, we saw that it is more costly to offer a higher rate than a lower rate compared to the optimal rate. This is due to two reasons – a lower probability of accepting the loan hurts more than lower interest rate revenues and the default rate increases with the higher interest rate. Both effects are a consequence of decision-dependent uncertainty. This, again, underlines the need to capture the dependence between the offered interest rate and default probability correctly and not neglect the relationship.

In the final part of the thesis, we investigated the robustness of decision-dependent randomness stochastic programs via contamination. We extended the work of Noyan et al. [2018] and Luo and Mehrotra [2020], who defined some robustness techniques into the class of decision-dependent randomness stochastic programs, as they formulated the ambiguity sets for the min-max approach of Žáčková [1966]. Noyan et al. [2018] also constructed a distributionlly robust application of a practical problem of a single-machine scheduling. In our work, we focused on the contamination method, which employs a parametric programming approach. We extended results of Dupačová [1986, 1996] and Dupačová and Kopa [2012], who studied contamination in a class of stochastic programs with exogenous randomness. We also presented and proved various lower and upper bounds applicable for different types of decision-dependent stochastic programs and specifically commented on how they translate in the special case of a stochastic program with expectation-type objective function and constraints.

To illustrate the usefulness of contamination in financial practice, we applied the methodology to a real problem from financial industry. In a decision-dependent welfare maximization problem of a company issuing loans to customers formulated in the second part of the thesis, we saw that the company's profit from a loan is greatly affected by the customer's quality. We demonstrated that with the use of contamination, the company can very well estimate the effect of facing a "bad" customer compared to a "good" customer. This could be consequently of use when making a decision on their investment into credit-risk department. We believe that the developed theoretical results possess practical applicability, interpretation and solvability, which makes them preferable for stability and sensitivity analysis in decision-dependent randomness stochastic programs.

Bibliography

S. Ahmed. *Strategic Planning under Uncertainty: Stochastic Integer Programming Approaches*. PhD thesis, University of Illinois, 2000.

Y. Aït-Sahalia. Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach. *Econometrica*, 70(1):223–262, 2002.

A. Ang and M. Piazzesi. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics*, 50(4):745–787, 2003.

B. Basciftci, S. Ahmed, and S. Shen. Distributionally robust facility location problem under decision-dependent stochastic demand, 2021. URL `https://arxiv.org/abs/1912.05577`.

J. Birge and F. Louveaux. *Introduction to Stochastic Programming*. Springer Verlag, New York, 1997. ISBN 0-387-98217-5.

J. Blanchet and K. Murthy. Quantifying distributional model risk via optimal transport. *Mathematics of Operations Research*, 44(2):565–600, 2019.

J. F. Bonnans and A. Shapiro. Optimization Problems with Perturbations: A Guided Tour. *Society for Industrial and Applied Mathematics*, 40(2):228–264, 1998.

D. Brigo and F. Mercurio. *Interest Rate Models - Theory and Practice with Smile, Inflation and Credit (2nd ed. 2006)*. Springer Verlag, Berlin, 2001. ISBN 978-3-540-22149-4.

D. Broeders, A. Chen, and B. Koos. *An institutional evaluation of pension funds and life insurance companies*. De Nederlandsche Bank, 2009.

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

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

G. Consigli. Asset-liability management for individual investors. In S. A. Zenios and W. T. Ziemba, editors, *Handbook of Asset and Liability Management*, volume 2, pages 751–827. Elsevier, 2008.

G. Consigli and M. A. H. Dempster. Dynamic stochastic programming for asset-liability management. *Annals of Operations Research*, 81:131–162, 1998.

G. Consigli, V. Moriggia, and S. Vitali. Long-term individual financial planning under stochastic dominance constraints. *Annals of Operations Research*, 292: 973–1000, 2020. URL `https://doi.org/10.1007/s10479-019-03253-8`.

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

J. Cox, J. E. Ingersoll, and S. A. Ross. An intertemporal general equilibrium model of asset prices. *Econometrica*, 53:363–384, 1985b.

M. A. H. Dempster and A. M. Ireland. MIDAS: An expert debt management advisory system. In *Data, expert knowledge and decisions*, pages 116–127. Springer, 1988.

D. Dentcheva and A. Ruszczynski. Optimization with stochastic dominance constraints. *SIAM Journal on Optimization*, 14(2):548–566, 2003.

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

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

J. Dupačová and M. Kopa. Robustness in stochastic programs with risk constraints. *Annals of Operations Research*, 200(1):55–74, 2012.

J. Dupačová and M. Kopa. Robustness of optimal portfolios under risk and stochastic dominance constraints. *European Journal of Operational Research*, 234(2):434–441, 2014.

J. Dupačová and J. Polívka. Stress testing for var and cvar. *Quantitative Finance*, 7:411–421, 08 2007. doi: 10.1080/14697680600973323.

J. Dupačová. Optimization under exogenous and endogenous uncertainty. In *Proceedings of MME06*, 2006.

J. Dupačová. Stability in stochastic programming with recourse. Contaminated distributions. *Mathematical Programming Study*, 27:133–144, 1986.

J. Dupačová. Scenario-based stochastic programs: resistance with respect to sample. *Annals of Operations Research*, 64:21–38, 1996.

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

J. Dupačová and V. Kozmík. Structure of Risk-Averse Multistage Stochastic Programs. *OR Spectrum*, 37:559–582, 2015b.

J. Dupačová and J. Polívka. Stress testing fo VaR and CVaR. *Quantitative Finance*, 7:411–421, 2007.

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

P. M. Esfahani and D. Kuhn. Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming*, 171(1-2):115–166, 2018.

K. Fergusson. Asymptotics of bond yields and volatilities for extended vasicek models under the real-world measure. *Annals of Financial Economics*, 12(01): 1750005, 2017.

K. Fergusson. Fast maximum likelihood estimation of parameters for square root and bessel processes. *Studies in Nonlinear Dynamics & Econometrics*, 1 (ahead-of-print), 2020.

K. Fergusson and E. Platen. Application of maximum likelihood estimation to stochastic short rate models. *Annals of Financial Economics*, 10:27–53, 2015.

A. Geyer and W. T. Ziemba. The Innovest Austrian pension fund financial planning model InnoALM. *Operations Research*, 56(4):797–810, 2008.

V. Goel and I. E. Grossmann. A stochastic programming approach to planning of offshore gas field developments under uncertainty in reserves. *Computers and Chemical Engineering*, 28(8):1409–1429, 2004.

E. G. Gol'shtein. Vypukloje programmirovanije. *Elementy teoriji*, 1970. Translated in Theory of Convex Programming, translations of mathematical monographs. Vol. 36, American Mathematical Society, Providence 1972.

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

P. S. Hagan and G. West. Interpolation methods for curve construction. *Applied Mathematical Finance*, 13(2):89–129, 2006.

J. D. Hamilton. *Time Series Analysis*, volume 2. Princeton New Jersey, 1994.

J. D. Hamilton and J. C. Wu. Identification and estimation of gaussian affine term structure models. *Journal of Econometrics*, 168(2):315–331, 2012.

P. Harms, D. Stefanovits, J. Teichmann, and M. V. Wüthrich. Consistent recalibration of the discrete-time multifactor Vasiček model. *Risks*, 4(3):1–31, 2016.

P. Harms, D. Stefanovits, J. Teichmann, and M. V. Wüthrich. Consistent recalibration of yield curve models. *Mathematical Finance*, 28(3):757–799, 2018.

H. Held and D. L. Woodruff. Heuristics for multi-stage interdiction of stochastic networks. *Journal of Heuristics*, 11(5-6):483–500, 2005.

T. S. Ho and S.-B. Lee. Term structure movements and pricing interest rate contingent claims. *The Journal of Finance*, 41(5):1011–1029, 1986.

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

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

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

J. Hull and A. White. Pricing interest–rate derivative securities. *The Review of Financial Studies*, 3:573–592, 1990b.

J. Hull and A. White. Numerical procedures for implementing term structure models I: Single-factor models. *Journal of Derivatives*, 2(1):7–16, 1994.

J. Hull and A. White. Interest rate trees: extensions and applications. *Quantitative Finance*, 18(7):1199–1209, 2018.

J. Hull, A. Sokol, and A. White. Short-rate joint-measure models. *Risk*, pages 59–63, 2014.

S. Jain, P. Karlsson, and D. Kandhai. KVA, mind your P's and Q's! *Wilmott*, 2019(102):60–73, 2019.

F. Jamshidian. An exact bond option formula. *The Journal of Finance*, 44(1): 205–209, 1989.

R. Jiang and Y. Guan. Data-driven chance constrained stochastic program. *Mathematical Programming*, 158(1-2):291–327, 2016.

R. Jiang and Y. Guan. Risk-averse two-stage stochastic program with distributional ambiguity. *Operations Research*, 66(5):1390–1405, 2018.

T. Jonsbraten, R. Wets, and D. Woodruff. A class of stochastic programs with decision dependent random elements. *Annals of Operations Research*, 82:83–106, 1998.

S. Joslin, K. J. Singleton, and H. Zhu. A new perspective on gaussian dynamic term structure models. *The Review of Financial Studies*, 24(3):926–970, 2011.

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

P. Kall and J. Mayer. *Stochastic Linear Programming: Models, Theory and Computation*, volume 156. 01 2005.

D. H. Kim. Challenges in macro-finance modeling. *Federal Reserve Bank of St. Louis Review*, 91(5 Part 2):519–44, 2009.

K. Kladívko and T. Rusý. Maximum likelihood estimation of the Hull-White model. *Journal of Empirical Finance*, 2021. In review.

M. Kopa and T. Rusý. Robustness of stochastic programs with endogenous randomness via contamination. *European Journal of Operations Research*, 2021a. In review.

M. Kopa and T. Rusý. A decision-dependent randomness stochastic program for asset–liability management model with a pricing decision. *Annals of Operations Research*, 299:241–271, 2021b.

M. Kopa, V. Moriggia, and S. Vitali. Individual optimal pension allocation under stochastic dominance constraints. *Annals of Operations Research*, 260(1-2): 255–291, 2018.

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

D. Lando. *Credit Risk Modeling: Theory and Applications*. Princeton University Press, 2009.

F. Luo and S. Mehrotra. Distributionally robust optimization with decision-dependent ambiguity set, 2020.

V. Moriggia, M. Kopa, and S. Vitali. Pension fund management with hedging derivatives, stochastic dominance and nodal contamination. *Omega - The International Journal of Management Science*, 87:127–141, 2019.

N. Noyan, G. Rudolf, and M. Lejeune. Distributionally robust optimization with decision-dependent ambiguity set. 2018. URL `http://www.optimization-online.org/DB_FILE/2018/09/6821.pdf`.

K. Nyholm and R. Vidova-Koleva. Nelson–siegel, affine and quadratic yield curve specifications: which one is better at forecasting? *Journal of Forecasting*, 31 (6):540–564, 2012.

B. Øksendal. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, sixth edition edition, 2007.

G. Pflug and A. Pichler. *Multistage Stochastic Optimization*. 02 2016. ISBN 978-3-319-08842-6. doi: 10.1007/978-3-319-08843-3.

G. C. Pflug. On-line optimization of simulated markovian processes. *Mathematics of Operations Research*, 15(3):381–395, 1990.

S. R. Pliska and J. Ye. Optimal life insurance purchase and consumption/investment under uncertain lifetime. *Journal of Banking & Finance*, 31 (5):1307–1319, 2007.

K. Postek, D. den Hertog, and B. Melenberg. Computationally tractable counterparts of distributionally robust constraints on risk measures. *SIAM Review*, 58(4):603–650, 2016.

S. Robinson. Local Structure of Feasible Sets in Nonlinear Programming, Part III: Stability and Sensitivity. *Mathematical Programming Study*, 30:45–66, 1987.

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

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

W. Römisch. Stability of stochastic programming problems. In A. Ruszczynski and A. Shapiro, editors, *Stochastic Programming, Handbooks in Operations Research and Management Science*. Elsevier, Amsterdam, 2003.

J. O. Royset and R. J.-B. Wets. Variational theory for optimization under stochastic ambiguity. *SIAM Journal on Optimization*, 27(2):1118–1149, 2017.

V. Russo and F. J. Fabozzi. Pricing coupon bond options and swaptions under the one-factor hull–white model. *The Journal of Fixed Income*, 25(4):76–82, 2016.

V. Russo and G. Torri. Calibration of one-factor and two-factor hull–white models using swaptions. *Computational Management Science*, 16(1):275–295, 2019.

T. Rusý and M. Kopa. An asset – liability management stochastic program of a leasing company. *Kybernetika*, 54(6):1247 – 1263, 2018.

A. Ruszczyński and A. Shapiro. *Stochastic Programming, Handbook in Operations Research and Management Science*. 01 2003.

A. Shapiro, D. Dentcheva, and A. Ruszczynski. *Lectures on Stochastic Programming: Modeling and Theory, Second Edition*. Society for Industrial and Applied Mathematics, 2014. ISBN 978-1-61197-342-6.

B. Tarhan, I. E. Grossmann, and V. Goel. Stochastic programming approach for the planning of offshore oil or gas field infrastructure under decision-dependent uncertainty. *Industrial and Engineering Chemistry Research*, 48(6):3078–3097, 2009.

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

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

J. Vedani, N. El Karoui, S. Loisel, and J.-L. Prigent. Market inconsistencies of market-consistent european life insurance economic valuations: pitfalls and practical solutions. *European Actuarial Journal*, 7(1):1–28, 2017.

K. Vishwanath, S. Peeta, and S. Salman. Investing in the links of a stochastic network to minimize expected shortest path length. *Purdue University Economics Working Papers*, Paper no: 1167, 2004.

S. Vitali, V. Moriggia, and M. Kopa. Optimal pension fund composition for an italian private pension plan sponsor. *Computational Management Science*, 14 (1):135–160, 2017.

M. Vojtek. Calibration of interest rate models-transition market case. CERGE-EI Working Papers wp237, The Center for Economic Research and Graduate Education – Economics Institute, Prague, 2004.

J. Žáčková. On minimax solutions of stochastic linear programming problems. *Casopis pro pestovani matematiky*, 91(4):423–430, 1966.

M. R. Wagner. Stochastic 0–1 linear programming under limited distributional information. *Operations Research Letters*, 36(2):150–156, 2008.

J. Witzany. Valuation of volatility sensitive interest rate derivatives in an emerging market. *International Journal of Financial Markets and Derivatives*, 1(4): 438–451, 2010.

X. Yu and S. Shen. Multistage distributionally robust mixed-integer programming with decision-dependent moment-based ambiguity sets. *Mathematical Programming*, 2020. URL https://doi.org/10.1007/s10107-020-01580-4.

F. Zapletal, M. Šmíd, and M. Kopa. Multi-stage emissions management of a steel company. *Annals of Operations Research*, 292:735–751, 2020.

J. Zhang, H. Xu, and L. Zhang. Quantitative stability analysis for distributionally robust optimization with moment constraints. *SIAM J. Optim.*, 26:1855–1882, 2016.

Y. Zhang, R. Jiang, and S. Shen. Ambiguous chance-constrained binary programs under mean-covariance information. *SIAM Journal on Optimization*, 28(4): 2922–2944, 2018.

S. Zymler, D. Kuhn, and B. Rusten. Distributionally robust joint chance constraints with second-order moment information. *Mathematical Programming*, 137(1-2):167–198, 2013.

List of Figures

List of Tables

List of Publications

Impact Factor Journals

1. Tomáš Rusý, Miloš Kopa. An asset – liability management stochastic program of a leasing company, *Kybernetika*, 54(6): 1247-1263, 2018.

2. Miloš Kopa, Tomáš Rusý. A decision-dependent randomness stochastic program for asset-liability management model with a pricing decision, *Annals of Operations Research*, 299(1-2), pages 241-271, 2021

3. Miloš Kopa, Tomáš Rusý. Robustness of stochastic programs with endogenous randomness via contamination, *European Journal of Operations Research*, 2021. In review.

4. Kamil Kladívko, Tomáš Rusý. Maximum likelihood estimation of the Hull-White model, *Journal of Empirical Finance*, 2021. In review.

Others

5. Tomáš Rusý. Optimal loan performance management via stochastic programming. In L. Váchová and V. Kratochvíl, editors, *Proceedings from 36th International Conference on Mathematical Methods in Economics (MME)*: 476-481, 2018.

6. Tomáš Rusý, Kamil Kladívko. Interest rate modelling: maximum likelihood estimation of one-factor short-rate models. In M. Houda and R. Remeš, editors, *Proceedings from 37th International Conference on Mathematical Methods in Economics (MME)*: 374-379, 2019.