



Matti Koivu

A STOCHASTIC OPTIMIZATION APPROACH TO FINANCIAL DECISION MAKING

HELSINKI SCHOOL OF ECONOMICS

ACTA UNIVERSITATIS OECONOMICAE HELSINGIENSIS

A-234

Matti Koivu

A STOCHASTIC OPTIMIZATION APPROACH TO
FINANCIAL DECISION MAKING

HELSINKI SCHOOL OF ECONOMICS

ACTA UNIVERSITATIS OECONOMICAE HELSINGIENSIS

A-234

© Matti Koivu and
Helsinki School of Economics

ISSN 1237-556X
ISBN 951-791-841-0
ISBN 951-791-842-9 (Electronic dissertation)

Helsinki School of Economics -
HeSE print 2004

Title: A Stochastic optimization approach to financial decision making

Author: Matti Koivu
Helsinki School of Economics
Department of Economics and Management Science
P.O.Box 1210, 00101 Helsinki, Finland
matti.koivu@hkkk.fi

Date: April 2004

Abstract: Stochastic optimization is an effective tool for analyzing decision problems under uncertainty. In stochastic optimization a decision problem is formulated as an optimization problem, where the objective is to find an optimal decision, while considering all the possible scenarios for the uncertain factors and dependencies between the decision variables through time. In stochastic optimization the decision problem is solved numerically and there are only minor limitations for decision criteria, constraints and distributions of random factors that can be used in the formulations. This thesis consists of an introductory section and four articles. The introduction summarizes the contents and findings of the four articles and provides an introduction to the main issues in stochastic optimization: formulation of the decision problem as a stochastic program, econometric modeling of the stochastic factors and discretization of the problem for numerical solution. The first two articles are related to Asset-Liability Management (ALM) problem of a Finnish pension company. Article 1 develops a stochastic model for assets and liabilities of a pension company. The model is used for producing long term forecasts for asset returns as well as company's liabilities and cash-flows. The model is utilized in Article 2, where a stochastic optimization model for ALM of a Finnish pension company is developed. The model is used as a decision support tool for finding long-term dynamic investment decisions in an uncertain environment, where the aim is to cover the uncertain future liabilities with dynamic investment strategies. The last two Articles address the problem of discretization of stochastic programs for numerical solution. New scenario generation techniques based on deterministic and randomized integration quadratures, more precisely Quasi Monte Carlo methods, are developed and applied to financial portfolio optimization problems. Conditions that guarantee the convergence of the objectives and solutions of the discretized problems to the original one are derived for both, Quasi-Monte Carlo and Randomized Quasi-Monte Carlo methods.

Keywords: Stochastic optimization, Asset-Liability Management, Econometric modeling, Finance, Quasi-Monte Carlo methods, Discretization, Variance reduction techniques

Doctoral dissertation

Helsinki School of Economics
Department of Economics and Management Science

A Stochastic optimization approach to financial decision making

Author: Matti Koivu
Supervisor: Professor Markku Kallio

**Preliminary
Examiners:** Professor William T. Ziemba, University of British Columbia, Canada
Professor Hercules Vladimirov, University of Cyprus, Cyprus

Opponent: Professor William T. Ziemba, University of British Columbia, Canada

Type of Research: Based on essays

Essays: Matti Koivu, Teemu Pennanen and Antero Ranne, “Modeling assets and liabilities of a Finnish pension insurance company: a VEqC approach”

Petri Hilli, Matti Koivu, Teemu Pennanen and Antero Ranne, “A stochastic programming model for asset liability management of a Finnish pension insurance company”

Teemu Pennanen and Matti Koivu, “Epi-convergent discretizations of stochastic programs via integration quadratures”

Matti Koivu, “Variance reduction in sample approximations of stochastic programs”

Preface

This work has been carried out while I have been working with the Department of Economics and Management Science at the Helsinki School of Economics. I am grateful to my supervisor Professor Markku Kallio who initially encouraged me to start my doctoral studies and has relentlessly guided me through the entire process. I have really enjoyed working under his supervision and his scientific expertise has greatly improved this thesis. I am also very grateful to Teemu Pennanen who together with Markku started the, now successful, stochastic optimization research project here at the Department of Management Science during fall 2001. This work is the first doctoral thesis of the project. It has not only been inspiring to work with Teemu as a co-author, but I have greatly benefited from his sharp and detailed comments, friendly advice and productive discussions with him during the different stages of this project. I owe special thanks to my co-authors Petri Hilli and Antero Ranne for their significant contribution to this thesis. The constructive comments of the pre-examiners of this thesis, Professor Hercules Vladimirov and Professor William T. Ziemba, are greatly appreciated.

I wish to thank all my colleagues at the department for a friendly and supporting atmosphere. Especially, I would like to express my gratitude to Professor Merja Halme and Leena Tanner for their co-operation and discussions over the years. I would also like to thank Aleksi Mattila and Vesa Poikonen together with the whole Management Science personnel for making my doctoral studies here an enjoyable experience.

Financial support from the Helsinki School of Economics Center for Doctoral Program, the Foundation of the Helsinki School of Economics and the School of Statistical Information, Inference, and Data Analysis is gratefully acknowledged.

I wish to thank my family and friends for their support during my studies. I am indebted to my parents for their endless support and encouragement. My brother deserves special thanks for all the help he has given me. Finally, I wish to express my gratitude to my wife, Kati for all the joy she has brought into my life. Because of you and our lively children Petja and Inka, the completion of this thesis took longer than expected, but it has been worth it.

Helsinki, April 2004

Matti Koivu

1 Introduction

Most decision making problems are influenced by factors that are uncertain at the time of the decision. This is especially true for dynamic decision problems, where the uncertainty is related to future realizations of certain key variables. Typical examples include long term strategic investment decisions in banks, pension insurance companies and foundations, where the uncertain investment returns play a crucial role in portfolio allocation decisions. There are at least four related approaches for solving discrete-time dynamic decision problems under uncertainty, namely stochastic optimization, stochastic dynamic programming, stochastic optimal control and decision rules, and each of the approaches is effective in certain domains.

Decision rules are, possibly state dependent, functions for calculating the values of decision variables at each time period. The decision function may depend on certain key economic variables, which are known to influence the (optimal) decisions. Decision rules are quite easy to implement and they are intuitive for most managers. The most widely used decision rules in portfolio allocation applications are fixed-mix strategies, where the investment portfolio is always rebalanced to fixed proportions; see (Fleten et al. 2002). There have been attempts to optimize decision rules, but they usually lead to non-convex optimization problems, which are very hard to solve (Høyland 1998, Fleten et al. 2002, Kouwenberg 2001). The solutions obtained with decision rules are usually sub-optimal for two reasons: one, global optimality in non-convex optimization problems is difficult to assure and two, the decision rules usually cover only a subset of possible solutions to a decision problem.

In stochastic dynamic programming the problem is to find optimal decisions at discrete points in time. The actions taken in some state at a given time stage are influenced by realizations of random variables. The typical approach to solving dynamic programming problems is to form a backward recursion resulting in an optimal decision associated with each state at each time stage; see e.g. (Bertsekas 2000). The need to enumerate all decisions as well as outcomes of random variables limits this approach to decision problems in which the state space can be kept manageable, i.e. at most 3 or 4 driving variables. Dynamic programming based control remains tied to decision problems satisfying the Markovian property, where the decisions and outcomes depend only on the current state of the world, and not on the history of past states preceding the current one.

Discrete-time stochastic optimal control, (SC) is a general modeling framework, where the problem is to find optimal state dependent control rules for the SC problem. The method also applies to problems, where the decisions and states are Markovian and the state space is low dimensional, i.e. at most 3–4 driving variables. Moreover, it is difficult to incorporate complex constraints often inherent in practical applications to the SC models. These shortcomings limit the applicability of stochastic dynamic programming and stochastic control models considerably in large real life optimization applications under uncertainty.

Stochastic optimization or stochastic programming (SP), provides the most general modeling framework for decision problems under uncertainty. SP formulations are not tied to Markovian assumptions and the method can easily handle high dimensional state spaces. Real life features, such as transaction costs, risk aversion, taxes, asset allocation bounds, legal restrictions and other complex considerations, are easily addressed with stochastic optimization. In SP a decision problem is formulated as an optimization problem, where the objective is to find an optimal decision for the problem, while considering all the possible scenarios for the uncertain factors and dependencies between the decision variables through time. The generality of the approach is based on

numerical computations and the use of modern optimization techniques, and there are only minor limitations for decision criteria, constraints and distributions of random factors that can be used in the formulations. SP formulations can lead to very large scale optimization problems, which require efficient solution algorithms due to possibly enormous number of decision variables in real life multiperiod problems. For a general introduction to SP the reader is referred to the official stochastic programming web-site: www.stoprog.org.

During the last decade the number of stochastic optimization applications to complex real life decision problems has grown dramatically, mainly due to fast development of computers' calculation capacity and modern optimization techniques, and the experience has demonstrated that the quality of the decisions in complex situations can be improved with stochastic programming (Ziemba and Mulvey 1998). Stochastic optimization models have been successfully adapted to applications in energy (Wang and Fang 2002), financial planning (Mulvey and Vladimirou 1992, Cariño and Ziemba 1998, Cariño et al. 1998, Ziemba and Mulvey 1998, Høyland 1998, Kouwenberg 2001), telecommunications (Sen et al. 1994) and supply chain management (Santoso et al. 2003), to mention a few.

Consider a multiperiod decision making problem, where a decision maker has to choose a vector x_t (e.g. portfolio allocation) for every time period $t = 0, \dots, T$. The decisions are affected by the realizations of a discrete time stochastic process $\{\xi_t\}_{t=0}^T$ (e.g. investment returns), which are not completely known at the time of the decision. At time t the decision maker knows the realizations $\{\xi_0, \dots, \xi_t\}$ and the probability measure P for the stochastic process $\{\xi_t\}_{t=0}^T$. The decision x_t at time t depends on the realizations $\{\xi_0, \dots, \xi_t\}$, and on the expectations of the future values $\{\xi_{t+1}, \dots, \xi_T\}$ of the stochastic process.

Many decision problems can be written as a following stochastic optimization problem.

$$\text{maximize} \quad E^P \sum_{t=0}^T u_t[\xi_0, \dots, \xi_t, x_t(\xi_0, \dots, \xi_t)] \quad (1)$$

subject to

$$\begin{aligned} A_t(\xi_0, \dots, \xi_t)x_t(\xi_0, \dots, \xi_t) + B_t(\xi_0, \dots, \xi_t)x_{t-1}(\xi_0, \dots, \xi_{t-1}) \\ = b_t(\xi_0, \dots, \xi_t) \quad P\text{-a.s.} \quad \forall t = 0, \dots, T. \end{aligned} \quad (2)$$

Where E^P denotes expectation under the probability measure P , $u_t[\xi_1, \dots, \xi_t, x_t]$ measures the utility from choosing x_t at time t in state $\{\xi_0, \dots, \xi_t\}$. Note that functions u_t can take a value $-\infty$, which makes it possible to include many types of constraints for x_t in the formulation. Equation (2), where A_t and B_t are linear operators dependent on the realizations $\{\xi_0, \dots, \xi_t\}$ and b_t is a vector, describes dependencies between consecutive time periods (e.g. development of wealth or budget from one period to another). Decision variables $\{x_t\}_{t=0}^T$ are functions, which give as a solution to the problem, the optimal decision for every time period t in state $(\xi_0, \dots, \xi_T)$.

The solution of the decision problem by stochastic optimization can be divided into four phases:

1. Model formulation, which requires the selection of essential decision variables, constraints and stochastic components for the problem, and choosing an appropriate objective function;
2. Description of stochastic components, which requires the definition of an appropriate stochastic process and estimation of the model parameters;
3. Discretization of the problem for numerical solution;

4. Solution of the discretized problem with appropriate optimization algorithms.

This introduction covers phases 1–3 by providing an overview of the general modeling procedure for stochastic optimization employed in this thesis. Phases 1–3 are the most important parts in formulation and implementation of stochastic optimization problems. Once a decision problem is formulated as a stochastic programming problem, a stochastic process for the random factors has to be specified. This involves the selection of an appropriate stochastic model and estimation of the model parameters. In practice, the stochastic factors are usually described as continuous random variables. This results in infinite dimensional optimization problems, which are usually impossible to solve. In stochastic optimization the continuous probability distributions have to be discretized for numerical solution of the problem. The discretized distributions are usually described in a form of a scenario tree with finite number of realizations for the random variables. This results in finite dimensional, usually very large optimization problem, which can be solved with numerical optimization techniques.

The rest of this introduction is organized as follows. Section 2 discusses the modeling of stochastic components, in general and summarizes Article 1 related to modeling assets and liabilities of a Finnish pension insurance company. An Asset Liability Management (ALM) application for a Finnish pension insurance company, developed in Article 2, is described in Section 3. Section 4 provides an overview of the discretization (scenario generation) techniques suggested in the literature and summarizes the findings of Articles 3–4, where deterministic and randomized integration quadratures are applied to scenario generation for stochastic optimization problems.

2 Modeling stochastic factors

Many discrete-time stochastic processes $\{\xi_t\}_{t=1}^T$ can be written as

$$\xi_t = G_t(\xi_{t-1}, \dots, \xi_{t-k}, \epsilon_t, \dots, \epsilon_{t-l}),$$

where $\{\epsilon_t\}_{t=1}^T$ are independent random variables. The econometric literature offers a wide variety of discrete-time stochastic processes for describing the evolution of stochastic factors. These processes can be linear or non-linear and the choice for an appropriate stochastic model depends heavily on the considered application and its time horizon. In financial applications we are usually interested in describing the return distributions of certain asset classes e.g. stocks, bonds, through time. Linear time series models can be written as

$$\xi_t = d + \sum_{i=1}^k A_i \xi_{t-i} + \sum_{j=1}^l B_j \epsilon_{t-j} + \epsilon_t, \quad \epsilon_t \sim N(0, \Sigma), \quad (3)$$

where $\xi_t, \epsilon_t \in \mathbb{R}^n$ denote the vector of n dependent variables e.g. asset returns and error terms, respectively, $d \in \mathbb{R}^n$ is the vector of drift terms, $A_i, B_j \in \mathbb{R}^{n \times n}$, $i = 1, \dots, k, j = 1, \dots, l$, are estimated coefficient matrices for lagged dependent variables and error terms, respectively and $\Sigma \in \mathbb{R}^{n \times n}$ is the covariance matrix of error terms. Usually, in empirical work $\{\epsilon_t\}_{t=1}^T$ are assumed to be independent normally distributed random variables. The discrete time version of the most widely used stochastic model for asset prices in financial literature, (multivariate) geometric brownian motion, is obtained from (3) by setting $l = 0, k = 1, A_1 = I$ (identity matrix), and can be written as

$$\xi_t - \xi_{t-1} = \Delta \xi_t = d + \epsilon_t, \quad \epsilon_t \sim N(0, \Sigma),$$

where ξ denotes logarithmic asset prices. Many popular time series models used in modeling long term asset returns are obtained by a suitable parametrization of (3). Equation (3) presents a Vector Autoregressive Moving Average (VARMA) model, see e.g. Lütkepohl (1991) and by setting $l = 0$ we obtain a Vector Autoregression (VAR) model (Sims 1980), which has been widely used in long term analysis of asset returns (Campbell et al. 1997, Campbell and Shiller 1987, 1988). Vector Equilibrium Correction (VEqC) model (Engle and Granger 1987, Johansen 1995) is also a special case of (3), since any VEqC model can always be written as a VAR model.

Although many of the models commonly used in empirical finance are linear, the nature of short term financial data suggests that non-linear models are more appropriate for forecasting and describing asset returns and volatility.

It is well known, that short term (daily - monthly) asset returns are fat tailed, skewed and heteroskedastic, see e.g. Campbell et al. (1997). Therefore, short term evolution of asset returns are usually described with non-linear models. The following provides a short description of widely utilized non-linear models in financial econometrics literature. Autoregressive Conditional Heteroskedasticity (ARCH) (Engle 1982) and Generalized ARCH (GARCH) (Bollerslev 1986) models are designed to capture the serial correlation in variance of observed asset returns. The original formulations considered univariate ($n = 1$) ARCH(q) and GARCH(p, q) models, which are written as

$$\begin{aligned} \text{ARCH:} \quad \sigma_t^2 &= \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2, \\ \text{GARCH:} \quad \sigma_t^2 &= \omega + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2, \end{aligned}$$

where σ_t^2 is the variance of ξ_t at time t , ω, α_i and β_j are non-negative parameters and p and q are the number of lagged variances and squared error terms, respectively. The presented ARCH and GARCH models have multivariate counterparts, see e.g. (Engle and Kroner 1995, Engle 2002), which can be used to model time dependent variances and covariances of asset returns e.g. in (3) by allowing the covariance matrix to be time dependent (Σ_t).

Fat tails and asymmetries in financial data can be taken into account by replacing the (uni-)multivariate normal distribution in (3) with another, more fat tailed (t -distribution) or skewed probability distribution, like skew-normal or skew- t -distribution (Azzalini and Capitanio 1999, 2003). Another way of modeling asymmetric time series are so called regime-switching models (Franses and Dijk 2000). The basic idea of regime-switching models is that the stochastic process is time-invariant conditional on the regime prevailing at time t . Regime-switching models characterize a non-linear stochastic process as piecewise linear by restricting the model to be linear in each regime. The number of regimes is limited, they may depend on endogenous or exogenous variables and the models differ in their assumptions concerning the stochastic process generating the regime (Franses and Dijk 2000).

Although non-linear models describe short term financial data more accurately than linear models they also have some drawbacks. Parameter estimation and their interpretation in non-linear models is usually much more difficult than in linear models. The problems with fat tails and heteroscedasticity are much less pronounced in long term (quarterly, yearly) asset returns, which favor linear time series models in long term forecasting models.

This short overview already shows that there is a variety of models for describing the evolution of given stochastic factors. The selection of an appropriate model is always strongly dependent on the application and unfortunately, often a non-trivial task.

Article 1 analyzes a problem related to modeling assets and liabilities of a Finnish pension insurance company. Pension companies manage enormous investment funds and their strategic investment allocations and bonus strategies are usually updated on a yearly basis. These decisions are linked with the companies' liabilities which extend far into the future. On one hand, this link comes from the companies' goal of meeting the liabilities, and on the other, from the legislation that aims at regulating the solvency risks of the company. The main goal of the pension insurance companies is to invest policyholders' pension insurance premiums safely and profitably in order to meet their future liabilities. The duration of the liabilities is usually very long (over 20 years), which calls for realistic simulation models for long-term scenarios of investment returns and liability flows. Such models form the basis for pension insurance companies' asset and liability management (ALM). Typically, these models have a macroeconomic flavor in that they try to describe the development of larger investment classes like short and long term bonds and broad equity indices along with wage indices and inflation. Many existing stochastic models for pension companies have been based on the well known Wilkie's stochastic investment model; see e.g. Wilkie (1986, 1995), Yakoubov et al. (1999) and Ranne (1998). The drawback of these models is their cascade structure, which allows the modeling of one-way causalities only. VAR and their generalizations, VEqC models are linear time series models, which do not have this limitation. These models have become widely utilized tools in applied econometrics during the last two decades, see e.g. Campbell and Shiller (1987, 1988) and Campbell et al. (1997).

In Article 1 VAR and especially VEqC models are developed for modeling long term asset returns and liabilities of a Finnish pension insurance company. The considered stochastic factors describe the development of larger investment classes, namely, cash, bonds, equities, property and loans to policyholders as well as liabilities, cash flows and some technical quantities that have an important role in the Finnish statutory earnings-related pension scheme. All the stochastic variables in the model are expressed in terms of seven economic factors, namely short-term interest rate, bond yield, stock price -, dividend -, property price -, rental - and wage index. The separate treatment of prices and dividends enables the modeling of such terms as dividend yield (dividend-price ratio) that has been shown to have predictive power in describing future changes of dividends (Campbell and Shiller 1988).

Article 1 emphasizes a general model building procedure that combines statistical information with user-specified characteristics. The developed model incorporates statistical information with expert views and it has been successfully implemented in a case where the available data is scarce and subject to changing economic conditions. The expert views are given in the form of drift parameters (for such quantities as interest rates, equity and property prices), and certain long-run equilibrium relations (e.g. average levels of interest rate spread and dividend yield). This is especially important in situations where the available data displays characteristics that are believed to change in the future (e.g. declining interest rates in EU-area during the 90's). Indeed, according to Hendry and Doornik (1997), deterministic factors like drift parameters and equilibrium values matter most for predictive failure of economic time series. The model is evaluated against rival VAR and VEqC models in out-of-sample forecasting tests and long-term simulation experiments. The results support the importance of incorporating expert information and co-integration relations in long-term forecasting models.

3 A stochastic programming model for asset liability management of a Finnish pension insurance company

Stochastic programming has proven to be an efficient tool in designing good strategies in wealth- and asset liability management in practice. This is due to its ability to cope with the dynamics and complex constraint structures usually inherent in such problems. Successful applications of stochastic programming to asset liability management have been reported in Nielsen and Zenios (1996), Cariño and Ziemba (1998), Cariño et al. (1998), Høyland (1998), Consigli and Dempster (1998), Kouwenberg (2001), and Geyer et al. (2002).

Article 2 describes a stochastic programming model and its computer implementation for asset liability management of a Finnish pension insurance company. Finnish pension insurance companies are responsible for huge investment funds and, like most pension companies in Europe, they are facing major challenges with a large number of retiring employees in the near future. The Finnish statutory earnings-related pension scheme is based on the so called defined benefit rule, where the pension insurance companies guarantee pension payments which are tied to the development of the policyholder's salaries. Because the system is statutory, many of its characteristics are common to all the companies. For example, the contribution rates and the technical reserves are calculated according to common formulas confirmed by the Ministry of Social Affairs and Health. The technical interest rate for the reserves is also common to all the companies and its value for each year is confirmed by the Ministry.

The pension insurance companies are, however, able to choose their own investment policy. Since, the Finnish pension system is statutory the companies are frequently obliged to report their solvency ratios, the ratio between solvency capital and solvency border, to the Ministry of Social Affairs and Health. The solvency capital is approximately the difference between the company's investment capital and the discounted expected value of future liabilities. A fundamental concept in the system is the solvency border, whose value depends on the investment portfolio of the company and is calculated according to the formulas given in a government decree. The starting point of the system is that the probability of ruin in one year at the solvency border, i.e. when solvency capital equals solvency border, should be approximately 2.5%, and therefore the value of the border is required to be dependent on the investment portfolio. In short, the riskier the company's investment portfolio the higher the value of solvency border. Depending on the investment returns and solvency position, the company can give bonuses to its customers. These bonuses are paid as reductions of the contributions, and they are the most important element in the competition between the Finnish pension insurance companies.

To cope with the increasing pension expenditures, caused by the retirement of the large age groups, the pension insurance companies need to manage their massive wealth reserves efficiently and finding good investment strategies is essential for the success of the companies. Finding such strategies is however a very complicated task, since companies must simultaneously take into account the nature of its future liabilities as well as its solvency position and bonus strategies.

Article 2 develops a decision support tool, stochastic optimization model, to address these issues in long-term asset-liability management framework. The developed model describes a long-term dynamic investment problem where the aim is to cover the uncertain future liabilities with dynamic investment strategies in an uncertain environment. The assets are considered at the level of the larger investment classes of cash, bonds, stocks, property and loans to policyholders. In addition to investment decisions, the model looks for good bonus payment strategies and it takes explicitly

into account various portfolio and transaction restrictions as well as legal restrictions coming from the complex Finnish pension system. The legal restrictions form a unique part of the model not present in earlier applications of stochastic programming.

Particular attention is given to the description of uncertain factors in the model which include investment returns, cash-flows, and the so called technical reserves used in the definition of the statutory restrictions. This is important since the output of a stochastic optimization model depends usually heavily on the underlying stochastic model. The evolution of the stochastic factors is described with a VEqC model developed in Article 1. The model is implemented and tested against static fixed-mix and dynamic portfolio insurance strategies (Perold and Sharpe 1988, Black and Jones 1988). Fixed-mix strategies are simple decision rules that always rebalance the investment portfolio to fixed proportions. In dynamic portfolio insurance strategies the investment proportions depend on the company's solvency so that more wealth is allocated to risky assets, stocks, when the company's solvency level is high and the stock market exposure is reduced as the company approaches insolvency.

These decision strategies cannot be considered as fully realistic models for the behavior of a real pension insurance company. However, they are currently used for various simulation purposes in practice, which motivates their use as benchmarks. In out-of-sample tests, the strategies based on the developed stochastic optimization model clearly outperform both the fixed-mix and portfolio insurance strategies. Similar results have been obtained for the more sophisticated but computation-intensive benchmarks in Høyland (1998), Kouwenberg (2001) and Fleten et al. (2002).

4 Scenario generation for stochastic programs

In practice, the problem dependent stochastic factors are usually described as continuous random variables, like in Section 2. This results in infinite dimensional optimization problems, which are generally impossible to solve. In stochastic optimization the random factors are usually described in a form of a scenario tree with finite number of realizations for the random variables. The resulting finite dimensional optimization problem can be solved with numerical optimization techniques; see e.g. King (1996). The properties of the discrete approximation P^ν , where ν is the number of scenarios, of the true probability measure P have a strong impact on the optimal solution obtained with the discretization. In general, the aim is to generate approximations P^ν of P , so that the discretized problem approximates the true optimization problem as well as possible.

The simplest and perhaps the most widely studied and used scenario generation method in stochastic programming applications is Monte Carlo (random) sampling, see Cariño and Ziemba (1998), Cariño et al. (1998), Kouwenberg (2001), Shapiro (2000). In Monte Carlo (MC) sampling the discrete probability measures P^ν are generated as a random sample from P . It is clear that a random sample can lead to a bad approximation of P , which in turn, may lead to an equally bad approximation of the optimization problem. The optimal values and solutions obtained with a specific discretization are often highly dependent on the sample. There have been attempts to improve the accuracy of crude Monte Carlo sampling by using ideas from importance sampling technique; see Infanger (1992) and Dempster and Thompson (1999). An idea originally proposed by Cariño et al. (1998) and refined by Høyland and Wallace (2001) is to use moment matching, where P^ν is constructed so that it has the first few moments of the original distribution; see also Høyland et al. (2003). In barycentric approximation, one constructs P^ν so that, under some convexity

properties of the objective function, the optimum value of the discretized problem provides an upper/lower bound for the true optimum value; see Frauendorfer (1992). Pflug (2001) proposed to construct discrete measures P'' so that they are as close as possible to P in the sense of the so called Wasserstein-distance.

Given that, the expectation in (1) usually involves the computation of multidimensional integrals it is surprising that modern numerical integration methods have not been extensively applied to scenario generation. In Articles 3 and 4 modern integration quadratures, more precisely Quasi Monte Carlo (QMC) methods, see e.g. Niederreiter (1992), are applied to discretization of one-stage stochastic programs. QMC methods can be seen as a deterministic counterpart to MC. They are designed to give weakly convergent discrete measures that approximate a given probability measure as well as possible. Moreover, they are just as easy to use as crude Monte Carlo and they are very fast compared to methods like barycentric approximation, moment matching or that in Pflug (2001).

Article 3 derives conditions that guarantee the epi-convergence of the objectives of the discretized problems to the original one. Epi-convergence (see e.g. Rockafellar and Wets (1998)) of the objectives is a minimal property that should be satisfied by any approximation scheme for optimization problems in order to get asymptotic convergence of the optimal values and solutions.

Epi-convergence of stochastic programs with respect to perturbations in the probability measure has been studied, for example, by Birge and Wets (1986), Robinson and Wets (1987), Dupačová and Wets (1988), Kall et al. (1988), Lucchetti and Wets (1993), Artstein and Wets (1994), Zervos (1999), Schultz (2000) and Vogel and Lachout (2003). In these studies, weak convergence of the approximating measures P'' to the original measure P has been found an important property. The epi-convergence result of Article 3 is closely related to the ones in the above references but it is easier to apply to discretizations and it does not require the feasible set to be independent of the probability measure. In Article 3, epi-convergence of three different models of portfolio management is proved and their behavior is studied numerically. Besides MC, the developed discretizations are the only existing ones with guaranteed epi-convergence for these problem classes. In the test problems, integration quadratures seem to result in faster convergence of optimal values than MC.

A disadvantage of deterministic QMC methods is that, the computation of accurate error estimates, which are usually deemed important in real life applications, is very difficult. L'Ecuyer and Lemieux (2002) review several QMC constructions and their randomizations that have been proposed to provide unbiased estimators and for error estimation. Randomized Quasi Monte Carlo (RQMC) methods can be regarded as variance reduction techniques with respect to MC, as well as some better known techniques, like importance sampling and antithetic variates; see e.g. (Bratley et al. 1987). RQMC methods can be used just like MC in estimating confidence intervals and variances for sample approximations in numerical integration, where RQMC often result in significant variance reduction with respect to MC; see (Lemieux and L'Ecuyer 2000).

Article 4 studies the use of RQMC methods and other variance reduction techniques in sample approximations of stochastic programs. It is shown in Article 4, that the epi-convergence result derived in Article 3 for QMC, also applies to discretizations generated with RQMC methods. Variance reduction techniques, like antithetic variates, importance - and latin hypercube sampling have been used in stochastic optimization e.g. in Kouwenberg (2001), Infanger (1992), Hight (1998) and Linderroth et al. (2002). The results of these studies have shown that variance reduction techniques improve the accuracy of the sample approximations over MC. However, RQMC methods have not been previously applied to stochastic optimization. These methods can be viewed as an

alternative to MC in computing so called statistical bounds, as in Shapiro (2003). In five different numerical test problems of portfolio management, the lower bounds for the optimal values obtained with RQMC are consistently much tighter than those obtained with MC. Moreover, RQMC methods significantly reduces the sample variance of the optimal values compared to MC and antithetic variates, thus providing an efficient sampling method for stochastic optimization.

References

- Z. Artstein and R. J.-B. Wets. Stability results for stochastic programs and sensors, allowing for discontinuous objective functions. *SIAM J. Optim.*, 4(3):537–550, 1994.
- A. Azzalini and A. Capitanio. Statistical applications of the multivariate skew normal distribution. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 61(3):579–602, 1999.
- A. Azzalini and A. Capitanio. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t -distribution. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 65(2):367–389, 2003.
- D. P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 1. Academic Press, 2nd edition, 2000.
- J. R. Birge and R. J.-B. Wets. Designing approximation schemes for stochastic optimization problems, in particular for stochastic programs with recourse. *Math. Programming Stud.*, (27):54–102, 1986.
- F. Black and R. Jones. Simplifying portfolio insurance for corporate pension plans. *Journal of Portfolio Management*, 14:33–37, 1988.
- T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *J. Econometrics*, 31(3):307–327, 1986.
- P. Bratley, B. L. Fox, and L. E. Schrage. *A guide to simulation*. Springer-Verlag, 2 edition, 1987.
- J. Y. Campbell, A. W. Lo, and A. C. MacKinley. *The Econometrics of Financial Markets*. Princeton University Press, USA, 1997.
- J. Y. Campbell and R. Shiller. Cointegration and tests of present value models. *Journal of Political Economy*, 95(5):1062–1088, 1987.
- J. Y. Campbell and R. Shiller. Stock prices, earnings and expected dividends. *Journal of Finance*, 43(3):661–676, 1988.
- D. R. Cariño, D. H. Myers, and W. T. Ziemba. Concepts, technical issues, and uses of the Russell-Yasuda Kasai financial planning model. *Operations Research*, 46(4):450–462, 1998.
- D. R. Cariño and W. T. Ziemba. Formulation of the Russell-Yasuda Kasai financial planning model. *Operations Research*, 46(4):433–449, 1998.
- G. Consigli and M. A. H. Dempster. Dynamic stochastic programming for asset-liability management. *Ann. Oper. Res.*, 81:131–161, 1998.

- M. A. H. Dempster and R. T. Thompson. EVPI-based importance sampling solution procedures for multistage stochastic linear programmes on parallel MIMD architectures. *Ann. Oper. Res.*, 90:161–184, 1999.
- J. Dupačová and R. J.-B. Wets. Asymptotic behavior of statistical estimators and of optimal solutions of stochastic optimization problems. *Ann. Statist.*, 16(4):1517–1549, 1988.
- R. F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4):987–1007, 1982.
- R. F. Engle. Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *J. Bus. Econom. Statist.*, 20(3):339–350, 2002.
- R. F. Engle and C. W. J. Granger. Co-integration and error correction: representation, estimation, and testing. *Econometrica*, 55(2):251–276, 1987.
- R. F. Engle and K. F. Kroner. Multivariate simultaneous generalized arch. *Econometric Theory*, 11(1):122–150, 1995.
- S.E. Fleten, K. Høyland, and S. W. Wallace. The performance of stochastic dynamic and fixed mix portfolio models. *European J. Oper. Res.*, 140(1):37–49, 2002.
- P. H. Franses and D. van Dijk. *Non-Linear Time Series Models in Empirical Finance*. Cambridge University Press, 2000.
- K. Frauendorfer. *Stochastic two-stage programming*, volume 392 of *Lecture Notes in Economics and Mathematical Systems*. Springer-Verlag, Berlin, 1992.
- A. Geyer, W. Herold, K. Kontriner, and W. T. Ziemba. The innovest austrian pension fund financial planning model InnoALM. Working paper, University of British Columbia, 2002.
- D. F. Hendry and J. A. Doornik. The implications for economic modelling of forecast failure. *Scottish Journal of Political Economy*, 44:437–461, 1997.
- J. L. Higle. Variance reduction and objective function evaluation in stochastic linear programs. *Journal on Computing*, 10:236–247, 1998.
- K. Høyland. Asset liability management for a life insurance company: A stochastic programming approach. Ph.D. Thesis, Norwegian University of Science and Technology, 1998.
- K. Høyland, M. Kaut, and S. Wallace. A heuristic for moment-matching scenario generation. *Computational Optimization and Applications*, 24:169–186, 2003.
- K. Høyland and S. W. Wallace. Generating scenario trees for multistage decision problems. *Management Science*, 47(2):295–307, 2001.
- G. Infanger. Monte Carlo (importance) sampling within a Benders decomposition algorithm for stochastic linear programs. *Ann. Oper. Res.*, 39:69–95, 1992.
- S. Johansen. *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford University Press, New York, 1995.

- P. Kall, A. Ruszczyński, and K. Frauendorfer. Approximation techniques in stochastic programming. In *Numerical techniques for stochastic optimization*, volume 10 of *Springer Ser. Comput. Math.*, pages 33–64. Springer, Berlin, 1988.
- A. J. King. Approximation and computation in stochastic programming. *Mathematical Programming B*, 75(2), 1996.
- R. Kouwenberg. Scenario generation and stochastic programming models for asset liability management. *European J. Oper. Res.*, 134(2):279–292, 2001.
- P. L’Ecuyer and C. Lemieux. Recent advances in randomized quasi-monte carlo methods. In M. Dror, P. L’Ecuyer, and F. Szidarovszki, editors, *Modeling Uncertainty: An Examination of Stochastic Theory, Methods, and Applications*, pages 419–474. Kluwer Academic Publishers, 2002.
- C. Lemieux and P. L’Ecuyer. A comparison of Monte Carlo, lattice rules and other low-discrepancy point sets. In *Monte Carlo and quasi-Monte Carlo methods 1998*, pages 326–340. Springer, Berlin, 2000.
- J. Linderoth, A. Shapiro, and S. Wright. The empirical behavior of sampling methods for stochastic programming. Optimization technical report 02-01, Computer Sciences Department, University of Wisconsin-Madison, 2002.
- R. Lucchetti and R. J.-B. Wets. Convergence of minima of integral functionals, with applications to optimal control and stochastic optimization. *Statist. Decisions*, 11(1):69–84, 1993.
- H. Lütkepohl. *Introduction to multiple time series analysis*. Springer-Verlag, Berlin, 1991.
- J.M. Mulvey and H. Vladimirov. Stochastic network programming for financial planning problems. *Management Science*, 38:1642–1664, 1992.
- H. Niederreiter. *Random number generation and quasi-Monte Carlo methods*, volume 63 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.
- S. S. Nielsen and S. A. Zenios. A stochastic programming model for funding single premium deferred annuities. *Math. Programming*, 75(2, Ser. B):177–200, 1996.
- A. F. Perold and W. F. Sharpe. Dynamic strategies for asset allocation. *Financial Analysts Journal*, 44:16–27, 1988.
- G. Pflug. Scenario tree generation for multiperiod financial optimization by optimal discretization. *Math. Program.*, 89(2, Ser. B):251–271, 2001.
- A. Ranne. The finnish stochastic investment model. In *Transactions of the 26th ICA*, pages 213–238, 1998.
- S. M. Robinson and R. J.-B. Wets. Stability in two-stage stochastic programming. *SIAM J. Control Optim.*, 25(6):1409–1416, 1987.

- R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998.
- T. Santoso, S. Ahmed, M. Goetschalckx, and A. Shapiro. A stochastic programming approach for supply chain network design under uncertainty. *Stochastic Programming E-print Series*, 2003.
- R. Schultz. Some aspects of stability in stochastic programming. *Ann. Oper. Res.*, 100:55–84, 2000.
- S. Sen, R.D. Doverspike, and S. Cosares. Network planning with random demand. *Telecommunication Systems*, 3, 1994.
- A. Shapiro. Stochastic programming by monte carlo simulation methods. *Stochastic Programming E-Print Series*, 2000.
- A. Shapiro. Inference of statistical bounds for multistage stochastic programming problems. *Mathematical Methods of Operations Research*, 58:57–68, 2003.
- C. A. Sims. Macroeconomics and reality. *Econometrica*, 48(1):1–48, 1980.
- S. Vogel and P. Lachout. On continuous convergence and epi-convergence of random functions. II. Sufficient conditions and applications. *Kybernetika (Prague)*, 39(1):99–118, 2003.
- X. Wang and K. T. Fang. The effective dimension and quasi-monte carlo integration. *Journal of Complexity*, 19:101–124, 2002.
- A. D. Wilkie. A stochastic investment model for actuarial use. *Transactions of the Faculty of Actuaries*, 39:341–403, 1986.
- A. D. Wilkie. More on a stochastic asset model for actuarial use. *British Actuarial Journal*, 1: 777–964, 1995.
- Y.H. Yakoubov, M. H. Teeger, and D. B. Duval. Stochastic investment model for asset and liability management. Proceedings of the 9th AFIR International Colloquium, 1999.
- M. Zervos. On the epiconvergence of stochastic optimization problems. *Math. Oper. Res.*, 24(2): 495–508, 1999.
- W. T. Ziemba and J. M. Mulvey, editors. *Worldwide asset and liability management*, volume 10 of *Publications of the Newton Institute*. Cambridge University Press, Cambridge, 1998.

Modeling assets and liabilities of a Finnish pension insurance company: a VEqC approach

Matti Koivu* Teemu Pennanen
Department of Management Science
Helsinki School of Economics
PL 1210
00101 Helsinki, Finland
`[hilli,koivu,pennanen]@hkkk.fi`

Antero Ranne
Actuarial Department
Ilmarinen Mutual Pension Insurance Company
`antero.ranne@ilmarinen.fi`

Abstract

This paper develops a stochastic model for assets and liabilities of a Finnish pension insurance company. The assets and liabilities are expressed in terms of seven economic factors from Finland and the EU-area. The development of these factors is modeled with a Vector Equilibrium Correction model, that incorporates statistical information with expert views in the form of user specified growth rates and long term equilibria. The forecast performance of the resulting model is tested and the model is used in long-term solvency and asset liability simulations.

Key words: time series analysis, VAR, Vector equilibrium correction model, asset liability management.

*This research was supported by The foundation for the Helsinki School of Economics under grant number 9941075.

1 Introduction

Pension companies manage enormous investment funds. Their main goal is to invest policyholders' pension premiums safely and profitably in order to meet their liabilities in the future. The duration of the liabilities is usually very long (over 20 years), which calls for realistic models for long-term scenarios of investment returns and liability flows. Such models form the basis for pension companies' asset and liability management (ALM). Typically, these models have a macroeconomic flavor in that they try to describe the development of larger investment classes like interest rates and broad equity indices along with wage indices and inflation.

To a large extent, the existing models for pension companies have been based on the well known Wilkie's stochastic investment model; see e.g. Wilkie (1986, 1995), Yakoubov et al. (1999) and Ranne (1998). The drawback of these models is their cascade structure, which allows the modeling of one-way causalities only. Vector autoregression (VAR) and their generalizations, vector equilibrium correction (VEqC) models, do not have this limitation; see for example Sims (1980), Engle and Granger (1987), Johansen (1995) and Clements and Hendry (1999). VAR-models for asset liability management have been used for example in Dert (1998), Wright (1998) and Harris (1999), but to our knowledge, only Boender et al. (1998) have proposed using a VEqC model for these purposes.

This paper develops a model for assets and liabilities of a Finnish pension insurance company. The stochastic variables in the model are expressed in terms of seven economic factors whose development is modeled with a linear time series model in VEqC form. The model incorporates statistical information with expert views and it has been successfully implemented in a case where the available data is scarce and subject to changing economic conditions. The expert views are given in the form of drift parameters (for such quantities as interest rates, equity and property prices), and certain long-run equilibrium relations (e.g. average levels of interest rate spread and dividend yield). This is especially important in situations where the available data displays characteristics that are believed to change in the future (e.g. declining interest rates in EU-area during the 90's). Indeed, according to Hendry and Doornik (1997), deterministic factors like drift parameters and equilibrium values matter most for predictive failure.

We will consider assets in the level of larger investment classes, namely, cash, bonds, equities, property and loans to policyholders. We treat the cash-flow and the change-in-value components of total asset returns separately. This is essential in the presence of significant transaction costs. Also, this enables the modeling of such terms as dividend yield (dividend-price ratio) that has been shown to have predictive power in describing future changes of dividends; see Campbell and Shiller (1988). On the liability side, we model the cash-flows of a pension insurance company and some technical quantities that have an important role in the Finnish statutory earnings-related pension scheme. In general, these terms depend on the development of salaries of the pensioners as well as population dynamics. These two factors are assumed independent, so that their development can be modeled separately. In this paper, we concentrate on modeling salaries and their co-movement with asset returns.

All the stochastic parameters in our model will be expressed in terms of the following seven factors: short-term interest rate, bond yield, stock price index, dividend yield, property price index, rental yield and wage index. These (or more precisely, simple transformations of them) will, in turn, be modeled by a structured VEqC model that allows for simple economic interpretations. The data for estimation is taken from Finland and the EU-area that are of greatest interest to Finnish pension

insurance companies.

The rest of this paper is organized as follows. In the next section, we describe the asset classes and liabilities and show how their development can be described with the above seven factors. In Section 3, we describe the data and study its stationarity properties. We start the development of our time series model in Section 4, where we build a VAR model for differences. The purpose there is to show how to specify the average drift in the model and to demonstrate the importance of correct specifications through simulations. In Section 5, we augment the VAR model with equilibrium correction terms thus obtaining our complete model in VEqC form. Section 6 compares the forecast performance of the developed VEqC-model with three rival models in out-of-estimation-sample forecast experiment. In Section 7, we examine by simulation the statistical properties of the long-term asset returns produced by the model and calculate the company's long-term cash-flows and reserves. In Section 8 the developed model is used in an ALM framework. We compare different dynamic portfolio allocation strategies and evaluate the company's long-term solvency and bankruptcy risks. Concluding remarks are presented in Section 9.

2 Assets and liabilities in terms of economic time series

Finnish pension insurance companies update their investment and bonus strategies on a yearly basis. These decisions are linked with the company's liabilities which extend far into the future. On one hand, this link comes from the company's goal of meeting the liabilities, and on the other, from the legislation that aims at regulating the solvency risks of the company. This section describes the main investment classes, the liabilities, and other quantities that are of interest in the strategic financial planning of a Finnish pension insurance company. Our aim is to express all these quantities in terms of seven economic factors that will then be modeled with a time series model.

2.1 Assets

Pension insurance companies' assets can be divided into five main investment classes: cash, long-term bonds, stocks, property and loans. Our goal is to model the return per wealth invested in each asset class. The total returns on the assets are split between cash income and change in value components, which, in general, require separate treatment due to transaction costs etc.

Cash: Pension companies keep a proportion of their assets in cash (short-term deposits) to ensure a reasonable level of liquid financial resources. Because of the short term nature of these investments, the change in value can be ignored and the whole return can be modeled as cash income. The return on cash investments can be well approximated by the 3-month Euribor.

Bonds: Currently, about one half of the Finnish pension insurance companies' wealth is invested in long-term bonds. The primary source of income on bond investments are the coupon payments, which is cash income. Usually, newly issued bonds sell at par, which implies that coupon payments equal the current yield. We approximate the cash flow component of the whole bond portfolio by the bond yield which is denoted by br_t .

According to (Campbell et al., 1997, page 408) the price of a bond can be approximated by

$$\ln B_t \approx D[k + (1 - \rho)c - \ln(1 + br_t)], \quad (1)$$

where D and c are the duration and coupon payments, respectively, of the portfolio, and k and ρ are constants. When the bond is selling at par, $\rho = 1/(1 + br_t) \approx 1$; see (Campbell et al., 1997, page 407). If we assume in addition that the bond portfolio is updated so that its duration is fairly constant (which is not far from reality in Finnish pension insurance companies),

$$\ln B_t - \ln B_{t-1} \approx -D \ln \frac{1 + br_t}{1 + br_{t-1}}.$$

Based on this, we approximate the value change of the bond portfolio by

$$\frac{B_t}{B_{t-1}} \approx \left(\frac{1 + br_{t-1}}{1 + br_t} \right)^D.$$

The duration will be set equal to the duration of the bond portfolio of a company being modeled. We will use the yield on German government benchmark bond, whose duration is close to D . In our calculations we will use $D \approx 5$ years.

Stocks: The riskiest but historically the most profitable long-term investment class is stocks. In stocks, the majority of the total return comes from the change in value and the dividend payments constitute the cash income component. We will model the change in value component with a stock price index and the cash income with the corresponding dividend yield.

Finnish pension insurance companies invest in stocks mainly in Finland and the rest of the European Union (EU) area. The development of the value of the company's stock portfolio is modeled with a "fixed mix" stock price index S which gives the value of a portfolio that is sequentially rebalanced to have a fraction θ^F of stock investments in Finland and θ^E in the EU area. The quarterly change in the fixed mix index is

$$\frac{S_t}{S_{t-1}} = \theta^F \frac{S_t^F}{S_{t-1}^F} + \theta^E \frac{S_t^E}{S_{t-1}^E}$$

where S^F and S^E are stock indices in Finland and EU, respectively. For S^E , we use the Datastream Europe market index, and for S^F the Helsinki Stock Exchange (HEX) portfolio price index.

The annual dividend yield corresponding to the fixed mix stock portfolio is calculated as a weighted average of dividends from Finland and EU as

$$Y_t^S = \theta^F Y_t^F + \theta^E Y_t^E.$$

The Finnish dividend yield Y^F is based on the HEX portfolio price index and the European yield Y^E is based on the Datastream Europe market index. We assume that $\theta^F = \theta^E = \frac{1}{2}$, which has the interpretation that a company's stock portfolio is split evenly between Finland and the rest of the EU.

Property: As an investment class, property resembles stocks in many ways. The return on property investments consists of potentially large price fluctuations and fairly stable cash income. Therefore, the return components on property investments are modeled similarly to stocks: the change in value component is modeled with a property price index and the cash income is modeled with a rental yield. The main difference from stocks is that the cash income component forms the majority of the total gross return on property investments.

Finnish pension insurance companies invest in property mainly domestically. Although the companies mostly invest in commercial property, we will use the Finnish residential property price index to model the price fluctuations of all property investments. This is because the commercial property market is rather illiquid and the property prices are hard to estimate accurately.

The cash income component will be modeled as the difference between the residential rental yield Y_t^P and the maintenance costs. The rental yield gives the rent paid per wealth invested and the maintenance costs are assumed to be a constant 3% of the property value. The property price index P_t (€ / m^2) and the corresponding rental index R_t (€ / m^2) are available from Statistics Finland. The rental yield is given by

$$Y_t^P = \frac{R_t}{P_t}. \quad (2)$$

Loans: The Finnish pension insurance companies invest part of their funds by giving loans to policyholders. In the past, these formed a great majority of all investments, but currently they account for about 10 %. There are two kinds of loans, premium loans and investment loans. The premium loans are an arrangement where a customer can borrow back part of the paid premium according to fixed rules. For the investment loans, the terms are agreed freely between the company and the borrower. In the model, the two kinds of loans are combined to form one investment class. The change in value component for loans is zero. The cash income component will be approximated by a moving average of bond yield. This is based on the fact that the interest on newly given loans is usually set equal to current bond yield.

2.2 Liabilities

The Finnish pension insurance companies are responsible for managing the statutory earnings-related pension scheme. Because the system is statutory, many of its characteristics are common to all the companies. For example, the amount of the pension is determined by fixed defined-benefit rules independent on the company where the person is insured. Also the contribution rates and the technical reserves are calculated according to common formulas confirmed by the Ministry of Social Affairs and Health. The technical interest rate for the reserves is also common to all the companies and its value for each year is confirmed by the Ministry.

Pension insurance companies are, however, able to choose their own investment policies. Depending on the investment returns, companies can give bonuses to their customers. These bonuses are paid as reductions of the contributions, and they are the most important element in the competition between the companies. The planning of the investment strategy is therefore essential for the success of an individual company. For this, the company must take into account the nature of its liabilities as well as its solvency position.

Reserves: The Finnish statutory earnings-related pension scheme is partly funded. Only part of the total amounts of old age, disability and unemployment pensions are funded, and the part time and survivors' pensions are not funded at all. As a whole, about 25% of the total pension expenditure is paid by the funded part. The rest is financed as a pay-as-you-go system.

In the model, the average amounts of the funded pensions are calculated by age and sex. The technical reserves before increases at the end of the year (see below) are

$$L = \sum_{x,s} \mathcal{V}(s,x) \mathcal{E}(s,x) a(s,x),$$

where summation is by sex s and age x , $\mathcal{V}(s,x)$ is the number of pensioners or active workers, $\mathcal{E}(s,x)$ is the yearly average of the funded pension and $a(s,x)$ is the actuarial present value (APV) function depending on mortality and a discount rate of 3%. The definition of the APV function is different for the old age, disability and unemployment pensions and it is given in the actuarial basis confirmed by the Ministry of Social Affairs and Health.

Because the funded pension is based on the salaries of the insured persons, the reserves in the model are dependent on the development of the wage index. The reserves are also dependent on the technical interest rate in a way explained below. Otherwise, the reserves in the model are deterministic.

Besides the old age, disability and unemployment pension reserves, the model contains some other special reserves:

- an equalization reserve for buffering the yearly surplus/deficit of the insurance business
- a clearing reserve for the pay-as-you-go part of the pension expenditure
- a bonus reserve for the bonuses paid to customers as reductions of pension contributions.

Technical interest rate: Besides the discount rate of 3% used for calculating the actuarial present value functions, there is a higher rate, the technical interest rate whose value varies yearly. The amount of the reserves is increased at the end of each year corresponding to the difference between the technical interest rate and the 3% discount rate. This means that the technical interest rate determines the actual total interest rate for the reserves, and the 3% discount rate is its minimum value.

The technical interest rate (r^{tech}) is calculated by a formula dependent on the average solvency level of all the pension companies and funds. In the model, this formula cannot be used because the solvency level is one of the company's decision variables. Besides, the model contains only the company's own solvency and not its level in the whole TEL pension scheme. For these reasons, an approximation is used based on the investment variables, since these are correlated with the general solvency level of the system. The formula used in the model is

$$r_t^{tech} = \max\{3\%, \gamma_0 + \gamma_1 br_t + \gamma_2(\ln S_t - \overline{\ln S_t}) + \gamma_3(\ln P_t - \overline{\ln P_t})\},$$

where br_t denotes the bond yield, $\ln S_t$ the logarithm of the stock price index and $\ln P_t$ the logarithm of the property price index at time t . The $\overline{\ln S_t}$ and $\overline{\ln P_t}$ are the expected values of the variables $\ln S_t$ and $\ln P_t$ for year t calculated using average growth rates, defined in section 4. The formula was found using a separate, more detailed simulation model where the actual formula for the technical interest rate could be calculated. The approximation formula was fitted to the results of this model. The estimated parameters γ_i , $i = 0, \dots, 3$ are all positive suggesting, that the technical interest rate follows the long term bond yield and increases when the stock and property prices are above their expected values. The technical interest rate plays a crucial part in the model because, to a great extent, it determines the correlations between the investment variables and the reserves.

Cash flows: Besides the investment yields, money flows in the company as paid contributions and out as pension expenditure. The pension expenditure is calculated depending on the number of pensioners and the average funded pensions. The contributions depend on the total salaries of the insured persons. The contribution rates, which vary by age and sex, are confirmed by the Ministry of Social Affairs and Health. The combined cash flow depends on the development of the

wage index. The Finnish wage index from Statistics Finland is used to represent the average wage development.

Solvency capital: The solvency capital is the amount by which the total assets of the company exceed the sum of its reserves. It functions as a buffer against the variation of the investment results. Because the reserves of a pension insurance company must always be fully covered by its assets, a non-positive solvency capital means a bankruptcy. Therefore, the development of the solvency capital is an important factor in policy evaluations.

The legislation prescribes various minimum and target levels for the solvency capital of a pension insurance company. The basic quantity is the solvency border, which depends on the structure of the company's investment portfolio. The lower border of the target zone is twice the amount of the solvency border. The position of the solvency capital relative to these levels is an indicator of the solvency risk of the company.

3 Time series data

3.1 Historical data

As described above, the assets and liabilities of a Finnish pension insurance company can be approximately expressed in terms of the following seven economic factors

1. Three month Euribor sr ;
2. Five year German government bond br ;
3. Fixed mix stock index S ;
4. Fixed mix dividend yield Y^S ;
5. Property price index P ;
6. Rental yield Y^P ;
7. Wage index W .

Our data set consists of quarterly observations of these factors between 1991/1 - 2001/4. We have chosen such a short period because of the capital movement liberalization in the EU area during 1990, which resulted in significant changes in economic conditions. The data prior to 1991 corresponds to a more regulated economy.

Three month Euribor has been quoted only since the beginning of 1999. We extend this series backwards by using the German 3 month interest rate which is available from Datastream. Rental index R_t , used in (2) is reported only once a year. Therefore linear interpolation for $\ln R_t$ is used to fill in the gaps in the time series. For the wage index we use a seasonally adjusted series in the model. We thus obtain an approximation of the full set of quarterly data for all the seven factors between 1991/1 - 2001/4. We take this as a description of the statistical parameters in our asset and liability model; see Figure 1.

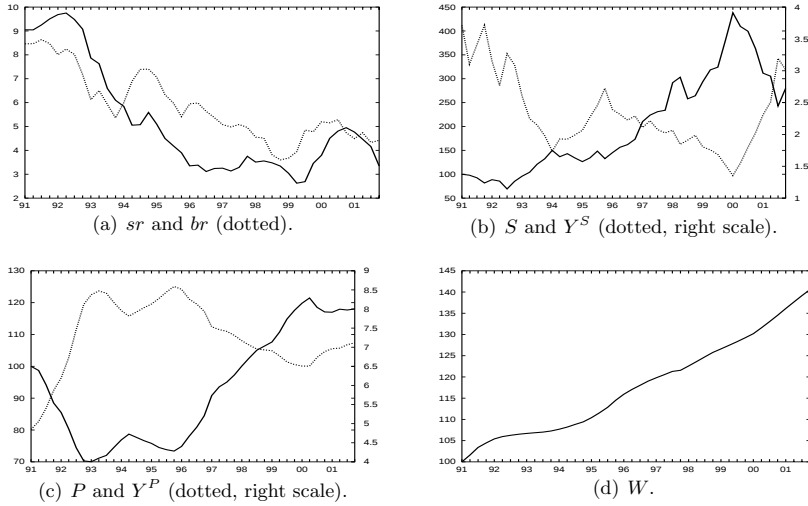


Figure 1: Historical time series.

3.2 Data transformations

The variations in dividend yield are roughly inversely proportional to the variations in the stock price index; see Figure 1. This is due to the fact that the dividend yield is, by definition, the dividend obtained per wealth invested in stocks where the latter follows the stock price index. Such multiplicative effects are not well modeled by the linear time series models that we are about to build. We will thus transform the dividend yield into the *dividend index*

$$D_t = S_t Y_t^S.$$

Similarly, instead of modeling the rental yield directly, we model the *rental index*

$$R_t = P_t Y_t^P.$$

We perform one more transformation, which is to take natural logarithms of all the seven time series, short-term interest rate, bond rate, stock price-, stock dividend-, property price-, property rental- and wage indices. This guarantees that the model never predicts negative indices or interest rates. The logarithmic time series are displayed in Figure 2.

3.3 Unit root tests

Before building an econometric model for the time series, we have to study their stationarity properties. We perform five unit root tests on x and its first difference. The tests are the augmented Dickey-Fuller test (ADF) (Dickey and Fuller, 1981), P_T and DF-GLS tests by Elliot et al. (1996) and Q_T and DF-GLSu tests, suggested by Elliot (1999). In the ADF test the lag length has been selected according to Schwarz information criterion with a maximum of five lags. The selected lag

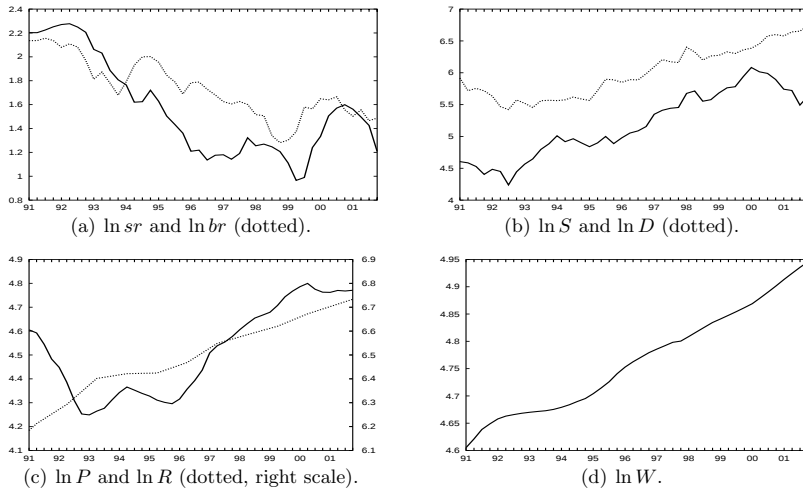


Figure 2: Transformed time series.

length is subsequently applied in the other reported tests. The P_T and DF-GLS tests are known to have improved power and better small sample properties compared to the ADF test (Elliot et al., 1996). The Q_T and DF-GLSu tests differ from the P_T and DF-GLS tests in the way the initial observation is treated in the derivation of the test statistics. In the Q_T and DF-GLSu tests the initial observation is drawn from its unconditional distribution whereas in the P_T and DF-GLS tests it is set to zero; see Elliot (1999). The results of the unit root tests were robust against different lag length selection methods, such as Akaike information criterion and General to simple, where the strategy is to select the highest significant lagged difference length e.g. in the ADF regression, less than or equal to some initial value. The value of $z(t)$ in Table 1 indicates the deterministic terms included in the unit root regressions. When $z(t) = 1$ a constant is included and with $z(t) = (1, t)$ a constant and a trend are included.

The results of the unit root test are displayed in Table 1. They clearly indicate that $\ln sr$, $\ln br$, $\ln S$ and $\ln D$ need to be differenced once in order to achieve stationarity. This confirms the findings of Hall et al. (1992), Sherris et al. (1999), Kanioura (2001) and Montoro (2001) regarding the interest rates. With $\ln P$ and $\ln R$ the evidence is not so clear. The non-stationarity of these series cannot be rejected at 5% significance level, except according to QT statistic the $\ln R$ is found to be trend stationary. The analysis of $\Delta \ln P$ and $\Delta \ln R$ cannot reliably reject the non-stationarity of the first differences either, which is not surprising considering the data used. The assumption, that the first difference of the logarithmic price index is stationary seems reasonable on economic grounds. Also, using quarterly data from 1970/1 to 1997/4, Barot and Takala (1998) concluded that $\Delta \ln P$ is stationary. The problem with property data are the long cycles that follow closely the general economic conditions in Finland. During the deep recession of the 1990's nominal property prices fell almost 40% by 1993. As a consequence of strong economic boom the property prices started to recover a few years later. These large long term fluctuations have caused the observed problems in unit root testing. We follow Barot and Takala (1998) and treat $\ln P$ as a difference stationary

process. Accordingly, $\ln R$ is treated similarly. All the unit root tests suggest that the logarithmic wage index is trend stationary. However, since all the other time series are treated as difference stationary we adopt the same strategy for $\ln W$. This assumption is supported by Q_T , DF-GLSu and ADF tests, Table 1. Moreover, Clements and Hendry (2001) argues that difference stationary models are considerably more adaptive forecasting tools compared to trend stationary models, when deterministic shifts occur during the forecast period.

Table 1: Unit root test statistics. ***, **, * indicate the rejection of the unit root null at 1%, 5%, 10% significance level.

Time series	Number of lags	Detrending, $z(t) =$	PT	DF-GLS	Q_T	DF-GLSu	ADF
$\ln sr$	1	(1)	11.07	-0.34	10.92	-1.35	-1.43
$\ln br$	1	(1)	10.48	-0.62	9.55	-1.67	-1.68
$\ln S$	0	(1,t)	9	-1.69	5	-1.75	-1.62
$\ln D$	0	(1,t)	20.85	-1.38	8.82	-2.32	-3
$\ln P$	1	(1,t)	6.74*	-2.08	3.51	-2.66	-3.22*
$\ln R$	1	(1,t)	5.88*	-2.32	2.60**	-2.7	-3.30*
$\ln W$	1	(1,t)	2.47***	-3.81***	1.31***	-3.88***	-3.5*
$\Delta \ln sr$	0	(1)	1.25***	-3.88***	2.06***	-3.82***	-3.97***
$\Delta \ln br$	0	(1)	0.69***	-5.18***	1.27***	-5.39***	-5.37***
$\Delta \ln S$	0	(1)	0.70***	-4.94***	1.34***	-5.85***	-6.22***
$\Delta \ln D$	1	(1)	0.44***	-6.04***	0.86***	-6.05***	-5.67***
$\Delta \ln P$	0	(1)	7.1	-1.2	5.48*	-2.78*	-2.44
$\Delta \ln R$	0	(1)	6.07	-1.25	5.33*	-2.5*	-2.42
$\Delta \ln W$	3	(1)	4.95	-1.37	1.69***	-2.92**	-5.2***

4 A VAR-model with specified drift

Denote the vector of logarithmic variables by

$$x_t = \begin{bmatrix} \ln sr_t \\ \ln br_t \\ \ln S_t \\ \ln D_t \\ \ln P_t \\ \ln R_t \\ \ln W_t \end{bmatrix}.$$

Based on the above observations, we assume that Δx_t is stationary. Our first attempt consists of building the VAR model

$$\Delta_d x_t = \sum_{i=1}^k A_i \Delta_d x_{t-i} + \epsilon_t, \quad \epsilon_t \sim N(0, \Sigma), \quad (3)$$

where $A_i \in \mathbb{R}^{7 \times 7}$, $\Sigma \in \mathbb{R}^{7 \times 7}$ and Δ_d denotes the shifted difference operator

$$\Delta_d x_t := \Delta x_t - d$$

with $d \in \mathbb{R}^7$; see also (Clements and Hendry, 1998, page 160). This format is convenient in that the parameter vector d determines the average drift in simulations. Indeed, if $\Delta_d x_t$ is stationary, (3) gives

$$E[\Delta_d x_t] = \left(\sum_{i=1}^k A_i \right) E[\Delta_d x_t],$$

so if (3) is free of unit roots, $E[\Delta_d x_t] = 0$, or

$$E[\Delta x_t] = d. \tag{4}$$

The above format is particularly natural for modeling indices.

Example 1 If x_t is the scalar process $\ln S_t$, and $A_i = 0$, (3) becomes

$$\Delta \ln S_t = d + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma),$$

which is a discrete-time version of the geometric Brownian motion model of stock price; see for example Hull (2000).

Moreover, according to Hendry and Clements (2001) modeling differences instead of levels gives protection against structural breaks in data generating process. VAR-models for differences of logarithms of economic time series have been built for example by Eitrheim et al. (1999) for the Central Bank of Norway.

Looking only at our (far from ideal) data, might suggest that there is a strong negative drift in the interest rates. However, we believe that $E[\Delta sr_t] = E[\Delta br_t] = 0$ in the long run, so we choose $d_{sr} = d_{br} = 0$ rather than estimating these parameters from the data. On the other hand, the dividend yield satisfies

$$\Delta \ln Y_t^S = \Delta \ln D_t - \Delta \ln S_t = \Delta_d \ln D_t - \Delta_d \ln S_t + d_D - d_S,$$

so if $\Delta_d \ln D_t$ and $\Delta_d \ln S_t$ follow (3), and if (3) is stationary, we have

$$E[\Delta \ln Y_t^S] = d_D - d_S.$$

Since there is no reason to believe that the dividend yield would have a consistent drift, one way or the other, we require $d_D = d_S$. Similar reasoning for the rental yield suggests $d_R = d_P$. It seems thus reasonable to assume that d has the form

$$d = \begin{bmatrix} 0 \\ 0 \\ d_S \\ d_S \\ d_P \\ d_P \\ d_W \end{bmatrix}. \tag{5}$$

Simply estimating d from the data would not result in a vector of the form (5). This is a clear case where, “expert” information seems more reliable than statistical information.

The choice of the values of the remaining drift parameters d_S , d_P and d_W is not quite as clear. Bewley (2000) and Landon-Lane (2000) have presented methods to restrict d when estimating the parameters of (3). In our case, one could make the restriction that d has to be of the form (5) and then the remaining parameters could be estimated from the data. In practice, however, pension insurance companies’ managers often have their own estimates for the average drifts for the future development of various time series. Such estimates are rarely based on statistical data alone. We take d_S , d_P and d_W as *user-specified parameters*. This not only provides a convenient way of incorporating expert views into the model, but it also simplifies the estimation process considerably; see below.

Our experiments below use

$$d = \begin{bmatrix} 0 \\ 0 \\ 0.0114 \\ 0.0114 \\ 0.007 \\ 0.007 \\ 0.009 \end{bmatrix}.$$

The value of d_S corresponds to 4.6% average of yearly log-return¹. It was argued in Barot and Takala (1998), that in the long run, there is no excess return in residential property prices over inflation rate. However, we believe that property investments will gain some real return over inflation and set the yearly growth rate of $\ln P$ to 2.8%, which is 0.8% over the inflation target of the European Central Bank. Finally, the yearly growth rate of $\ln W$ is set to 3.6%, which is 1.6% over the inflation target.

4.1 Estimation

Having specified the drift parameter d , it remains to choose the lag length k and the matrices A_i and Ω in model (3). We use PcFiml 9.0 for computing test statistics and parameter estimates; see Doornik and Hendry (1997). We start by selecting the appropriate lag length k in our VAR-model. Table 2 presents the lag length reduction test results starting from $k = 4$. Since, the sequential F-tests cannot reject the reductions to $k = 1$ and since the Schwarz (SC) and Hannan-Quinn (HQ) information criteria in Table 3 have minimum values at $k = 1$, we select $k = 1$.

Table 2: Lag length reduction tests

System reduction		Test statistic	p-value
from $k =$	to $k =$		
2	1	$F(49, 126) = 1.342$	0.0982
3	2	$F(49, 90) = 0.8912$	0.6663
4	3	$F(49, 55) = 1.0998$	0.3645

¹In our model, the usual return S_t/S_{t-1} is log-normally distributed with mean $\exp(d_S + \frac{1}{2}\sigma_S^2)$, where σ_S is the stock price volatility (Hull, 2000). Our choice of d_S gives roughly 7% average yearly return when $\sigma_S = 20\%$.

Table 3: Information criteria values

$k =$	np	log-likelihood	SC	HQ
1	49	1167.7489	-48.865	-50.115
2	98	1214.5735	-46.779	-49.28
3	147	1258.4598	-44.56	-48.31
4	196	1334.5532	-43.805	-48.805

Note: np = number of estimated parameters

We then estimate A_1 and Σ with the method of maximum likelihood (ML), starting with an unrestricted model and carry out an iterative procedure, where one insignificant parameter per iteration is removed and ML estimates for the remaining parameters are recomputed until all insignificant coefficients at 5% significance level have been removed. For a comparison and discussion of different model selection criteria; see Brüggemann and Lütkepohl (2000) and Lütkepohl (1991). The resulting estimates of A_1 , their standard errors $SE(A_1)$, residual correlation matrix C and residual standard deviations σ are

$$\begin{aligned}
A_1 &= 10^{-1} \begin{bmatrix} 3.665 & 4.887 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 7.392 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2.832 & 0 & 9.122 & 0 & 0 \\ 0 & 0 & 0.571 & 0 & 6.999 & 0 & 0 \\ 0 & -0.162 & 0 & 0 & 0 & 8.538 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0.825 & 8.424 \end{bmatrix} \\
SE(A_1) &= 10^{-1} \begin{bmatrix} 1.269 & 1.428 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3.234 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.842 & 0 & 2.811 & 0 & 0 \\ 0 & 0 & 0.273 & 0 & 0.868 & 0 & 0 \\ 0 & 0.080 & 0 & 0 & 0 & 0.619 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.246 & 0.521 \end{bmatrix} \\
C &= \begin{bmatrix} 1 & & & & & & \\ 0.0761 & 1 & & & & & \\ -0.0157 & -0.0201 & 1 & & & & \\ -0.1976 & -0.3084 & 0.4780 & 1 & & & \\ -0.0341 & -0.0683 & 0.3192 & -0.0275 & 1 & & \\ 0.1519 & 0.1084 & -0.1433 & -0.0111 & 0.0300 & 1 & \\ -0.1108 & -0.2356 & 0.0680 & 0.4057 & -0.0222 & -0.2113 & 1 \end{bmatrix} \\
\sigma &= 10^{-2} \begin{bmatrix} 7.5639 & 7.6128 & 11.1010 & 7.7848 & 1.9880 & 0.4381 & 0.1869 \end{bmatrix}
\end{aligned}$$

The likelihood ratio test of over-identifying restrictions $\chi^2(38) = 37.26[0.503]$, clearly accepts the made reductions. Table 4 reports the equation residual test results, where the numbers are the p-values of the test statistics. The reported tests are the F -test for 4th-order residual autocorrelation, χ^2 normality test, F -test for autocorrelated squared residuals and F -test for residual heteroscedasticity respectively, see Doornik and Hendry (1997). The results reveal some autocorrelation problems in $\Delta \ln sr$, $\Delta \ln S$ and $\Delta \ln D$, which is quite typical for financial time series data. The autocorrelation problems were persistent even in models with longer lag lengths, suggesting that VARMA models, (see e.g. Lütkepohl and Poskitt (1996), Lütkepohl and Claessen (1997))

could be more appropriate for the given data set. The residual normality assumption is rejected for $\Delta \ln R$ due to one negative outlier. The residual distribution of $\Delta \ln W$ has fatter tails compared to normal distribution, which would imply larger variance for the wage index forecasts than observed with normally distributed errors. This in turn would mean that the reserves would exhibit little larger fluctuations. However, we believe that this will not considerably affect the efficient asset allocation decisions obtained for the ALM problem in Section 8.

Table 4: Equation residual diagnostics

Equation	AR 1-4 F	Norm χ^2	ARCH 1-4 F	HET F
$\Delta \ln sr$	0.0069	0.5134	0.7492	0.0984
$\Delta \ln br$	0.2533	0.6668	0.7999	0.9590
$\Delta \ln S$	0.0150	0.2902	0.9625	0.7573
$\Delta \ln D$	0.0393	0.2335	0.8165	0.9430
$\Delta \ln P$	0.0630	0.3898	0.1250	0.1043
$\Delta \ln R$	0.0727	0.0000	0.5932	0.5895
$\Delta \ln W$	0.1273	0.0042	0.7088	0.6960

4.2 Simulation experiment

To test the long term behavior of the model we performed 250 twenty-year simulations with the estimated model (DVAR_{mod}) started from

$$x_0 = \ln \begin{bmatrix} 3.35 \\ 4.42 \\ 279.6 \\ 843.7 \\ 118.0 \\ 839.8 \\ 140.6 \end{bmatrix}, \quad x_{-1} = \ln \begin{bmatrix} 4.16 \\ 4.33 \\ 242.9 \\ 776.0 \\ 117.7 \\ 831.3 \\ 139.1 \end{bmatrix},$$

which was the situation in the beginning of 2002. For comparison we performed equivalent simulations with a model (DVAR_{sys}) where all the regressors are retained and the drift parameters are estimated without restrictions. The outcomes of the simulations for the DVAR_{mod} and DVAR_{sys} models are displayed in Figures 3 and 4, respectively. The outcomes of the DVAR_{sys} simulations highlights the importance of correctly specifying the form of the drift vector, see Figure 4. The short term interest rate is rapidly declining throughout the simulation period due to estimated negative drift. Also, the log interest rate spread, log dividend and rental yields are trending due to differences in the estimated drifts of the underlying time series. The average drifts of the DVAR_{mod} were what we desired, but in some respects the model behaves strangely: in many scenarios the logarithmic short rate, log interest rate spread, log dividend and rental yields have deviated unrealistically from their usual values; see Figure 3.

Despite the fact that the drift parameters for interest rates were set to zero in DVAR_{mod} , the short term interest rate is generally declining throughout the simulation period. This shows how strongly the future distributions depend on the initial values of the variables in a VAR model. This phenomenon could be avoided by changing the initial values, so that the simulation starts from some neutral conditions, as suggested by Lee and Wilkie (2000), but then the relevant market information essential to the present investment decisions is lost.

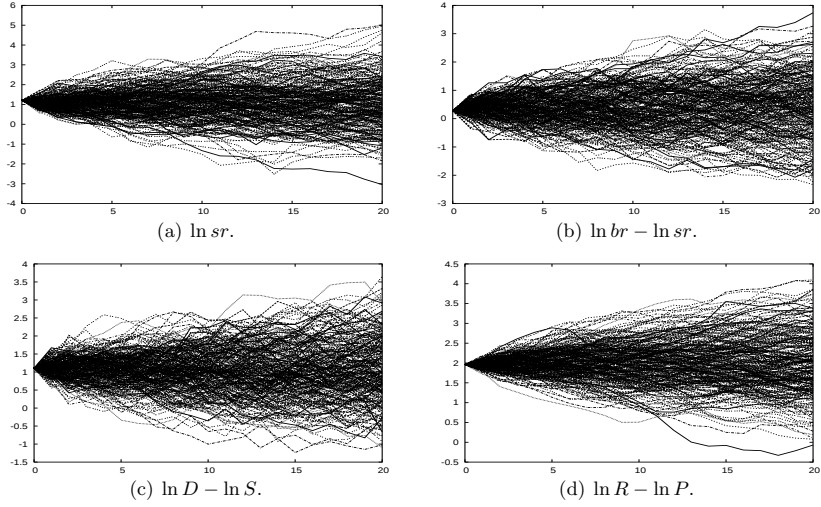


Figure 3: $DVAR_{mod}$ simulations.

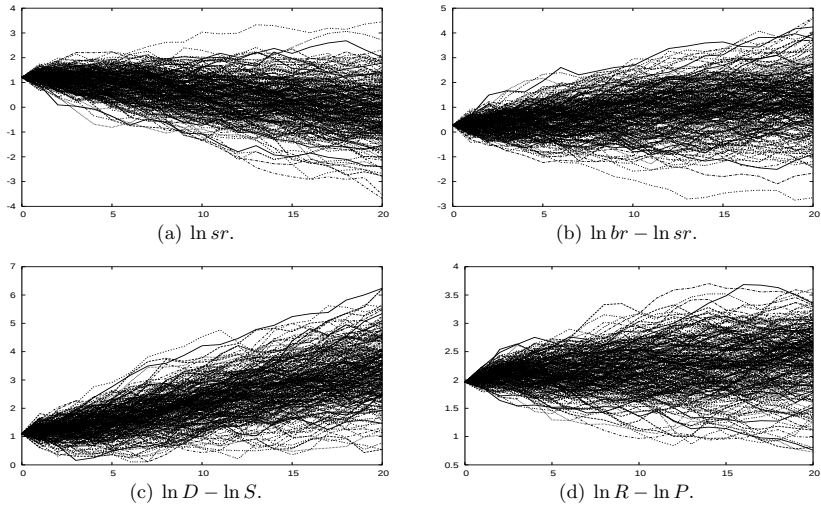


Figure 4: $DVAR_{sys}$ simulations.

The problems even after specifying the drift vector result from the fact that the above models only look at the differences Δx_t and completely ignores the actual values of x_t , which is what we are really interested in. As demonstrated by the simulations, the average values for x_t are largely determined by the initial values x_0 and x_{-1} , which may be poor estimates of the future. This leads

us to consider VEqC-models, which avoid these shortcomings.

5 A VEqC-model with specified drift and cointegration relations

A *vector equilibrium correction model* is obtained from the VAR-model for differences by adding an “equilibrium correction term” to the right-hand side of the equations. In the case of our drift-specified VAR-model we get the model

$$\Delta_d x_t = \sum_{i=1}^k A_i \Delta_d x_{t-i} + \alpha(\beta' x_{t-1} - \mu) + \epsilon_t, \quad \epsilon_t \sim N(0, \Sigma), \quad (6)$$

where $\beta \in \mathbb{R}^{7 \times l}$, $\mu \in \mathbb{R}^l$ and $\alpha \in \mathbb{R}^{7 \times l}$. The additional term takes into account the long-term behavior of x_t around statistical equilibria described by the linear equations $\beta' x = \mu$. It is assumed that, in the long run,

$$E[\beta' x_t] = \mu, \quad (7)$$

and that if x_t deviates from the equilibria (due to shocks in economic conditions) it will tend to move back towards them. The matrix α determines the speed of adjustment towards the equilibria. In this sense, VEqC-models incorporate long-run equilibrium relationships (often derived from economic theory) with short-run dynamic characteristics deduced from historical data. VEqC-models for logarithms of economic time series have been built for example by Eitrheim et al. (1999) for the Central Bank of Norway and by Anderson et al. (2000) for the Federal Reserve Bank of St. Louis. The results of Eitrheim et al. (1999) indicate that the inclusion of equilibrium-correction feedbacks may improve the forecast accuracy of VAR-models for differences, especially in the long-run.

The equilibrium correction term is particularly convenient when modeling interest rates.

Example 2 If x_t is the scalar process $\ln r_t$, $d = 0$ and $A_i = 0$, the model becomes

$$\Delta \ln r_t = \alpha(\ln r_{t-1} - \mu) + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma).$$

With $\alpha < 0$, this is a discrete-time version of the mean-reverting interest rate model of Black and Karasinski (1991). With $\alpha = -1$, we obtain the memoryless model

$$\ln r_t = \mu + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma).$$

Besides mean reversion effects for interest rates, the equilibrium correction term is useful also in controlling long-term averages of the interest rate spread and the yields, which behaved unrealistically in the simple VAR-model of the previous section. We follow the two-step methodology of Engle and Granger (1987) by first specifying the equilibrium relations, and then estimating A , α and Σ from the data. By (4) and (7), the matrix β must satisfy the consistency condition

$$\beta' d = \beta' E[\Delta x_t] = E[\beta' \Delta x_t] = E[\Delta(\beta' x_t)] = 0. \quad (8)$$

Based on our experiences with the VAR-model, we propose the following four equilibrium relations.

1. $\ln sr_t = \mu_{mr}$. Similarly to mean reverting interest rate models, this suggests that, in the long run, the short rate drifts towards certain equilibrium level. Although, the unit root tests in Section 4 indicated $\ln sr$ to be non-stationary (due to the changing economic conditions during the 1990's), many studies have concluded that interest rates are mean reverting and stationary in the long run; see for example Wu and Zhang (1996) and Fama and Bliss (1987).
2. $\ln br_t - \ln sr_t = \mu_{sp}$. This relation means that the “geometric interest rate spread” $\frac{br_t}{sr_t}$ has a long term equilibrium value. Various studies have concluded that the difference $br - sr$ of the long and short term interest rates is stationary; see for example Campbell and Shiller (1987), Bradley and Lumpkin (1992) and Hall et al. (1992). Campbell et al. (1997) found also the logarithmic transformation $\ln(1 + br) - \ln(1 + sr)$ of the interest rate spread to be stationary. To our knowledge, only (Kanioura, 2001, page 5) has studied the geometric spread. She found it to be stationary in the United States.
3. $\ln D_t - \ln S_t = \mu_{dy}$. Writing this as $\ln Y_t^S = \mu_{dy}$, we see that it corresponds to the existence of an equilibrium value for the dividend yield. This is supported by the findings of Campbell and Shiller (1988), Campbell et al. (1997) and Wilkie (1986).
4. $\ln R_t - \ln P_t = \mu_{ry}$. Similarly to dividend yield, this can be written as $\ln Y_t^P = \mu_{ry}$ which corresponds to a stationary rental yield.

These choices correspond to

$$\beta = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \quad (9)$$

This satisfies (8) with any d of the form (5). In fact, any d satisfying (8) for this choice of β has to be of the form (5). Similarly to the drift vector d , we take the equilibrium values μ in the cointegration relations as *user-specified parameters*.

The historical values of βx_t are displayed in Figure 5. In our time frame 1991/1 – 2001/4, these series do not pass the stationarity tests on conventional significance levels. However, we believe that these series will be stationary in the long run.

In our experiments, we use the β in (9) and

$$\mu = \begin{bmatrix} \ln 3.7 \\ \ln 1.2 \\ \ln 2.5 \\ \ln 7.0 \end{bmatrix}.$$

This corresponds to long term equilibrium values of 3.7%, 1.2%, 2.5% and 7% for short rate, geometric interest rate spread, dividend yield and rental yield, respectively.

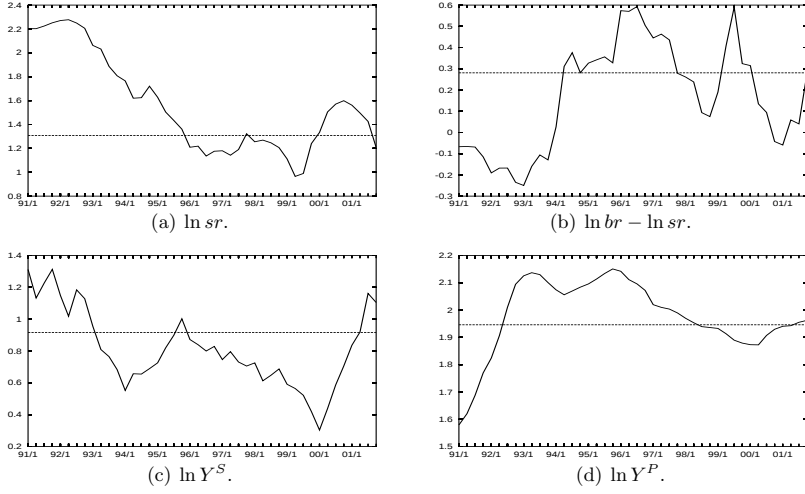


Figure 5: Historical values of the cointegration vectors. The horizontal lines mark the expected equilibrium levels μ .

5.1 Estimation

Having specified the drift parameter d and the cointegration relations, we find the maximum likelihood estimates of the remaining parameters A , α and Σ , using the iterative model reduction procedure like in section 4.1. This results in the following values.

$$A_1 = 10^{-1} \begin{bmatrix} 3.672 & 3.467 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2.855 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -59.11 \\ 0 & 0 & -2.425 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.629 & 0 & 3.617 & 0 & 0 \\ 0 & -0.209 & 0 & 0 & -0.663 & 8.533 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0.638 & 8.712 \end{bmatrix}$$

$$SE(A_1) = 10^{-1} \begin{bmatrix} 1.222 & 1.466 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.469 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 26.50 \\ 0 & 0 & 0.836 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.231 & 0 & 1.065 & 0 & 0 \\ 0 & 0.082 & 0 & 0 & 0.222 & 0.682 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.246 & 0.589 \end{bmatrix}$$

$$\begin{aligned}
\alpha = 10^{-1} & \begin{bmatrix} 0 & 0.964 & 0 & 0 \\ -1.061 & -1.499 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1.449 & 0 \\ -0.238 & 0 & 0 & 0.637 \\ 0 & 0.080 & 0 & 0 \\ 0 & 0 & -0.024 & 0 \end{bmatrix} \\
SE(\alpha) = 10^{-1} & \begin{bmatrix} 0 & 0.415 & 0 & 0 \\ 0.414 & 0.781 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0.352 & 0 \\ 0.067 & 0 & 0 & 0.160 \\ 0 & 0.032 & 0 & 0 \\ 0 & 0 & 0.011 & 0 \end{bmatrix} \\
C = & \begin{bmatrix} 1 & & & & & & \\ 0.1308 & 1 & & & & & \\ -0.0606 & 0.1332 & 1 & & & & \\ -0.2764 & -0.1978 & 0.5302 & 1 & & & \\ 0.0092 & 0.0184 & 0.354 & -0.0796 & 1 & & \\ 0.0911 & 0.1526 & -0.1329 & -0.0675 & -0.0102 & 1 & \\ -0.174 & -0.1929 & 0.0641 & 0.3586 & 0.0113 & -0.2396 & 1 \end{bmatrix} \\
\sigma = 10^{-2} & \begin{bmatrix} 7.3288 & 7.4645 & 10.804 & 7.8889 & 1.6589 & 0.4085 & 0.1826 \end{bmatrix}
\end{aligned}$$

Some remarks on α :

- The first cointegration vector has a significant negative coefficient in bond rate equation. This can be interpreted as a reaction to interest rate expectations: when the short rate is above its long term average μ_{sr} , it is expected to decline in the long run, which causes a drop in the bond rate. Similarly, short rate being below its average, pushes the bond rate up. The first cointegration vector appears also in the property price equation with a negative sign. This implies that low interest rates increases property prices and vice versa. This may result from the fact that low interest rates decrease the loan servicing costs, which encourages people/companies to invest in properties.
- The geometric interest rate spread enters the short rate equation with a positive sign (which is in line with the expectations hypothesis; see e.g. Campbell et al. (1997)), and the bond rate equation with a negative sign (which in turn contradicts the expectations hypothesis). If the spread is above its average value μ_{sp} , these terms push the interest rates closer to each other, and if it is below μ_{sp} , they push the short rate down and the bond rate up. This is in line with the findings of Campbell (1995) and (Campbell et al., 1997, Section 10.2.2).
- The third cointegration vector, the log dividend yield, appears in the dividend index equation with a negative sign, so large values of the dividend yield cause a decrease in the dividend index, and vice versa. This effect is similar to findings in Campbell and Shiller (1988). It causes the dividend index to follow the movements in the stock price index, keeping the dividend yield in a reasonable range.

- The log rental yield enters the property price equation with a positive sign, with the interpretation that large values of the rental yield anticipates an increase in property prices, and vice versa.

Table 5 reports the equation residual test results for the VEqC-model. Again, the numbers denote the p-values of the different test statistics. The results are similar to those obtained with the VAR-model in Section 4. The tests reveal some autocorrelation problems in $\Delta \ln sr$, $\Delta \ln S$ and $\Delta \ln D$ and again the normality assumption is rejected in the residuals of $\Delta \ln R$ and $\Delta \ln W$ due to few outliers.

Table 5: VEqC-model equation residual diagnostics

Equation	AR 1-4 F	Norm χ^2	ARCH 4 F	HET F
$\Delta \ln sr$	0.0029	0.3471	0.9232	0.8173
$\Delta \ln br$	0.1698	0.8191	0.9030	0.9862
$\Delta \ln S$	0.0070	0.2516	0.8918	0.8439
$\Delta \ln D$	0.0037	0.3504	0.6661	0.8877
$\Delta \ln P$	0.0551	0.2290	0.6195	0.8498
$\Delta \ln R$	0.1725	0.0004	0.5072	0.9563
$\Delta \ln W$	0.0530	0.0028	0.5326	0.9123

5.2 Simulation experiment

We computed 250 twenty-year simulations with the above VEqC-model (VEqC_{mod}) started from the same initial values as in Section 4.2. The equilibrium correction terms effectively control the interest rates and the yields that were problematic in the DVAR models; see Figure 6. The mean reversion apparent in Figure 6 is caused by the inclusion of the equilibrium correction terms, which considerably reduce the variance of the interest rates as well as the dividend and rental yields.

6 Forecast tests

We test and compare the performance of our VEqC_{mod} model with three rival models in an out-of-estimation-sample forecast experiment. The forecast test period covers seven new quarterly observations from 2002/1 to 2003/3. Although the short test period does not allow us to draw significant conclusions concerning the forecast performance of different models, the test gives us an indication how the developed VEqC_{mod} model would have performed in volatile financial markets during 2002–2003. For comparison, we report the results of the forecast tests for DVAR_{mod} , DVAR_{sys} and VEqC_{sys} models. VEqC_{sys} denotes an unrestricted vector equilibrium correction system where the drifts and the equilibrium values for the equilibrium correction relations of Section 5 are estimated from historical data without restrictions.

We compare the four models' forecast accuracy by performing tests for structural stability (see Lütkepohl (1991)) during the forecast period $T+1, \dots, T+h$, where T denotes the forecast origin and $h=7$ is the length of the forecast horizon. We calculate a test statistic of the form

$$\lambda_h = \sum_{i=1}^h u'_{T+i} \Sigma^{-1} u_{T+i} \sim \chi^2(Nh), \quad (10)$$

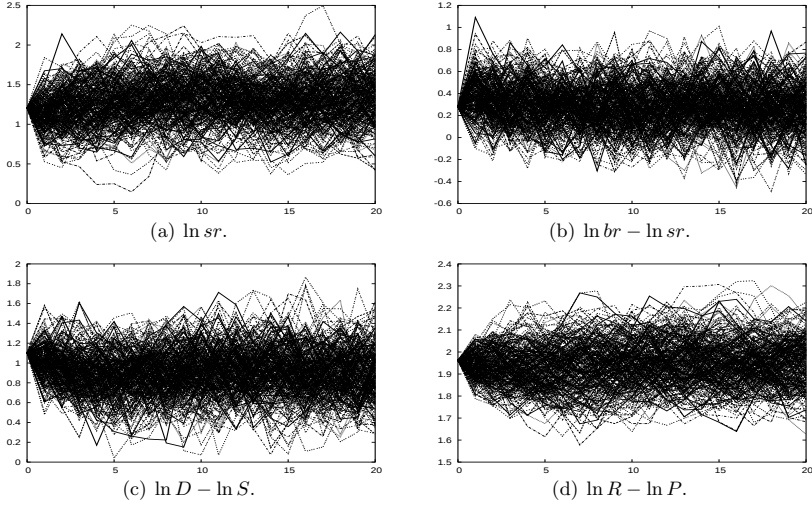


Figure 6: VEqC_{mod} simulations.

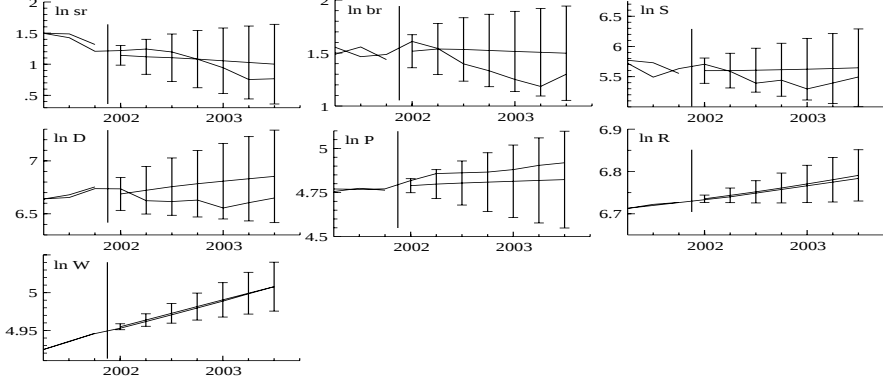
where u_{T+i} can be interpreted as the 1-step ahead forecast errors as we move through the forecast period, Σ is the residual covariance matrix as in (3) and $N = 7$ is the dimension of the model. The null hypothesis for the test is that the process generating $(\Delta x_{T+1}, \dots, \Delta x_{T+h})$ is the same as that which generated $(\Delta x_1, \dots, \Delta x_T)$ and the hypothesis is rejected if the forecasts differ too much from the actually observed values, see e.g. Lütkepohl (1991). The values of the approximate χ^2 and an F variant test statistics, where the unknown quantities in (10) are replaced by estimated values (see e.g. Lütkepohl (1991) or Clements and Hendry (1998)) together with the p-values of the tests are reported in Table 6.

Table 6: Forecast test statistics

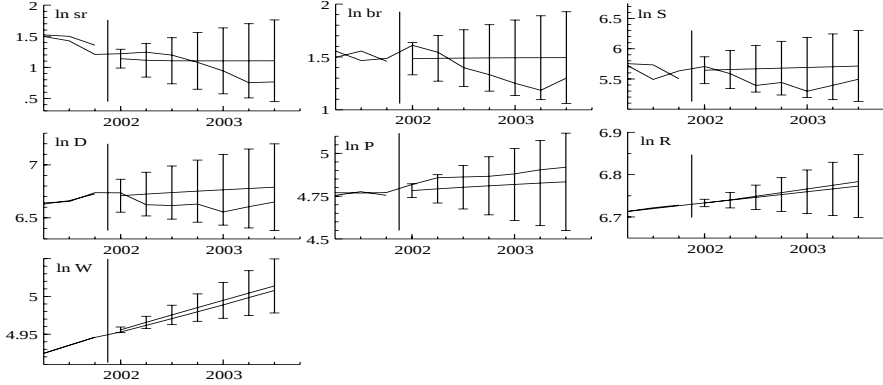
Model	χ^2	p-value	F	p-value
DVAR _{sys}	46.11	0.59	0.94	0.58
DVAR _{mod}	54.94	0.26	1.12	0.35
VEqC _{sys}	94.39	0.00	1.93	0.03
VEqC _{mod}	47.47	0.54	0.97	0.55

The null hypothesis of structural stability is rejected only for VEqC_{sys} at the 1% and 5% significance levels according to χ^2 and F statistics, respectively, and VEqC_{mod} and DVAR_{sys} seems to produce the smallest forecast errors during the test period. Figure 7 displays 1- to 7-step ahead forecasts with their approximate 95% confidence intervals and the actually observed values of x_t during the forecast period for all the four models. The main reason for the forecast failure of VEqC_{sys} is apparent from Figure 7(c), where the actually observed value for $\ln sr$ and $\ln S$ are clearly outside their forecast confidence intervals. For the other three models there are no striking differences in the forecast performance, although the confidence intervals for the VAR-models (Figures 7(a) – 7(b)) are wider than for the VEqC-models (Figures 7(c) – 7(d)). These findings together with the results of the long term simulation experiments of Sections 4.2 and 5.2 give support to our approach of

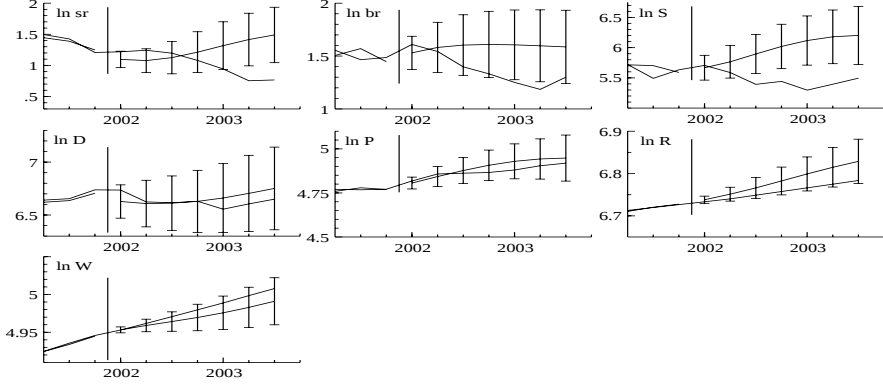
specifying the drifts and equilibrium values for the equilibrium relations. It is also worth noting that the VEqC-models contain more economic insight than the pure DVAR models and especially the parameters of the VEqC_{mod} model are easy to interpret. In the next Section we will use the developed VEqC_{mod}-model in long term return and liability simulation.



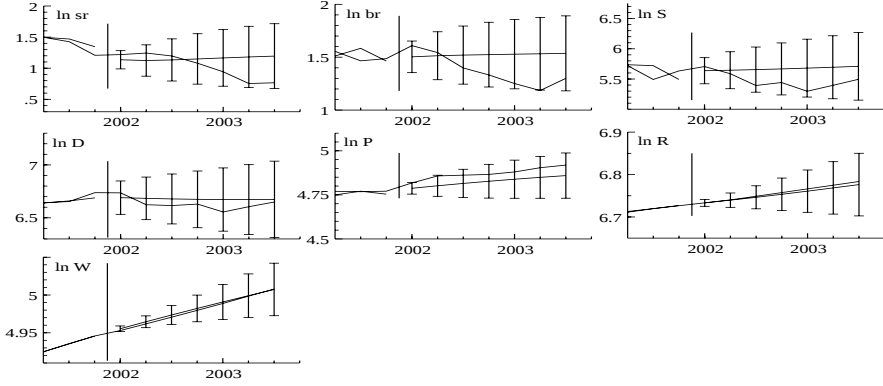
(a) DVAR_{sys}.



(b) DVAR_{mod}.



(c) VEqC_{sys} .



(d) VEqC_{mod} .

Figure 7: 1- to 7-step ahead forecasts with 95% confidence intervals and observed values of x_t for the four models.

7 Long-term return and liability simulations

We will first study the behavior of total returns of the considered asset classes. This will be done by performing 1000 twenty-year simulations with the above model started from the initial values given in Section 4.2, and computing the corresponding yearly total returns for each asset class. The total return of an asset is defined as the sum of the change-in-value and the cash income components.

Using the approximations of Section 2.1, we get the following expressions for the total returns

$$\begin{aligned}
\text{Cash: } & \sqrt{sr_t sr_{t-1}}, \\
\text{Bonds: } & \left(\frac{1 + br_{t-1}}{1 + br_t} \right)^D + \frac{1}{2}(br_{t-1} + br_t), \\
\text{Stocks: } & \frac{S_t}{S_{t-1}} + \frac{1}{2} \left(\frac{D_{t-1}}{S_{t-1}} + \frac{D_t}{S_t} \right), \\
\text{Property: } & \frac{P_t}{P_{t-1}} + \frac{1}{2} \left(\frac{R_{t-1}}{P_{t-1}} + \frac{R_t}{P_t} \right) - 0.03, \\
\text{Loans: } & \frac{1}{2}(br_{t-1} + br_t).
\end{aligned}$$

Figure 8 displays the development of the means and standard deviations of the yearly total returns for cash, bonds, stocks, property and loans, based on 1000 twenty-year simulations. During the first few years, the average returns go through large changes, after which they converge to their equilibrium values. The variations in the average returns towards the end of the simulation horizon are simply effects of the finite sample size.

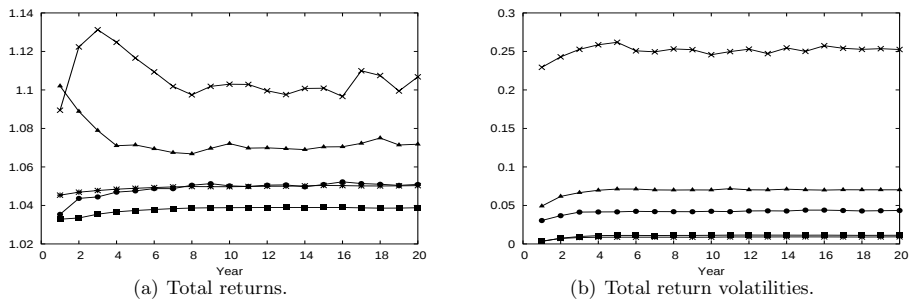


Figure 8: Simulated total returns and their volatilities for Cash (■), Bonds (●), Stocks (×), Property (▲) and Loans (✱) with VEqC(mod)

The initial conditions for the simulation affect the returns and correlation structures considerably in the first few years as the model starts from a disequilibrium. The correlations between the total returns of the asset classes at the end of years 1 and 20 are shown in Figure 9. Even the signs of some correlation coefficients change between years 1 and 20. Once the model converges back to an equilibrium the yearly correlation structure becomes stable.

The reserves and cash-flows can be computed based on the values of the time series according to the rules outlined in Section 2.2. The results of 1000 20-year simulations are displayed in Figure 10. The reserves grow consistently over time but the variation in the projected cash flows at the horizon are substantial. The decreasing trend in cash flows after a few years results from the retirement of the large age groups.

	Cash	Bonds	Stocks	Property	Loans
Cash	1				
Bonds	-0.530	1			
Stocks	0.014	-0.153	1		
Property	-0.123	-0.007	0.553	1	
Loans	0.530	-0.999	0.153	0.007	1

(a) Year 1.

	Cash	Bonds	Stocks	Property	Loans
Cash	1				
Bonds	0.240	1			
Stocks	-0.060	-0.133	1		
Property	-0.340	-0.113	0.515	1	
Loans	0.763	0.229	0.036	-0.218	1

(b) Year 20.

Figure 9: Simulated total return correlations.

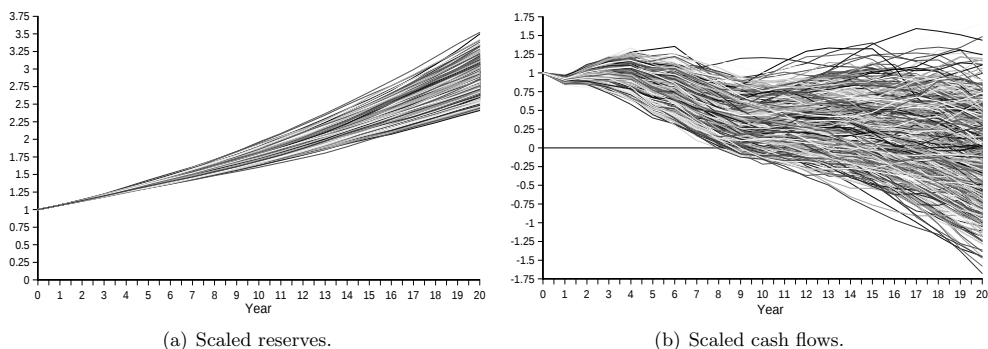


Figure 10: Simulated Reserves and Cash flows.

8 Evaluation of dynamic portfolio allocation strategies

Given a stochastic model for the asset returns and the liabilities, we would like to compute the distribution of a company's solvency in the future. This is a nontrivial task since the future solvency depends on the values of the reserves and cash flows as well as the investment strategies that the company employs now and in all the possible states of the world in the future. Moreover, the Finnish legislation imposes complicated regulations that the companies' must take into consideration in their strategic asset allocation, see Hilli et al. (2003). We evaluate the company's long term solvency by considering two widely studied decision rules for dynamic portfolio allocation, namely *fixed-mix* and *portfolio insurance* strategies; see e.g. Perold and Sharpe (1988), Cesari and Cremonini (2003).

In fixed-mix strategy the portfolio is always rebalanced to a given (fixed) asset distribution (mix). So a fixed-mix strategy is given by a vector of numbers giving the fixed percentages that the asset allocations should satisfy now and in the future. In the present setting this is a vector of five numbers giving the portfolio weights for cash, bonds, stocks, property and loans.

The proportion of loans in the investment portfolio each year is kept fixed at 0.145% of the

reserves, which corresponds to 11.5% weight in the initial asset portfolio. This means that the proportion of loans in the portfolio decreases if the total value of the investments increases faster than the value of the reserves and vice versa. We will examine by simulation the performance of different asset mixes obtained by different combinations of the following weights applied to the remaining portfolio.

$$\begin{aligned}\text{Cash: } w_C &\in \{0, 0.01, \dots, 0.03\}, ; \\ \text{Stocks: } w_S &\in \{0, 0.025, \dots, 0.5\}, ; \\ \text{Property: } w_P &\in \{0.1, 0.15, \dots, 0.4\}, ; \\ \text{Bonds: } w_B &= 1 - w_C - w_S - w_P.\end{aligned}$$

The upper bounds for stocks and property are statutory restrictions and the bond investments are chosen so that the total weights in the remaining portfolio sum up to 100%.

The used portfolio insurance (PI) strategy is based on the constant proportion portfolio insurance framework of Perold and Sharpe (1988) and Black and Jones (1988). The portfolio weights for cash and property are varied according to the same rules as in the fixed-mix case. The rest of the wealth is allocated between the more liquid assets, bonds and stocks. The proportion of stocks in the portfolio at time t is given by,

$$w_{S,t} = \begin{cases} \min \left\{ (1 - w_C - w_P) \min \left\{ \rho \left(\frac{W_t - L_t}{W_t} \right), 1 \right\}, 0.5 \right\} & \text{if } W_t - L_t \geq 0, \\ 0 & \text{if } W_t - L_t < 0, \end{cases}$$

where ρ is a risk tolerance parameter indicating how the proportion invested in stocks increases with the company's solvency ratio, $(W_t - L_t)/W_t$, where W_t and L_t denote the values of the company's assets and reserves in the beginning of year t , respectively. The percentage invested in stocks is a constant multiple of the company's solvency ratio, which was close to 22% initially, with higher values of ρ resulting in higher stock market allocations and again at most 50% of the total wealth can be invested in stocks. When the company's wealth W_t is less than the value of its reserves L_t (floor) the stock market allocation is set to zero and the remaining wealth is invested in bonds. In general, PI strategies are fairly realistic decision rules for pension insurance companies because they allocate more wealth to risky assets, stocks, when the companies' solvency ratios improve and reduce the stock market exposure as the companies approach insolvency.

For each fixed-mix portfolio combination and for PI strategies with varying risk tolerances, $\rho \in \{1, 1.5, \dots, 20\}$, we perform 1000 simulations with a 20-year time horizon. Figure 11 displays the average solvency capital/reserves ratio in 20 years versus the insolvency probability for each fixed-mix portfolio and PI strategy based on the sample of 1000 scenarios. Insolvency means that the solvency capital has become negative at least once during the 20-years. The best PI strategies clearly dominate the best performing fixed-mix strategies at all reasonable risk levels.

The lower boundaries of the clouds of points can be interpreted as the *efficient frontiers* of the fixed-mix and PI strategies, which are displayed in Figure 12. For insolvency probabilities between 0–5% the efficient PI strategies improve the average solvency capital-reserves ratios from 15 to 35% compared to fixed-mix portfolios, and the performance gap between the two methods decreases as the bankruptcy risk increases.

The compositions of efficient fixed-mix portfolios and initial portfolio weights for efficient PI strategies are displayed in Figure 13 for varying insolvency probabilities. All the efficient fixed-mix

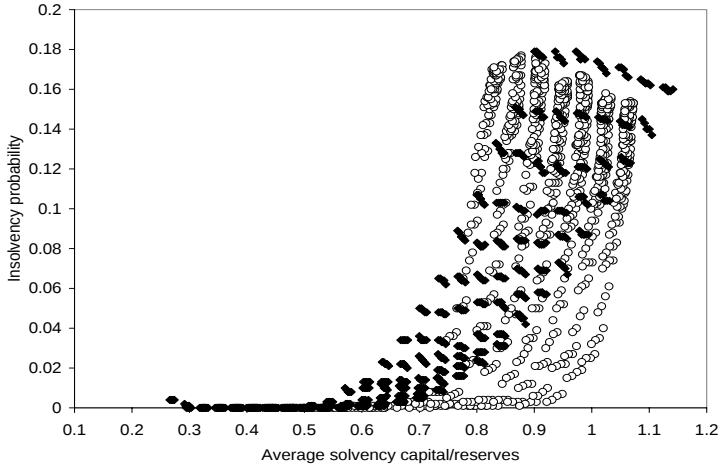


Figure 11: Average solvency capital/reserves at the horizon against insolvency probability for different fixed-mix ($\blacklozenge$) and PI ($\circ$) strategies.

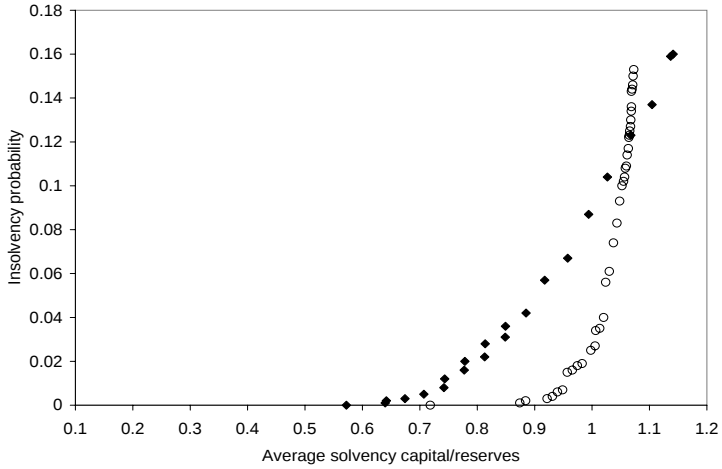


Figure 12: Efficient frontier of fixed-mix ($\blacklozenge$) and PI ($\circ$) strategies.

and initial PI portfolios have at least 35% of the wealth invested in property and in some portfolios a small fraction of money invested in cash. Due to PI strategies' ability to react dynamically to changing solvency situations, the initial stock market allocation in PI portfolios can be kept much higher compared to fixed-mix portfolios with similar insolvency probabilities.

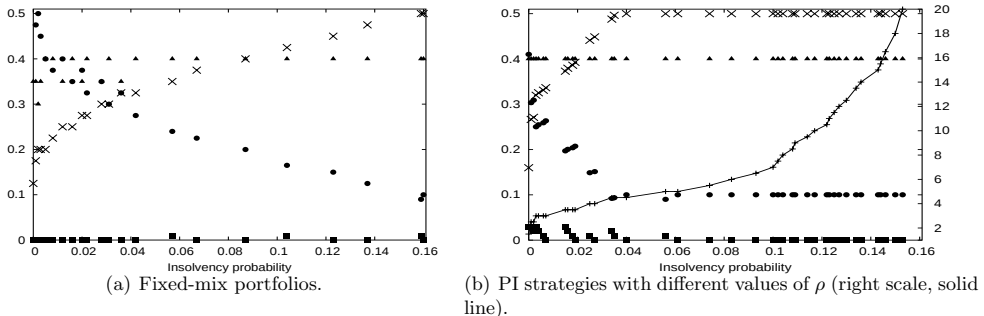


Figure 13: Efficient portfolios' weights for Cash (■), Bonds (●), Stocks (×) and Property (▲).

9 Conclusions

This paper proposed a stochastic model for future development of the main stochastic factors that are of interest in asset liability management of a Finnish pension insurance company. Some of the most critical parameters in the model, namely the drift rates and certain long-term equilibrium values, are taken as user-specified instead of relying completely on statistical information. This is essential when the available data displays drifts or other characteristics that are believed to change in the future. The cointegration relations allow the modeling of causalities derived from economic theories and/or statistical studies.

The presented model should, of course, not be taken as the only possible model of reality. We would like to emphasize more the general model building procedure that combines statistical information with user-specified characteristics. In the proposed approach many variations are possible. For example, the model for bond investments returns is only a crude approximation of reality and it could probably be made more accurate by more careful analysis. Also, in modeling property investments, one could try to replace the residential property price index by something that better describes the value of the property investments of a Finnish pension insurance company. Finally, instead of modeling nominal values of the time series, one could incorporate inflation to the model and model the real values of the time series². An alternative possibility for modeling the relation between inflation and interest rates would be to use the fact that inflation is strongly related to logarithmic changes in the wage index which is already in our model. However, our model does not show any direct relation between the interest rates and the wage index. This is probably due to the fact that the three month Euribor series was extended backwards by using data from Germany while the wage index is that of Finland.

The decision rules for dynamic asset allocation considered in Section 8 give a fairly good approximation of the company's expected future solvency and bankruptcy risks, but are not of course the best way of designing investment strategies for a pension insurance company. In Hilli et al. (2003) we describe an optimization model that better takes into account the freedom to update the portfolio in the future as well as all the relevant constraints that the Finnish legislation imposes. In this approach, known as stochastic programming, one tries to find the best initial portfolio given

²The authors are grateful to Professor David F. Hendry for suggesting this.

the objectives of the company, various portfolio (and other) constraints and the stochastic model for the uncertain factors.

References

- Anderson, R. G., Hoffman, D., and Rasche, R. H. (2000). A Vector Error-Correction Forecasting Model of the U.S. Economy. Working paper 1998-008b, current version July 2000, Federal Reserve Bank of St. Louis.
- Barot, B. and Takala, K. (1998). House Prices and Inflation: A Cointegration Analysis for Finland and Sweden. Discussion paper 12/98, Bank of Finland.
- Bewley, R. (2000). Controlling spurious drift in macroeconomic forecasting models. The University of New south Wales, <http://economics.web.unsw.edu.au/people/rbewley/papers/Drift.pdf>.
- Black, F. and Jones, R. (1988). Simplifying portfolio insurance for corporate pension plans. *Journal of Portfolio Management*, 14:33–37.
- Black, F. and Karasinski, P. (1991). Bond and option pricing when short rates are lognormal. *Financial Analysts Journal*, July-August:52–59.
- Boender, C. G. E., van Aalst, P., and Heemskerk, F. (1998). Modelling and management of assets and liabilities of pension plans in the netherlands. In Ziemba, W. T. and Mulvey, J. M., editors, *Worldwide asset and liability management*, pages 561–580. Cambridge University Press.
- Bradley, M. G. and Lumpkin, S. A. (1992). The treasury yield curve as a cointegrated system. *The Journal of Financial and Quantitative Analysis*, 27(3):449–463.
- Brüggemann, R. and Lütkepohl, H. (2000). Lag selection in subset var models with an application to a u.s. monetary system. Discussion paper, Humboldt-University.
- Campbell, J. Y. (1995). Some lessons from the yield curve. *Journal of Economic Perspectives*, 9(3):129–152.
- Campbell, J. Y., Lo, A. W., and MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton University Press, USA.
- Campbell, J. Y. and Shiller, R. (1987). Cointegration and tests of present value models. *Journal of Political Economy*, 95(5):1062–1088.
- Campbell, J. Y. and Shiller, R. (1988). Stock prices, earnings and expected dividends. *Journal of Finance*, 43(3):661–676.
- Cesari, R. and Cremonini, D. (2003). Benchmarking portfolio insurance and technical analysis: a monte carlo comparison of dynamic strategies of asset allocation. *Journal of Economic Dynamics and Control*, 27:987–1011.
- Clements, M. P. and Hendry, D. F. (1998). *Forecasting economic time series*. Cambridge University Press, Cambridge.

- Clements, M. P. and Hendry, D. F. (1999). *Forecasting non-stationary economic time series*. MIT Press, Cambridge, MA. With a foreword by Katarina Juselius.
- Clements, M. P. and Hendry, D. F. (2001). Forecasting with difference-stationary and trend-stationary models. *Econometrics Journal*, 4(1):1–19.
- Dert, C. L. (1998). A dynamic model for asset liability management for defined benefit pension funds. In Ziemba, W. T. and Mulvey, J. M., editors, *Worldwide asset and liability management*, pages 501–536. Cambridge University Press.
- Dickey, D. A. and Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49:1057–1072.
- Doornik, J. and Hendry, D. (1997). *Modelling Dynamic Systems Using PcFiml 9.0 for Windows*. International Thomson Publishing, London, UK.
- Eitrheim, O., Husebø, T. A., and Nymoen, R. (1999). Equilibrium-correction versus differencing in macroeconomic forecasting. *Economic Modelling*, 16:515–544.
- Elliot, G. (1999). Efficient tests for a unit root when the initial observation is drawn from its unconditional distribution. *International Economic Review*, 40:767–783.
- Elliot, G., Rothenberg, T. J., and Stock, J. H. (1996). Efficient tests for an autoregressive unit root. *Econometrica*, 64:813–836.
- Engle, R. F. and Granger, C. W. J. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica*, 55(2):251–276.
- Fama, E. F. and Bliss, R. R. (1987). The information in long-maturity forward rates. *The American Economic Review*, 77(4):680–692.
- Hall, A. D., Anderson, H. M., and Granger, C. W. J. (1992). A cointegration analysis of treasury bill yields. *The Review of Economics and Statistics*, 74, No. 1:116–122.
- Harris, G. (1999). Markov chain monte carlo estimation of regime switching vector autoregressions. *ASTIN Bulletin*, 29:47–79.
- Hendry, D. F. and Clements, M. (2001). Economic forecasting: some lessons from recent research. Working paper no. 82, European Central Bank.
- Hendry, D. F. and Doornik, J. A. (1997). The implications for economic modelling of forecast failure. *Scottish Journal of Political Economy*, 44:437–461.
- Hilli, P., Koivu, M., Pennanen, T., and Ranne, A. (2003). A stochastic programming model for asset liability management of a finnish pension company. *Stochastic Programming E-Print Series*.
- Hull, J. C. (2000). *Options, Futures, and Other Derivatives*. Prentice Hall, fourth edition.
- Johansen, S. (1995). *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford University Press, New York.

- Kanioura, A. (2001). A cointegration analysis of us interest rates. Discussion papers series: 2001-03, The University of Sheffield.
- Landon-Lane, J. (2000). Imposing restrictions on drift in a var. The University of New south Wales.
- Lee, P. and Wilkie, A. (2000). A comparison of stochastic asset models. In *Proceedings of the 10th AFIR Colloquium*, pages 407–445.
- Lütkepohl, H. (1991). *Introduction to multiple time series analysis*. Springer-Verlag, Berlin.
- Lütkepohl, H. and Claessen, H. (1997). Analysis of cointegrated varma processes. *Journal of Econometrics*, 80:223–239.
- Lütkepohl, H. and Poskitt, D. (1996). Specification of echelon-form varma models. *Journal of Business and Economic Statistics*, 14:69–79.
- Montoro, V. F. (2001). Expectations and the behaviour of spanish treasury bill rates. Discussion papers in economics and econometrics no. 0112, University of Southampton.
- Perold, A. F. and Sharpe, W. F. (1988). Dynamic strategies for asset allocation. *Financial Analysts Journal*, 44:16–27.
- Ranne, A. (1998). The finnish stochastic investment model. In *Transactions of the 26th ICA*, pages 213–238.
- Sheris, M., Tedesco, L., and Zehnirith, B. (1999). Investment returns and inflation models: Some australian evidence. *British Actuarial Journal*, V, Part I:237–268.
- Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, 48(1):1–48.
- Wilkie, A. D. (1986). A stochastic investment model for actuarial use. *Transactions of the Faculty of Actuaries*, 39:341–403.
- Wilkie, A. D. (1995). More on a stochastic asset model for actuarial use. *British Actuarial Journal*, 1:777–964.
- Wright, I. D. (1998). A stochastic asset model using vector auto-regression. Actuarial research report no. 108, City University.
- Wu, Y. and Zhang, H. (1996). Mean reversion in interest rates: New evidence from a panel of OECD countries. *Journal of Money, Credit and Banking*, 28, No. 4(1):604–621.
- Yakoubov, Y., Teeger, M. H., and Duval, D. B. (1999). Stochastic investment model for asset and liability management. AFIR Colloquium. www.actuaries.org/members/en/AFIR/colloquia/Tokyo/papers.cfm.

A stochastic programming model for asset liability management of a Finnish pension insurance company

Petri Hilli[†] Matti Koivu[†] Teemu Pennanen

Department of Management Science

Helsinki School of Economics

PL 1210

00101 Helsinki, Finland

[hilli,koivu,pennanen]@hkkk.fi

Antero Ranne

Actuarial Department

Ilmarinen Mutual Pension Insurance Company

antero.ranne@ilmarinen.fi

Abstract

This paper describes a stochastic programming model that was developed for asset liability management of a Finnish pension insurance company. In many respects the model resembles those presented in the literature, but it has some unique features stemming from the statutory restrictions for Finnish pension insurance companies. Particular attention is paid to modeling the stochastic factors, implementation and to numerical testing. Out-of-sample tests clearly favor the strategies suggested by our model over static fixed-mix and dynamic portfolio insurance strategies.

1 Introduction

Stochastic programming has proven to be an efficient tool in designing effective strategies in wealth- and asset liability management in practice. This is due to its ability to cope with the dynamics and complex constraint structures usually inherent in such problems. In principle, stochastic programming is not tied to any particular form of objective function or model of the stochastic factors as long as the distribution of the stochastic factors is independent of the decision variables in the model. Successful applications of stochastic programming to asset liability management have been reported at least in Nielsen and Zenios (1996), Cariño et al. (1998), Cariño and Ziemba (1998), Høyland (1998), Consigli and Dempster (1998), Kouwenberg (2001), and Geyer et al. (2003). See also the excellent volumes by Ziemba and Mulvey (1998) and Zenios and Ziemba (2004) and the references therein. For a general introduction to stochastic programming see the (COSP) stochastic programming site: www.stoprog.org.

[†]Financial support from the Foundation for the Helsinki School of Economics under grants number 9981114 and 9981117 is gratefully acknowledged.

This paper describes a stochastic programming model and its computer implementation for asset liability management of a Finnish pension insurance company. Finnish pension companies manage large investment funds and, like most pension companies in Europe, they are facing a large number of retiring policyholders at around 2010–2020. Our model addresses a long term dynamic investment problem where the aim is to cover the uncertain future liabilities with dynamic investment strategies. The assets are considered as the aggregate investment classes of cash, bonds, stocks, property and loans to policyholders. In addition to investment decisions, our model looks for optimal bonus payments and it takes explicitly into account various portfolio and transaction restrictions as well as some legal restrictions coming from the complex pension system in Finland which is based on the defined benefit rule. The legal restrictions form a unique part of the model not present in earlier applications of stochastic programming.

We pay particular attention to describing the uncertain factors in the model which include investment returns, cash-flows, and the technical reserves used in the definition of the statutory restrictions. This is important since the output of a stochastic programming model depends usually heavily on the underlying model for the stochastic factors. Our approach consists of first building a time series model, which is then discretized into scenario trees appropriate for numerical solution of the optimization model. This is convenient for the user who only needs to come up with an appropriate econometric description of the stochastic factors. The discretization is fully automated and hidden from the user.

The model was implemented and tested against static fixed-mix and dynamic portfolio insurance strategies. Fixed-mix strategies are simple decision rules that always rebalance the investment portfolio to maintain fixed asset proportions. The used portfolio insurance strategies are based on the constant proportion portfolio insurance framework of Perold and Sharpe (1988) and Black and Jones (1988), where the proportion of risky assets is kept as a constant multiple of the difference between the portfolio value and a protective floor. If the portfolio value hits or falls below the floor all the funds are invested in riskless assets.

These decision strategies are not the most realistic models for the behavior of a real pension insurance company. However, they are often used for various simulation purposes in practice, which motivates their use as benchmarks. Other, more sophisticated but computation-intensive, choices of benchmarks have been used in Høyland (1998), Kouwenberg (2001); see also Fleten et al. (2002). We used the out-of-sample testing procedure recommended e.g. by Dardis and Mueller (2001) of Tillinghast-Towers Perrin. In the tests, the strategies based on our stochastic programming model clearly outperform both the fixed-mix and portfolio insurance strategies. Similar results have been obtained for the more sophisticated benchmarks in Høyland (1998), Kouwenberg (2001), Fleten et al. (2002).

The rest of the paper is organized as follows. A mathematical model of the ALM problem is presented in Section 2. A model for the underlying stochastic factors and its discretizations (scenario trees) are described in Section 3. Section 4 outlines a computer implementation of our model and reports the results of numerical tests including an extensive out-of-sample simulation.

2 The optimization model

Our model is a multistage stochastic program where a sequence of decisions (asset allocations etc.) is interlaced with a sequence of observations of random variables (asset returns etc.). At each stage, decisions are made based on the information revealed up to that point, so the decision variables at a

stage are functions of the random variables observed up to that stage. This kind of interdependent dynamics of information and decisions is typical in sequential decision making under uncertainty, which is what ALM and many other wealth management problems are; see for example Ziemba and Mulvey (1998) or Föllmer and Schied (2002).

The decision stages are indexed by $t = 0, 1, \dots, T$, where $t = 0$ denotes the present time, and the set of assets is indexed by $j \in J$, with

$$J = \{\text{cash, bonds, stocks, property, loans to policyholders}\}.$$

The decision variables characterize the asset management strategy as well as the company's solvency situation and the bonus strategy. Uncertainties result from random future investment returns as well as from random cash flows and technical reserves described below. There are several constraints stemming from the regulations of the Finnish pension system. The objective is to optimize the development of the company's solvency situation as described by the Ministry of Social Affairs and Health as well as the amount of bonuses paid to policyholders.

We will first describe the asset management model, followed by the model of statutory restrictions and finally the objective. Decision variables are random variables for all t except for $t = 0$. For parameters, randomness is indicated explicitly.

2.1 Asset management

Asset management constitutes a central part of the model. The following formulation is fairly standard in asset management applications of stochastic programming.

Inventory constraints describe the dynamics of holdings in each asset class:

$$\begin{aligned} h_{0,j} &= h_j^0 + p_{0,j} - s_{0,j} \\ h_{t,j} &= R_{t,j}h_{t-1,j} + p_{t,j} - s_{t,j} \quad t = 1, \dots, T-1, \quad j \in J, \end{aligned}$$

where

$$\begin{aligned} h_j^0 &= \text{initial holdings in asset } j, \\ R_{t,j} &= \text{return on asset } j \text{ over period } [t-1, t] \text{ (random)} \end{aligned}$$

are parameters, and

$$\begin{aligned} p_{t,j} &= \text{(nonnegative) purchases of asset } j \text{ at time } t, \\ s_{t,j} &= \text{(nonnegative) sales of asset } j \text{ at time } t, \\ h_{t,j} &= \text{holdings in asset } j \text{ in period } [t, t+1] \end{aligned}$$

are decision variables. As usual, we do not allow portfolio rebalancing at the horizon, which is why the index t goes only up to $T-1$ in the inventory constraints. The company does not have control over the loans since, according to the Finnish law the policyholders have the right to borrow money from the company against their paid pension premiums. The amount invested in loans is thus determined by the policyholders. Holdings in loans are stochastic and we will assume them to be proportional to the technical reserves; see Section 2.2.1 below.

Budget constraints guarantee that the total expenses do not exceed revenues:

$$\begin{aligned} \sum_{j \in J} (1 + c_j^p) p_{0,j} + H_{-1} &\leq \sum_{j \in J} (1 - c_j^s) s_{0,j} + F_0, \\ \sum_{j \in J} (1 + c_j^p) p_{t,j} + \tau_t H_{t-1} &\leq \sum_{j \in J} (1 - c_j^s) s_{t,j} + \sum_{j \in J} D_{t,j} h_{t-1,j} + F_t \quad t = 1, \dots, T-1, \end{aligned}$$

where

$$\begin{aligned} c_j^p &= \text{transaction cost for buying asset } j, \\ c_j^s &= \text{transaction cost for selling asset } j, \\ \tau_t &= \text{length of period } [t-1, t] \text{ in years,} \\ H_{-1} &= \text{transfers to the bonus reserve a year before stage } t = 0, \\ D_{t,j} &= \text{dividend paid on asset } j \text{ over period } [t-1, t] \text{ (random),} \\ F_t &= \text{cash flows in period } [t-1, t] \text{ (random)} \end{aligned}$$

are parameters and

$$H_t, \quad t = 0, \dots, T-1 = \text{transfers to the bonus reserve per year during period } [t, t+1]$$

are decision variables. The net cash flow F_t is the difference between pension contributions and expenditures during period $[t-1, t]$. The company can pay a proportion of its accumulated wealth as bonuses to its policyholders. These bonuses are paid as reductions of the pension contributions. The amount of the total bonuses is determined at the end of each year, and the sum is transferred to the bonus reserve. The whole bonus reserve is then paid out during the following year. For periods longer than one year, we assume that H_t is kept constant throughout the period, hence $\tau_t H_{t-1}$ gives the value of bonuses paid to policyholders during period $[t-1, t]$.

Portfolio constraints give bounds for the allowed range of portfolio weights:

$$l_j w_t \leq h_{t,j} \leq u_j w_t \quad t = 0, \dots, T-1, \quad j \in J,$$

where

$$w_t = \sum_{j \in J} h_{t,j} = \text{total wealth at time } t = 0, \dots, T-1,$$

and

$$\begin{aligned} l_j &= \text{lower bound for the proportion of } w_t \text{ in asset } j, \\ u_j &= \text{upper bound for the proportion of } w_t \text{ in asset } j \end{aligned}$$

are parameters whose values are given in Table 1.

Table 1: Lower and upper bounds for investment proportions

j	l_j	u_j
Cash	0.01	1
Bonds	0	1
Stocks	0	0.5
Property	0	0.4

The upper bounds for stocks and property are statutory restrictions. The lower bound for cash investments is set to guarantee sufficient liquidity.

The total wealth w_t at stage $t = 0, \dots, T - 1$ is computed after portfolio rebalancing. At the horizon, there is no rebalancing so

$$w_T = \sum_{j \in J} (R_{T,j} + D_{T,j}) h_{T-1,j} + F_T - \tau_T H_{T-1}.$$

Transaction constraints bound the sales and purchases to a given fraction of w_t :

$$\begin{aligned} p_{t,j} &\leq \tau_t b_j^p w_t \quad t = 0, \dots, T - 1, \quad j \in J, \\ s_{t,j} &\leq \tau_t b_j^s w_t \quad t = 0, \dots, T - 1, \quad j \in J, \end{aligned}$$

where

$$\begin{aligned} b_j^p &= \text{upper bound for purchases of asset } j \text{ per year as a fraction of total wealth,} \\ b_j^s &= \text{upper bound for sales of asset } j \text{ per year as a fraction of total wealth} \end{aligned}$$

are parameters. The values of b_j^p and b_j^s are displayed in Table 2. The tight rebalancing restrictions for property are set because of illiquidity of the Finnish property markets. For other asset classes the yearly rebalancing is restricted to be at most 20% of the total wealth. These restrictions model the policies of the company as well as the requirement that the size of transactions should be kept at levels that do not affect market prices.

Table 2: Upper bounds for transactions

j	b_j^p	b_j^s
Cash	0.2	0.2
Bonds	0.2	0.2
Stocks	0.2	0.2
Property	0.01	0.01

2.2 Statutory restrictions

The statutory restrictions for Finnish pension insurance companies are quite strict, and they form a unique part of our stochastic programming model. Besides imposing constraints on the decision variables, these rules form the basis for defining the objective function in our model.

2.2.1 Solvency capital

The Finnish pension insurance companies are obliged to comply with several restrictions described in the legislation, government decrees or regulations given by the Ministry of Social Affairs and Health. A fundamental restriction is that the assets of a company must always cover its *technical reserves* L_t , which corresponds to the present value of future pension expenditure discounted with the so called “technical interest rate”. A detailed description for determining the value of L_t is given in Koivu et al. (2003). The assets include, besides the total amount of investments w_t , a transitory item of the net amount of other debts and credits in the balance sheet. This relatively small amount is calculated approximately as a fixed proportion c^G of the technical reserves. The difference

$$C_t = w_t + c^G L_t - L_t = w_t - (1 - c^G) L_t$$

of assets and the technical reserves is called the *solvency capital*. If at any time, C_t becomes negative, the company is declared bankrupt.

2.2.2 Solvency limits

Besides bankruptcy, ($C_t \leq 0$), there are several target levels that have been set to characterize the pension insurance companies’ solvency situation. These levels form an early warning system, so that the company and the supervising authorities can take action before a bankruptcy actually happens. A fundamental concept in the system is the *solvency border* $\tilde{B}_t$, defined in (1) below. If the solvency capital C_t falls below this limit the financial position is considered to be at risk, and the company is required to present to the authorities a plan for recovering a safe position. In addition, the company is not allowed to give any bonuses to its policyholders.

The target zone for the ratio $C_t/\tilde{B}_t$ is $[2, 4]$. In this zone, the financial position of a company is considered to be quite good. There is still discussion about how strictly the upper limit should be observed (in practice, no company has yet exceeded the upper limit). Therefore, we will ignore the upper limit in the model.

The concept of the solvency border corresponds to the solvency requirements in the European Union (EU) insurance directives. There is, however, an essential difference in the calculation method. The Finnish solvency border is based on the investment portfolio of a company. The fluctuation of the solvency capital is mainly caused by the investment market, and therefore the risk of going bankrupt is strongly dependent of the company’s investment risk. The starting point of the Finnish system is that the probability of ruin in one year at the solvency border should be approximately 2.5%, and therefore the value of the border is required to be dependent on the investment portfolio. In contrast, the EU directives take no account of the company’s investments. It is widely regarded that the EU regulations are insufficient, and a project is now established to renew the EU solvency requirements. The solvency border $\tilde{B}_t$ is

$$\tilde{B}_t = \left(a \sum_{j \in J} m_j h_{t,j} + b \sqrt{\sum_{j,k \in J} \sigma_{j,k} h_{t,j} h_{t,k}} \right) \frac{(L_t + H_t)}{w_t}, \quad (1)$$

where $a = -0.972/100$, $b = 1.782/100$, and the parameters

$$m = \begin{bmatrix} 0.18 \\ 0.66 \\ 6.20 \\ 3.70 \\ 0.72 \end{bmatrix}, \quad \sigma = \begin{bmatrix} 0.93 & 0.01 & 3.08 & 1.05 & -0.02 \\ 0.01 & 11.47 & 12.80 & -3.62 & 11.19 \\ 3.08 & 12.80 & 460.51 & 91.50 & 9.67 \\ 1.05 & -3.62 & 91.50 & 176.55 & -1.31 \\ -0.02 & 11.19 & 9.67 & -1.31 & 11.18 \end{bmatrix}$$

give the means and covariances for the asset classes (in the order: cash, bonds, stocks, property, loans to policyholders), according to the government decree, of one-year rate of returns over the technical interest rate. For asset classes like stocks, the parameter $\sigma_{j,j}$ is substantially larger than for safer classes like bonds. In reality, the values of m and σ are not fixed for eternity, but are updated by the Ministry of Social Affairs and Health on an irregular basis. The current values were set in 1999. We keep the values m and σ fixed in our optimization model partly because of the infrequent updating and also because any uncertainty in these parameters would be hard to model. Note that $\tilde{B}_t$ is a nonconvex function of the variables in the model.

2.2.3 Upper bound for bonuses

Finnish pension insurance companies compete with each other by paying out bonuses to their policyholders. To attract new customers companies would like to keep the amount of bonuses very high, but because the pension system is statutory, the government has aimed to restrict the amount of bonuses so that a sufficient proportion of the assets is preserved in the system to guarantee future pensions. Therefore, the Ministry of Social Affairs and Health imposes a formula for the maximum amount of each year's bonus transfers. The maximum depends on the solvency capital C_t and the solvency border $\tilde{B}_t$ of the company according to the formula

$$\tilde{H}_t^{\max} = \phi(C_t/\tilde{B}_t) (C_t - \tilde{B}_t)$$

where $\phi(z)$ is a piecewise linear function which has the minimum value of 0 when $z \leq 1$ and the maximum value of 0.04 when $z \geq 4$. It follows that $\tilde{H}_t^{\max}$ is also a nonconvex function of the variables in the model.

2.2.4 Convex approximations

In the optimization model, the nonconvex solvency border is replaced by

$$B_t = a \sum_{j \in J} m_j h_{t,j} + b \sqrt{\sum_{j,k \in J} \sigma_{j,k} h_{t,j} h_{t,k}},$$

which is convex in the variables. We have $B_t \geq \tilde{B}_t$ since $(L_t + H_t)/w_t \leq 1$ unless the company is bankrupt. Replacing $\tilde{B}_t$ by B_t in the model, makes the constraints in the model more restrictive, so we will stay on the safe side, except when the company is bankrupt. In the case of bankruptcy, the solvency border is underestimated by a factor of $(L_t + H_t)/w_t$.

We will also replace the nonconvex function $\tilde{H}_t^{\max}$ by a convex approximation, namely,

$$H_t^{\max} = 0.03 \max\{C_t - B_t, 0\}.$$

This is based on the fact that the historical average of $\phi(z)$ has been close to 0.03.

2.3 Objective function

There are many possibilities for measuring the performance of a company by an objective function. Natural candidates would be expected utility of wealth or solvency capital under various utility functions. Here, we will describe a utility function that takes explicitly into account the unique features of the Finnish pension system.

As described in Section 2.2.2, the Ministry of Social Affairs and Health measures pension insurance companies' solvency situation by the ratio $C_t/\tilde{B}_t$ of the solvency capital and the solvency border. The Ministry defines four zones according to which companies' solvency situation is classified:

$$\begin{aligned} C_t/\tilde{B}_t \in [2, \infty) &: \text{target} \\ C_t/\tilde{B}_t \in [1, 2) &: \text{below target} \\ C_t/\tilde{B}_t \in [0, 1) &: \text{crisis} \\ C_t/\tilde{B}_t \in (-\infty, 0) &: \text{bankrupt.} \end{aligned}$$

We replace $\tilde{B}_t$ throughout by its convex approximation B_t given above, and we define three shortfall variables:

$$\begin{aligned} SF_{t,1} &\geq 2B_t - C_t & t = 1, \dots, T-1, \\ SF_{t,2} &\geq B_t - C_t + H_t/0.03 & t = 1, \dots, T-1, \\ SF_{t,3} &\geq -C_t & t = 1, \dots, T, \end{aligned}$$

each of which gives the amount by which a zone is missed. These will be penalized in the objective function. The inequality for $SF_{t,2}$ incorporates the constraint

$$H_t \leq H_t^{\max}$$

for bonus transfers. The penalty for $SF_{t,2}$ will be chosen large enough to guarantee that, at the optimum, the upper bound is satisfied.

For $t = 0, \dots, T-1$, the state of the company will be evaluated by the utility function

$$u(C_t, B_t, H_t, L_t) = C_t/L_t - \sum_{z=1}^3 \gamma_z SF_{t,z}/L_t + u^b(H_t/L_t),$$

where γ_z are positive parameters and u^b is a nondecreasing concave function that will be specified according to the preferences of the company. However, the choice of u^b has to be made in accordance with the penalty parameter γ_2 in order to guarantee that the upper bound for H_t is not violated at the optimum. At stage T , the utility is measured by

$$u_T(C_T, L_T) = C_T/L_T - \gamma_3 SF_{T,3}/L_T.$$

The overall objective function in our model is the discounted expected utility

$$E^P \left\{ \sum_{t=1}^{T-1} d_t u(C_t, B_t, H_t, L_t) + d_T u_T(C_T, L_T) \right\},$$

where d_t is the discount factor for stage t . The problem is to maximize this expression over all the decision variables and subject to all the constraints described above.

2.4 Problem summary

Deterministic parameters:

- h_j^0 = initial holdings in asset j ,
- c_j^p = transaction cost for buying asset j ,
- c_j^s = transaction cost for selling asset j ,
- l_j = lower bound for wealth in asset j as a fraction of total wealth,
- u_j = upper bound for wealth in asset j as a fraction of total wealth,
- b_j^p = upper bound for purchases of asset j per year as a fraction of total wealth,
- b_j^s = upper bound for sales of asset j per year as a fraction of total wealth,
- c^G = the amount of transitory items as a fraction of the technical reserves,
- a = the (negative) weight for the return component in the solvency border,
- b = the weight for the standard deviation component in the solvency border,
- m_j = mean yearly return of asset j according to the government decree,
- $\sigma_{j,k}$ = covariance of one-year returns according to the government decree,
- τ_t = length of period $[t - 1, t]$ in years,
- γ_z = penalty parameters in the objective function,

Stochastic parameters:

- $R_{t,j}$ = return on asset j over period $[t - 1, t]$,
- $D_{t,j}$ = dividend paid on asset j over period $[t - 1, t]$,
- F_t = cash flows from period $[t - 1, t]$,
- L_t = technical reserves at time t ,

Decision variables:

- $h_{t,j}$ = holdings in asset j from period t to $t + 1$,
- $p_{t,j}$ = purchases of asset j at time t ,
- $s_{t,j}$ = sales of asset j at time t ,
- w_t = total wealth at time t ,
- H_t = transfers to bonus reserve at time t ,
- C_t = solvency capital at time t ,
- B_t = solvency border at time t ,
- $SF_{t,z}$ = shortfall from zone z at time t .

The stochastic programming model is

$$\begin{aligned}
& \text{maximize} \quad E^P \left\{ \sum_{t=1}^{T-1} d_t u(C_t, B_t, H_t, L_t) + d_T u_T(C_T, L_T) \right\} \\
& \quad h_{0,j} = h_j^0 + p_{0,j} - s_{0,j}, \\
& \quad h_{t,j} = R_{t,j} h_{t-1,j} + p_{t,j} - s_{t,j}, \\
& \quad p_{t,j}, s_{t,j} \geq 0, \\
& \quad \sum_{j \in J} (1 + c_j^p) p_{0,j} + H_{-1} \leq \sum_{j \in J} (1 - c_j^s) s_{0,j} + F_0, \\
& \quad \sum_{j \in J} (1 + c_j^p) p_{t,j} + \tau_t H_{t-1} \leq \sum_{j \in J} (1 - c_j^s) s_{t,j} + \sum_{j \in J} D_{t,j} h_{t-1,j} + F_t, \\
& \quad w_t = \sum_{j \in J} h_{t,j}, \\
& \quad l_j w_t \leq h_{t,j} \leq u_j w_t, \\
& \quad p_{t,j} \leq \tau_t b_j^p w_t, \\
& \quad s_{t,j} \leq \tau_t b_j^s w_t, \\
& \quad C_t = w_t - (1 - c^G) L_t, \\
& \quad B_t \geq a \sum_{j \in J} m_j h_{t,j} + b \sqrt{\sum_{j,k \in J} \sigma_{j,k} h_{t,j} h_{t,k}}, \\
& \quad SF_{t,1} \geq 2B_t - C_t, \\
& \quad SF_{t,2} \geq B_t - C_t + 100H_t/3, \\
& \quad SF_{t,3} \geq -C_t, \\
& \quad \text{for all } t = 1, \dots, T-1, \quad j \in J, \\
& \quad w_T = \sum_{j \in J} (R_{T,j} + D_{T,j}) h_{T-1,j} + F_T - \tau_T H_{T-1}, \\
& \quad C_T = w_T - (1 - c^G) L_T, \\
& \quad SF_{T,3} \geq -C_T, \\
& \quad (h, p, s, w, H, C, B, SF) \in \mathcal{N}
\end{aligned}$$

where P is the probability distribution of the random parameters, E^P denotes the expectation operator, and the constraints are required to hold almost surely with respect to P . The symbol $\mathcal{N}$ stands for the subspace of nonanticipative decision rules, so the decision variables of a given stage are not allowed to depend on random variables whose values are observed only in later stages. The model is a convex program that is nonlinear both in the objective and the constraints. There are 19 decision variables in each stage $t = 0, \dots, T-1$ (recall that for loans to policyholders, $h_{t,j}$, $p_{t,j}$ and $s_{t,j}$ are determined by L_t) and 3 in the last stage.

3 Scenario tree generation

The probability distribution P of the random parameters is an important input to the model, and the solution will depend on it in an essential way. We assume that the random parameters follow

the stochastic model developed in Koivu et al. (2003). Numerical solution of the optimization problem is then done through discretization of the continuous distribution as in Pennanen and Koivu (2002). This results in a description of the stochastic elements in the form of a scenario tree. The stochastic model for assets and liabilities is briefly described in Section 3.1 and Section 3.2 outlines the discretization methods that produce the scenario trees.

3.1 Modeling the stochastic factors

The formulas for calculating $R_{t,j}$ and $D_{t,j}$ for each asset class are displayed in Table 3, where sr , br , S , Div , P and $Rent$ denote the short term interest rate, long term bond yield, stock price index, dividend index, property price index and rental index, respectively, and τ_t denotes the length of the time period in years. The parameter D_M denotes the average duration of the company's bond portfolio.

Table 3: Return and dividend formulas

Asset class	$R_{t,j}$	$D_{t,j}$
Cash	$((1 + sr_t)(1 + sr_{t-1}))^{\frac{\tau_t}{2}}$	1
Bonds	$\left(\frac{1+br_{t-1}}{1+br_t}\right)^{D_M}$	$\frac{1}{2}(br_{t-1} + br_t)\tau_t$
Stocks	$\frac{S_t}{S_{t-1}}$	$\frac{1}{2}\left(\frac{Div_{t-1}}{S_{t-1}} + \frac{Div_t}{S_t}\right)\tau_t$
Property	$\frac{P_t}{P_{t-1}}$	$\left(\frac{1}{2}\left(\frac{Rent_{t-1}}{P_{t-1}} + \frac{Rent_t}{P_t}\right) - 0.03\right)\tau_t$
Loans	1	$\frac{1}{2}(br_{t-1} + br_t)\tau_t$

The return for cash investments is approximated by the geometric average of the short term interest rate during the holding period. The formula for bond returns is based on a duration approximation as in Koivu et al. (2003); see also (Campbell et al., 1997, Chapter 10). The parameter D_M equals five years. The dividends for stock and property investments present the average dividend and rental yield, respectively, during the holding period. For property investments the maintenance costs, which are assumed to be a constant 3% of the property value, are deducted from the rental yield. Similarly to bonds, the cash income for loans is approximated by an average of bond yield. This is based on the fact that the interest on newly given loans is usually set equal to the current bond yield. The return for loans is equal to one because these instruments are not traded in the market.

The Finnish earnings-related pension scheme follows the defined benefit principle, where the pension insurance company guarantees the pension payments which are tied to the development of the policyholder's salaries. It follows that, the technical reserves L and cash flows F depend on policyholder's wages and population dynamics. These are assumed independent, so that their development can be modeled separately. The values of L and F depend also on the technical interest rate, which determines the total growth rate for the reserves. In the model, the technical interest rate is calculated based on recent asset returns and it is an important part of the model because, to a great extent, it determines the correlations between the investment variables and the reserves. The development of wages is described by the general Finnish wage index W . For a more detailed description of the development of L and F , see Koivu et al. (2003).

The stochastic variables in the model can thus be approximately expressed in terms of the seven economic factors, sr , br , S , Div , P , $Rent$ and W . The quarterly development of

$$x_t = \begin{bmatrix} \ln sr_t \\ \ln br_t \\ \ln S_t \\ \ln Div_t \\ \ln P_t \\ \ln Rent_t \\ \ln W_t \end{bmatrix}$$

will be described with a Vector Equilibrium Correction (VEqC) model, popularized by Engle and Granger (1987) and Johansen (1995). During the last decade VEqC models have been widely used in modeling and forecasting economic and financial time series, see e.g. Campbell and Shiller (1987), Anderson et al. (2000), Clements and Hendry (1998) and Clements and Hendry (1999). We consider a VEqC model

$$\Delta_\delta x_t = \sum_{i=1}^k A_i \Delta_\delta x_{t-i} + \alpha(\beta' x_{t-1} - \mu) + \epsilon_t, \quad (2)$$

where $A_i \in \mathbb{R}^{7 \times 7}$, $\beta \in \mathbb{R}^{7 \times l}$, $\mu \in \mathbb{R}^l$, $\alpha \in \mathbb{R}^{7 \times l}$, Δ_δ denotes the shifted difference operator

$$\Delta_\delta x_t := \Delta x_t - \delta$$

with $\delta \in \mathbb{R}^7$, and ϵ_t are independent normally distributed random variables with zero mean and variance matrix $\Sigma \in \mathbb{R}^{7 \times 7}$. When the model is stationary the parameter vector δ determines the average drift for the time series. The term $\alpha(\beta' x_{t-1} - \mu)$ takes into account the long-term behavior of x_t around statistical equilibria described by the linear equations $\beta' x = \mu$. It is assumed that, in the long run,

$$E[\beta' x_t] = \mu,$$

and that if x_t deviates from the equilibria it will tend to move back to them. The matrix α determines the speed of adjustment toward the equilibria. In a sense, VEqC-models incorporate long-run equilibrium relationships (often derived from economic theory) with short-run dynamic characteristics deduced from historical data.

We take δ and μ as user specified parameters. This enables the incorporation of expert information in specifying the expected growth rates for x_t as well as long term equilibrium values for such quantities as mean reversion levels, interest rate spread and dividend yield. In particular, this gives control over mean returns which have been shown (in the context of the Markowitz model) to have a big impact on the optimal portfolio choice, see Chopra and Ziemba (1993). The appropriate lag-length k is determined and the remaining parameters are estimated from quarterly data from Finland and the EU-area. The estimated parameter values used in the numerical tests of Section 4 are given in the Appendix. For a more detailed description of the model; see Koivu et al. (2003).

3.2 Discretization

In our optimization model, we are interested in the conditional distributions of x_{t+h} , given x_t , typically for $h \geq 4$. This can be calculated conveniently as follows. After specifying the model (2),

we write it as a Vector Auto-Regressive (VAR) model in levels

$$x_t = (I + A_1 + \Gamma)x_{t-1} + \sum_{i=2}^k (A_i - A_{i-1})x_{t-i} - A_k x_{t-k-1} + c + \epsilon_t,$$

where $\Gamma = \alpha\beta'$ and $c = -\alpha\mu + (I - \sum_{i=1}^k A_i)\delta$. This, in turn, can be written in the companion form

$$\bar{x}_t = \bar{A}\bar{x}_{t-1} + \bar{c} + \bar{\epsilon}_t,$$

where

$$\bar{x}_t = \begin{bmatrix} x_t \\ x_{t-1} \\ \vdots \\ x_{t-k} \end{bmatrix}, \quad \bar{A} = \begin{bmatrix} I + A_1 + \Gamma & A_2 - A_1 & \cdots & A_k - A_{k-1} & -A_k \\ I & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{bmatrix},$$

$$\bar{c} = \begin{bmatrix} c \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \bar{\epsilon}_t = \begin{bmatrix} \epsilon_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

It follows that

$$\bar{x}_{t+h} = \bar{A}^h \bar{x}_t + \sum_{i=1}^h \bar{A}^{h-i} \bar{c} + e_h, \quad (3)$$

where $e_h = \sum_{i=1}^h \bar{A}^{h-i} \bar{\epsilon}_i$. The random term e_h is normally distributed with zero mean, and from the independence of $\bar{\epsilon}_i$ it follows that e_h has the variance matrix

$$\bar{\Sigma}_h = \sum_{i=1}^h \bar{A}^{h-i} \begin{bmatrix} \Sigma & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} (\bar{A}^T)^{h-i}.$$

A convenient feature of (3) is that the dimension of the random term never exceed $7(k+1)$ even if h is increased. In the model of Koivu et al. (2003), $k=1$, so the dimension will be at most 14.

We discretize the model (3) using integration quadratures as described in Pennanen and Koivu (2002). This results in scenario trees that converge weakly to the original process as the number of branches is increased. This technique is just as easy to implement as the better known method of conditional sampling. Indeed, a scenario tree with a given period structure $(\tau_1, \dots, \tau_T)$ and branching structure $(\nu_1, \dots, \nu_T)$ can be generated as follows. For each $t = 0, \dots, T$, denote by $\mathcal{N}_t$ the set of nodes in the scenario tree at stage t . The set $\mathcal{N}_0$ consists only of the root node which is labeled by 0. The rest of the nodes will be labeled by positive integers in the order they are generated. The number $h_t = 4\tau_t$ gives the length of period $[t-1, t]$ in quarters.

Set $m := 0$, $\bar{x}_m :=$ the current state of the world, and $\mathcal{N}_0 := \{m\}$.

```

for  $t := 1$  to  $T$ 
   $\mathcal{N}_t := \emptyset$ 
  for  $n \in \mathcal{N}_{t-1}$ 
    Draw a random sample of  $\nu_t$  points  $\{e_{h_t}^i\}_{i=1}^{\nu_t}$  from  $N(0, \bar{\Sigma}_{h_t})$ 
    for  $i := 1$  to  $\nu_t$ 
       $m := m + 1$ 
       $\bar{x}_m = \sum_{i=1}^h \bar{A}^{h-i} \bar{c} + \bar{A}^h \bar{x}_n + e_{h_t}^i$ 
       $\mathcal{N}_t := \mathcal{N}_t \cup \{m\}$ 
    end
  end
end

```

The random samples required above are easily generated by computing the spectral decomposition

$$\bar{\Sigma}_h = \sum_{i=1}^{7(k+1)} \lambda_h^i u_h^i (u_h^i)^T,$$

where λ_h^i are the eigenvalues of $\bar{\Sigma}_{h_t}$ in decreasing order and u_h^i are the corresponding eigenvectors. If $\bar{\Sigma}_{h_t}$ has rank d_t ,

$$\bar{\Sigma}_h = C_h C_h^T,$$

where $C_h = [\sqrt{\lambda_h^1} u_h^1, \dots, \sqrt{\lambda_h^{d_t}} u_h^{d_t}]$, and the desired sample is obtained as

$$e_{h_t}^i := C_h F_{d_t}^{-1}(u_{h_t}^i),$$

where $\{u_{h_t}^i\}_{i=1}^{\nu_t}$ is a random sample from U_{d_t} , the d_t -dimensional uniform distribution on $[0, 1]^{d_t}$ and F_{d_t} is the distribution function of the d_t -dimensional standard normal distribution. An advantage of computing the spectral decomposition (instead of the Cholesky decomposition as e.g. in Høyland et al. (2003)) is that when $\bar{\Sigma}_{h_t}$ is singular, d_t gives the true dimension of the random term. For example, when $h = 1$, $d_t = 7$.

The random samples $\{u_{h_t}^i\}_{i=1}^{\nu_t}$ above can be viewed as discrete approximations of U_{d_t} . As in Pennanen and Koivu (2002), we will replace these random samples by low discrepancy point-sets that have been designed to give good approximations of U_{d_t} . In the numerical tests in the next section we will use point-sets from the Sobol sequence; see for example Jäkel (2002). This produces a scenario tree with the same branching structure as the above conditional sampling procedure but a potentially better approximation of the original stochastic process, because the low discrepancy points are constructed to be more evenly distributed over U_{d_t} than typical random points. The computation times with Sobol sequences is roughly equal to that with Monte Carlo. See Pennanen and Koivu (2002) for a numerical study of such scenario trees.

4 Numerical results

4.1 Implementation

Figure 1 sketches the structure of the overall optimization system. The scenario generator (written in the C programming language) takes as inputs the period and branching structures of the scenario

tree and the time series model for the stochastic factors and generates the scenario tree for the assets and liabilities. The tree can be visually and otherwise inspected e.g. in spreadsheet programs until the outcomes are satisfactory. The scenario tree is then written into a text file in AMPL format described in Fourer et al. (2002). The optimization model written in AMPL modeling language and the data from the scenario generator are processed in AMPL and fed to MOSEK MOSEK which is an interior-point solver for convex (nonlinear) programs. The solution details and statistics produced by AMPL/MOSEK can again be visualized e.g. in spreadsheet programs. The system can be used under most Unix and Windows platforms.

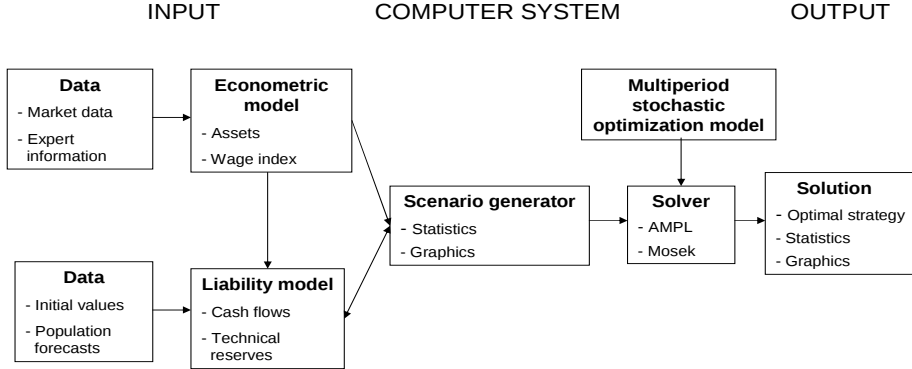


Figure 1: Stochastic optimization system

4.2 Computational experiments

We chose the beginning of year 2002 as the first stage $t = 0$ in our experiments. The initial values for the time series model and the model parameters

$$h_0 = (1563, 622, 5573, 3914, 2158)$$

and $H_{-1} = 151.341$ (million euros) were chosen accordingly. As an example, we generated a scenario tree with period structure (1, 3, 6) years and branching structure (25, 10, 10) (2500 scenarios). This takes less than a second on Intel Pentium 4, 2.33GHz, with 1Gb of SDRam. Figure 2 plots the values of some important parameters on the scenario tree. We solved the corresponding stochastic programming model for five sets of shortfall penalty coefficients given in Table 4.

Table 4: Shortfall penalty coefficients in the example

	γ_1	γ_2	γ_3
SP 1	1	10	10
SP 2	0.5	10	10
SP 3	1	1	1
SP 4	0.1	10	10
SP 5	0	0	0

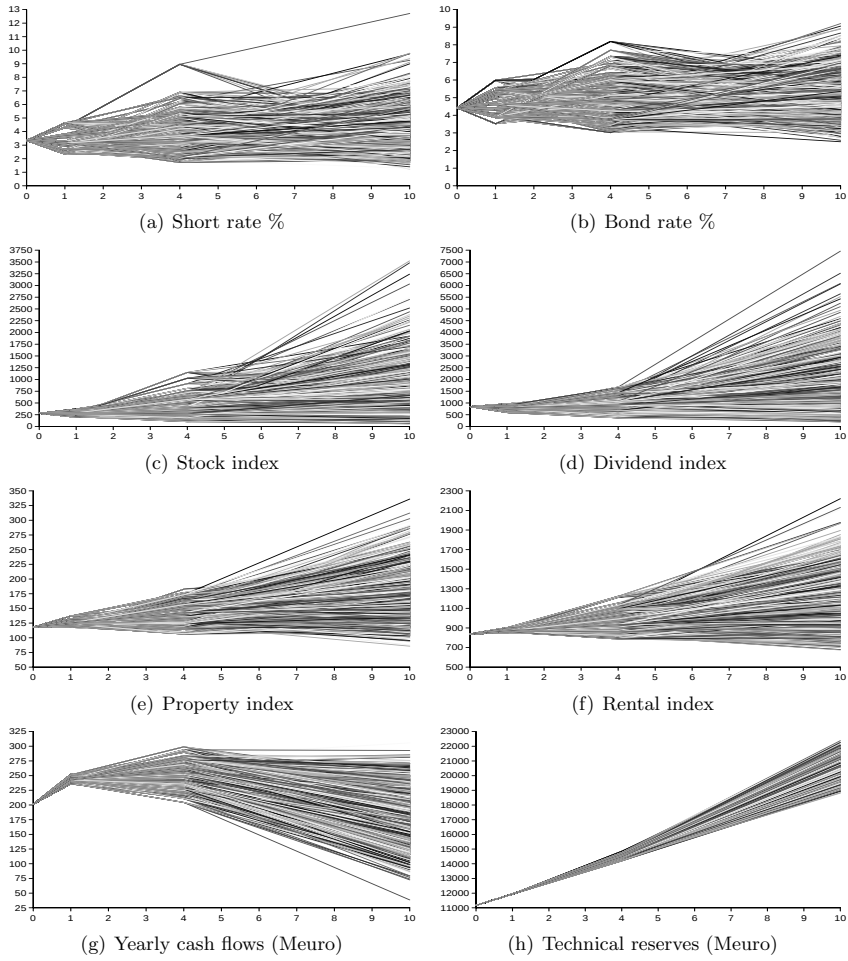


Figure 2: Scenario tree of the example.

These (somewhat arbitrarily chosen) values correspond to different attitudes towards the attainment of the various target zones described in Subsection 2.3. In all cases we used the piecewise linear utility function

$$u^b(\cdot) = 1.5\gamma_2 \min\{\cdot, 0.01\}$$

for bonuses. The solution of the corresponding optimization models takes less than 10 seconds each. Figure 3 displays the optimal portfolio weights in stage $t = 0$ for the five sets of parameter choices. The first column gives the actual portfolio of the company in the beginning of year 2002. One can also examine the development of the optimized decision variables along the scenario tree.

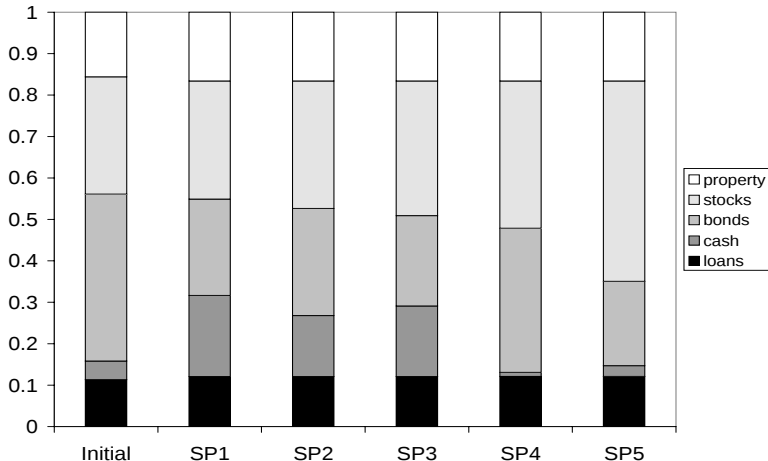


Figure 3: Initial portfolio h^0 and the optimal portfolios corresponding to the parameter values in Table 4.

Figures 4(a) and 4(b) plot the optimized C_t/L_t and H_t/L_t ratios, respectively, for SP1 of Table 4. The solvency capital C_t is always nonnegative (no bankruptcy) in every scenario while the bonus transfer/liability ratio H_t/L_t is equal to 0.01 in almost every scenario.

To gain some insight on the effect of the the shortfall penalties associated with the target zones, we solved the optimization model SP1 for varying levels of initial wealth w_0 . This was done by rescaling the initial portfolio so that the relative portfolio weights remained unchanged. The model was resolved and the optimized first stage portfolio recorded. The resulting portfolios are graphed in Figure 5(a) as a function of the ratio w_0/L_0 . For comparison, we did the same for SP5, where there is no penalty for the shortfalls; see Figure 5(b). The original wealth-liability ratio at the beginning of year 2002 was $w_0/L_0 = 1.238$.

Compared to SP5, the optimal portfolios for SP1 have considerably more wealth allocated to the short interest rate and bonds when $w_0/L_0 \leq 1.5$. This is natural since putting more wealth to the “safer” instruments reduces the solvency border and also the shortfalls. When w_0/L_0 approaches 2, the portfolios begin to look alike. This is caused by the fact that for high levels of initial wealth the probability of a shortfall is reduced and the effect of penalties becomes negligible. The most interesting phenomenon is that when the company approaches bankruptcy ($w_0/L_0 < 1$), it moves

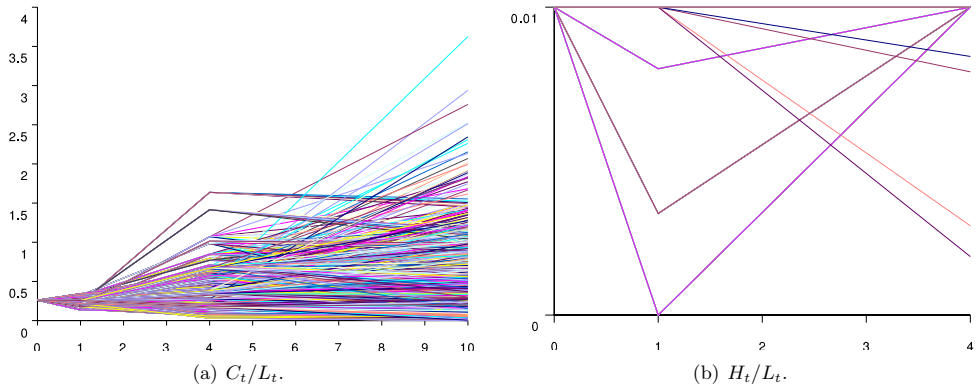


Figure 4: Optimized solvency capital and bonus ratios along the scenario tree for SP1.

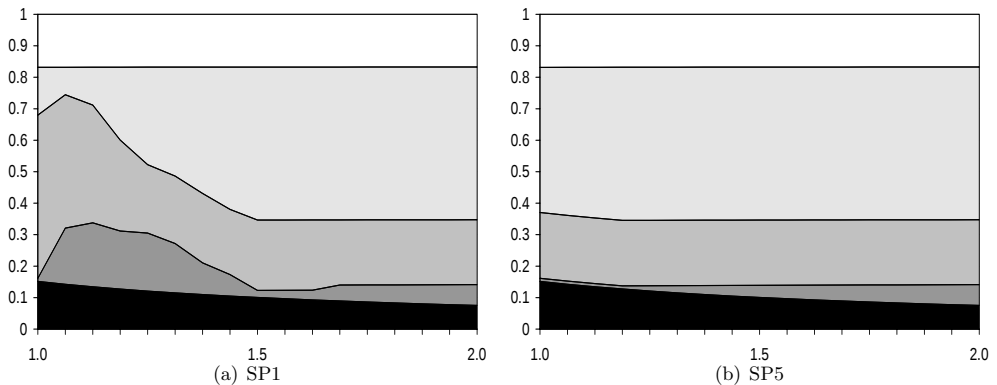


Figure 5: The optimal first stage portfolio as a function of w_0/L_0 (The legend of Figure 3 applies).

wealth from short interest rate to bonds and stocks, even though this results in higher solvency border and higher shortfall penalties for the first two zones. This is probably due to the fact that the company is anticipating the solvency situation in later periods and trying to make safe portfolio allocations by ignoring to some extent the recommendations embodied in the definition of the solvency border.

4.3 Convergence of discretizations

Being forced to approximate the continuous distribution of the uncertain parameters by finite distributions, it is natural to ask how the corresponding optimization problems depend on the number of scenarios. A simple test is to study the behavior of the optimal values as the number of

scenarios is increased. We will do the test for SP1 of Table 4 using the Sobol sequence as described in Section 3.2. For simplicity, we only considered fully symmetric scenario trees where each node has an equal number of branches, i.e. branching structure is (k, k, k) for $k = 1, 2, 3 \dots$. The solid line in Figure 6 plots the objective value as a function of the size of the scenario tree. For low

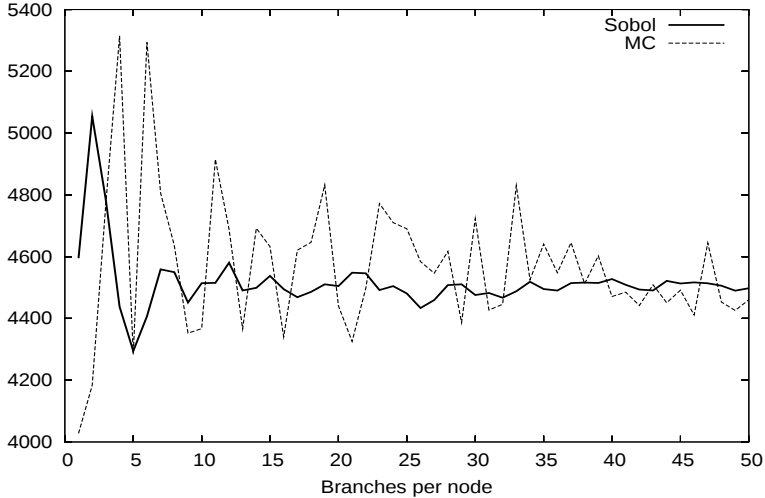


Figure 6: Convergence of the optimal value

values of k , the optimal value goes through large variations, but as k is increased the optimal value seems to stabilize. In fact, it stabilizes close to 4504 which is what we obtained with the branching structure $(25, 10, 10)$ in the above example. Convergence of discretizations of multistage stochastic programs has been studied for example by Olsen (1976) and Casey and Sen (2003), but at present, there do not exist analytical results that would explain the convergence of the optimal values in the present case.

For comparison, we did the same test using Monte Carlo sampling in generating the scenario trees. This resulted in the dotted line in Figure 6. The optimal values obtained with Monte Carlo seem to converge too but not nearly as fast as the optimal values obtained with the Sobol sequence.

4.4 Out-of-sample test

We implemented an out-of-sample testing procedure to evaluate the performance of our stochastic programming model. Optimized strategies corresponding to the five sets of shortfall penalty coefficients in Table 4 were compared with a variety of static fixed-mix and dynamic portfolio insurance (PI) strategies meeting the statutory restrictions of Table 1. The fixed-mix portfolio weights were chosen according to a grid in order to evenly cover the region of feasible portfolios.

In the PI strategies the portfolio weights for cash π_c and property π_p are varied according to the same rules as in the fixed-mix case. The rest of the wealth is divided between bonds and stocks

and the proportion of stocks in the portfolio at time t is given by,

$$\pi_{s,t} = \begin{cases} \min \left\{ (1 - \pi_c - \pi_p) \min \left\{ \rho \frac{C_t}{w_t}, 1 \right\}, 0.5 \right\} & \text{if } C_t \geq 0, \\ 0 & \text{if } C_t < 0, \end{cases}$$

where ρ is a risk tolerance parameter indicating how the proportion invested in stocks increases with the company's solvency ratio, C_t/w_t . The percentage invested in stocks is a constant multiple of the company's solvency ratio, which was close to 22% initially, with higher values of ρ resulting in higher stock market allocations. When the company's solvency capital is negative the stock market allocation is set to zero and the remaining wealth is invested in bonds. In general, PI strategies are suitable decision rules for pension insurance companies because they allocate more wealth to risky assets, stocks, when the companies' solvency ratios improve and reduce the stock market exposure as the companies approach insolvency.

As pointed out in the introduction, fixed-mix and PI strategies should not be considered as fully realistic decision rules. Rather, we view them as the first benchmarks that any practical decision support system should be able to outperform. Note however, that with these decision strategies, there is no guarantee that the transaction constraints will be satisfied. To simplify the comparison of different strategies, bonus transfers H_t were set to zero in each model. In addition, transaction costs were ignored in the case of fixed-mix and PI strategies to simplify computations. Note that this causes a bias in favor of the fixed-mix and PI strategies. The scenario trees used in optimization had the same structure as in the example of Section 4.1, that is, period structure (1, 3, 6) years and branching structure (25, 10, 10).

In the test, we evaluated the performance of each strategy over 325 randomly simulated scenarios of the stochastic parameters over 20 years. Portfolio rebalancing was made every year, i.e. fixed-mix portfolios are rebalanced to fixed proportions, PI portfolios are rebalanced and stochastic programming problems were solved with a new scenario tree, based on the current values of the stochastic parameters along the simulated scenario. We considered PI strategies with $\rho \in \{0.5, 1, \dots, 20\}$. The following describes the testing procedure. As outlined in Section 3, the stochastic factors in each year can be expressed in terms of a 14-dimensional vector. Below, $\bar{x}_{s,y}$ denotes the value of this vector in year y along a randomly generated scenario s .

for $s := 1$ **to** 325

Set $\bar{x}_{s,0} = \bar{x}_0$ (the current state of the world).

for $y := 0$ **to** 19

Generate a scenario tree rooted at $\bar{x}_{s,y}$.

Solve the corresponding optimization problems and rebalance all the portfolios.

Randomly sample $\bar{x}_{s,y+1}$ from the time series model and calculate the resulting portfolios and cash-flows at time $y + 1$.

end

end

Figure 7 plots the performance of all the fixed-mix and PI strategies and the 5 stochastic programming strategies with respect to the average solvency capital at the end of the simulation period versus the bankruptcy probability during the period. Considering the main risk of the company, bankruptcy, and average solvency capital, the stochastic programming strategies clearly dominate both the fixed-mix and PI strategies, even though the probability of bankruptcy was

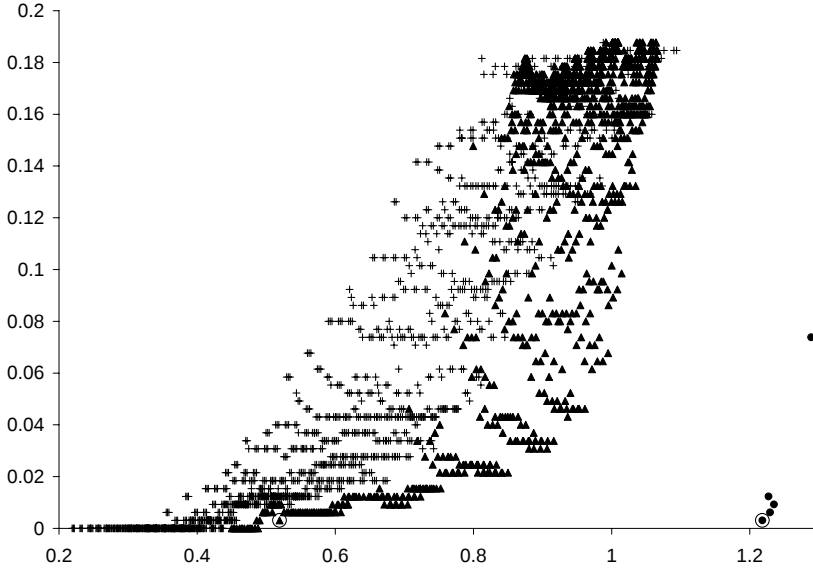


Figure 7: C_T/L_T vs. bankruptcy probability for fixed-mix (+), PI (▲) and stochastic programming (●) strategies.

not explicitly minimized. It is also worth noting that the best PI strategies outperform the best fixed-mix strategies at all reasonable risk levels. The riskiest stochastic programming strategy, SP5 of Table 4, went bankrupt in 25 simulations out of the 325 and the safest, SP1, in only one.

We will compare SP1 more closely with the best performing PI strategy circled in Figure 7, having the same bankruptcy probability as SP1. In the selected PI strategy $\pi_c = 0.04$, $\pi_p = 0.15$ and $\rho = 1$. The development of the solvency capital-reserves ratio for both strategies is described in Figure 8. The three lines represent the development of the sample average and the 95% confidence interval computed from the 325 scenarios. A higher mean and upwards skewed distribution indicates that the stochastic programming model can hedge against risks without losing profitability. Figure 9 shows the distribution of the solvency capital-solvency border ratio C_t/B_t at the beginning of the second year. Due to the aim for high investment returns, the stochastic programming strategy avoids unnecessarily high levels of C_t/B_t , and consequently, it hits the lower border of the target zones frequently.

Figure 10 displays the development of the distribution of C_t/B_t in the 325 scenarios over the four zones defined in Subsection 2.3. If we compare the two strategies according to the target zones, the PI strategy seems to perform better than SP1. However, in the long run the stochastic programming strategy produces superior returns compared to the PI strategy, without increasing the company's bankruptcy risk.

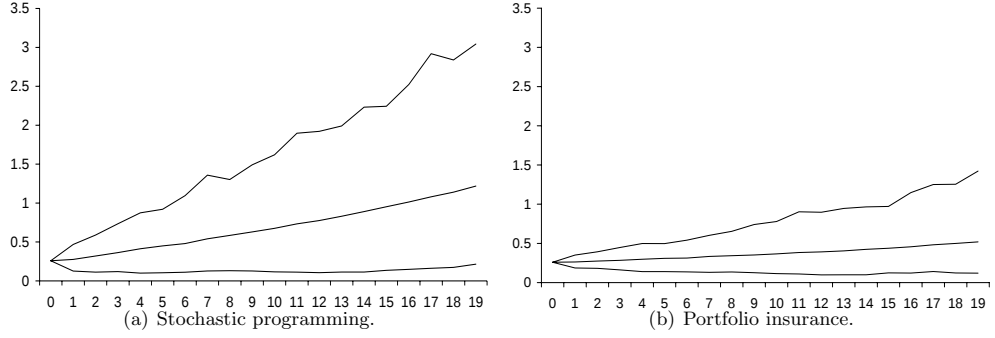


Figure 8: C_t/L_t averages and 95% confidence intervals.

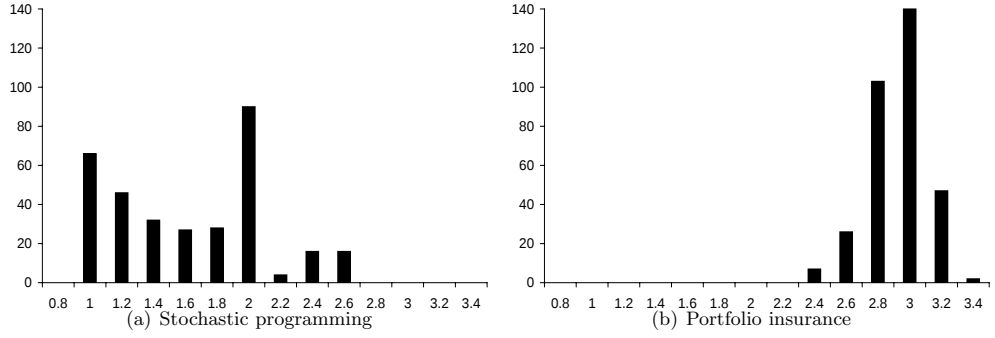


Figure 9: Distribution of C_2/B_2 at the beginning of the second period

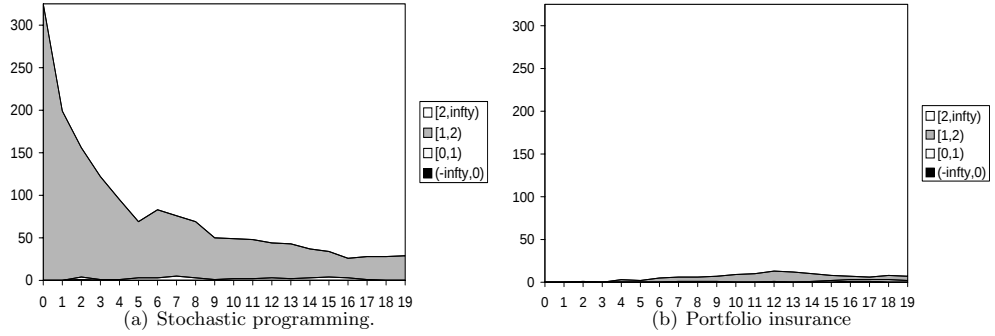


Figure 10: Development of the distribution of C_t/B_t over the different zones

Appendix

The parameters for the time series model described in Subsection 3.1 were estimated using full information maximum likelihood and are as follows. The number of lags $k = 1$,

$$\delta = \begin{bmatrix} 0 \\ 0 \\ 0.0114 \\ 0.0114 \\ 0.007 \\ 0.007 \\ 0.009 \end{bmatrix}, \quad \mu = \ln \begin{bmatrix} 3.7 \\ 1.2 \\ 2.5 \\ 7.0 \end{bmatrix},$$

$$A_1 = 10^{-1} \begin{bmatrix} 3.672 & 3.467 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2.855 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -59.11 \\ 0 & 0 & -2.425 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.629 & 0 & 3.617 & 0 & 0 \\ 0 & -0.209 & 0 & 0 & -0.663 & 8.533 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0.638 & 8.712 \end{bmatrix},$$

$$\alpha = 10^{-1} \begin{bmatrix} 0 & 0.964 & 0 & 0 \\ -1.061 & -1.499 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1.449 & 0 \\ -0.238 & 0 & 0 & 0.637 \\ 0 & 0.080 & 0 & 0 \\ 0 & 0 & -0.024 & 0 \end{bmatrix}, \quad \beta = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\Sigma = 10^{-4} \begin{bmatrix} 53.7113 & 7.155 & -4.7954 & -15.978 & 0.1119 & 0.2726 \\ 0.13079 & 55.719 & 10.741 & -11.647 & 0.2278 & 0.4652 \\ -0.0606 & 0.1332 & 116.73 & 45.187 & 6.3447 & -0.5866 \\ -0.2764 & -0.1978 & 0.5302 & 62.235 & -1.0418 & -0.2177 \\ 0.0092 & 0.0184 & 0.354 & -0.0796 & 2.7519 & -0.0069 \\ 0.09107 & 0.1526 & -0.1329 & -0.0675 & -0.0102 & 0.1668 \\ -0.174 & -0.1929 & 0.0641 & 0.3586 & 0.0113 & -0.2396 \end{bmatrix}.$$

The initial values for the time series at the beginning of year 2002 were

$$x_0 = \ln \begin{bmatrix} 3.35 \\ 4.42 \\ 279.6 \\ 843.7 \\ 118.0 \\ 839.8 \\ 140.6 \end{bmatrix}, \quad x_{-1} = \ln \begin{bmatrix} 4.16 \\ 4.33 \\ 242.9 \\ 776.0 \\ 117.7 \\ 831.3 \\ 139.1 \end{bmatrix}.$$

References

- R. G. Anderson, D. Hoffman, and R. H. Rasche. A vector error-correction forecasting model of the U.S. economy. Working paper 1998-008b, current version July 2000, Federal Reserve Bank of St. Louis, 2000.
- F. Black and R. Jones. Simplifying portfolio insurance for corporate pension plans. *Journal of Portfolio Management*, 14:33–37, 1988.
- J. Y. Campbell, A. W. Lo, and A. C. MacKinley. *The Econometrics of Financial Markets*. Princeton University Press, USA, 1997.
- J. Y. Campbell and R. Shiller. Cointegration and tests of present value models. *Journal of Political Economy*, 95(5):1062–1088, 1987.
- D. R. Cariño, D. H. Myers, and W. T. Ziemba. Concepts, technical issues, and uses of the Russell-Yasuda Kasai financial planning model. *Operations Research*, 46(4):450–462, 1998.
- D. R. Cariño and W. T. Ziemba. Formulation of the Russell-Yasuda Kasai financial planning model. *Operations Research*, 46(4):433–449, 1998.
- M. Casey and S. Sen. The scenario generation algorithm for multistage stochastic linear programs. Working paper, University of Arizona, 2003.
- V. K. Chopra and W. T. Ziemba. The effect of errors in means, variances, and covariances on optimal portfolio choice. *Journal of Portfolio Management*, 19(2):6–11, 1993.
- M. P. Clements and D. F. Hendry. *Forecasting economic time series*. Cambridge University Press, Cambridge, 1998.
- M. P. Clements and D. F. Hendry. *Forecasting non-stationary economic time series*. MIT Press, Cambridge, MA, 1999. With a foreword by Katarina Juselius.
- G. Consigli and M. A. H. Dempster. Dynamic stochastic programming for asset-liability management. *Ann. Oper. Res.*, 81:131–161, 1998.
- T. Dardis and H. Mueller. Reengineering ALM in North America. *Emphasis*, (1):22–25, 2001.
- R. F. Engle and C. W. J. Granger. Co-integration and error correction: representation, estimation, and testing. *Econometrica*, 55(2):251–276, 1987.
- S. E. Fleten, K. Høyland, and S. W. Wallace. The performance of stochastic dynamic and fixed mix portfolio models. *European J. Oper. Res.*, 140(1):37–49, 2002.
- H. Föllmer and A. Schied. *Stochastic finance, An introduction in discrete time*, volume 27 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 2002.
- R. Fourer, D. M. Gay, and B. W. Kernighan. *AMPL: A Modeling Language for Mathematical Programming*. Duxbury Press, 2nd edition, 2002.

- A. Geyer, W. Herold, K. Kontriner, and W. T. Ziemba. The Innovest Austrian pension fund financial planning model InnoALM. Working paper, University of British Columbia, 2003.
- K. Høyland. Asset liability management for a life insurance company: A stochastic programming approach. Ph.D. Thesis, Norwegian University of Science and Technology, 1998.
- K. Høyland, M. Kaut, and S. Wallace. A heuristic for moment-matching scenario generation. *Computational Optimization and Applications*, 24:169–186, 2003.
- P. Jäckel. *Monte Carlo Methods in Finance*. John Wiley & Sons, 2002.
- S. Johansen. *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford University Press, New York, 1995.
- M. Koivu, T. Pennanen, and A. Ranne. Modeling assets and liabilities of a Finnish pension company: a VEqC approach. Working paper, Helsinki School of Economics, 2003. URL www.hkkk.fi/~pennanen.
- R. Kouwenberg. Scenario generation and stochastic programming models for asset liability management. *European J. Oper. Res.*, 134(2):279–292, 2001.
- MOSEK. www.mosek.com.
- S. S. Nielsen and S. A. Zenios. A stochastic programming model for funding single premium deferred annuities. *Math. Programming*, 75(2, Ser. B):177–200, 1996.
- P. Olsen. Discretizations of multistage stochastic programming problems. *Math. Programming Stud.*, (6):111–124, 1976.
- T. Pennanen and M. Koivu. Integration quadratures in discretization of stochastic programs. *Stochastic Programming E-Print Series*, (11), 2002.
- A. F. Perold and W. F. Sharpe. Dynamic strategies for asset allocation. *Financial Analysts Journal*, 44:16–27, 1988.
- S. A. Zenios and W. T. Ziemba, editors. *Handbook of asset liability management*. Handbooks in Finance. North Holland, 2004.
- W. T. Ziemba and J. M. Mulvey, editors. *Worldwide Asset and Liability Management*, volume 10 of *Publications of the Newton Institute*. Cambridge University Press, Cambridge, 1998.

Epi-convergent discretizations of stochastic programs via integration quadratures

Teemu Pennanen* and Matti Koivu
Department of Management Science
Helsinki School of Economics
PL 1210, 00101 Helsinki, Finland
[pennanen,koivu]@hkkk.fi

Abstract

The simplest and the best-known method for numerical approximation of high-dimensional integrals is the Monte Carlo method (MC), i.e. random sampling. MC has also become the most popular method for constructing numerically solvable approximations of stochastic programs. However, certain modern integration quadratures are often superior to crude MC in high-dimensional integration, so it seems natural to use them also in discretization of stochastic programs. This paper derives conditions that guarantee the epi-convergence of the resulting objectives to the original one. Our epi-convergence result is closely related to some of the existing ones but it is easier to apply to discretizations and it allows the feasible set to depend on the probability measure. As examples, we prove epi-convergence of quadrature-based discretizations of three different models of portfolio management and study their behavior numerically. Besides MC, our discretizations are the only existing ones with guaranteed epi-convergence for these problem classes. In our tests, modern quadratures seem to result in faster convergence of optimal values than MC.

1 Introduction

Let X and Ξ be complete separable metric spaces, and Σ the Borel σ -algebra on Ξ . Let P be a probability measure on (Ξ, Σ) , and f an extended real-valued function on $X \times \Xi$, such that $f(x, \cdot)$ measurable for every $x \in X$. This paper studies numerical solution through discretization of stochastic programs (optimization problems) of the form

$$\underset{x \in X}{\text{minimize}} \quad E^P f(x) := \int_{\Xi} f(x, \xi) P(d\xi), \quad (SP)$$

where the integral is interpreted as $+\infty$ when $f(x, \cdot) \notin L^1(\Xi, \Sigma, P)$. Here the decision variable x is not a function of ξ , so (SP) represents a *static* (one-stage) stochastic program. By allowing f to take on the value $+\infty$ we can incorporate constraints into the objective, which makes (SP) a

*The work of this author was partially supported by The Finnish Foundation for Economic Education under grant number 21599

very general model for static decision making problems under uncertainty. Unlike most studies of stochastic programs, we do not assume the feasible set

$$\text{dom } E^P f(x) = \{x \in X \mid f(x, \cdot) \in L^1(\Xi, \Sigma, P)\}$$

to be known a priori. This is essential e.g. in stochastic programs without relatively complete recourse and in certain financial applications, where the determination of the feasible set is part of the problem rather than its statement; see Subsection 4.3.

In solving problems of the above form, a common approach is to replace P by a finitely supported measure of the form

$$P^\nu = \sum_{i=1}^{\nu} p_i^\nu \delta_{\xi_i^\nu},$$

where $\delta_{\xi_i^\nu}$ denotes the unit mass located at ξ_i^ν . This results in the problem

$$\underset{x \in X}{\text{minimize}} \quad E^{P^\nu} f(x) := \sum_{i=1}^{\nu} p_i^\nu f(x, \xi_i^\nu), \quad (SP^\nu)$$

which, for moderate values of ν , is usually easier to solve than (SP) . Several approaches for constructing the measures P^ν have been considered in the literature. In general, the aim is to choose P^ν so that (SP^ν) is a good approximation of (SP) and that the number ν of support points of P^ν is small enough to allow for numerical solution of (SP^ν) . Note that, since f is extended real-valued and since the containment $\text{dom } E^P f \subset X$ may be strict, it may happen that $\text{dom } E^{P^\nu} f \neq \text{dom } E^P f$.

The simplest and perhaps the most popular choice in applications is to use empirical measures of the form $P^\nu = \sum_{i=1}^{\nu} \frac{1}{\nu} \delta_{\xi_i^\nu}$, where $\{\xi_i^\nu\}_{i=1}^{\nu}$ is a random sample from P . Such random approximations of stochastic programs are known to be consistent as the sample size increases; see for example Artstein and Wets (1995) for asymptotic analysis under very mild conditions and Shapiro (2000) for more special results in the case where f is real-valued and the feasible set is known a priori ($\text{dom } E^P f = X$). However, a random sample can lead to a bad approximation of P , which in turn, may lead to an equally bad approximation of the optimization problem. There have been attempts to improve the accuracy of crude Monte Carlo sampling by using ideas from importance sampling technique; see Infanger (1992) and Dempster and Thompson (1999). Høyland and Wallace (2001) use moment matching where P^ν is constructed so that it has the first few moments of the original distribution; see also Høyland et al. (2003). In barycentric approximation, one constructs P^ν so that, under certain convexity properties of the function $f(\cdot, \cdot)$, the optimum value of (SP^ν) provides an upper/lower bound to that of (SP) ; see Frauendorfer (1992). Pflug (2001) proposed to construct discrete measures P^ν so that they are as close as possible to P in the sense of the so called Wasserstein-distance

This paper studies the use of modern integration quadratures in constructing the discretizations (SP^ν) . Such quadratures have the attractive feature that they have been designed to give discrete measures that approximate a given measure as well as possible. Moreover, they are just as easy to use as crude Monte Carlo and they are fast compared to methods like barycentric approximation, moment matching or that in Pflug (2001). We study the corresponding discretizations both analytically and numerically. The use of integration quadratures in solving stochastic programs have been considered in Deák (1988), Lepp (1990) and in Pennanen and Koivu (2002), but in Deák (1988),

integration quadratures were not considered as feasible methods for constructing approximations (SP^ν), and in Pennanen and Koivu (2002) no convergence analysis were given.

Since we are dealing with minimization problems, a natural framework for analyzing approximations is *epi-convergence*; see Attouch (1984) or Rockafellar and Wets (1998) for introduction to epi-convergence. Epi-convergence of the objectives is a minimal property that should be satisfied by any approximation scheme for optimization problems in order to get asymptotic convergence of optimal values and solutions. Epi-convergence of stochastic programs with respect to perturbations in the probability measure has been studied, for example, by Birge and Wets (1986), Robinson and Wets (1987), Dupačová and Wets (1988), Kall et al. (1988), Lucchetti and Wets (1993), Artstein and Wets (1994), Zervos (1999), Schultz (2000) and Vogel and Lachout (2003). In these studies, *weak convergence* of the approximating measures P^ν to the original measure P has been found an important property. In numerical integration, weak convergence corresponds to *consistency* which is a minimal requirement for any integration quadrature. We derive an epi-convergence result which is closely related to the ones in the above references but it is easier to apply to discretizations and it does not require the feasible set to be independent of the measure. As examples, we discretize three different models of portfolio optimization with integration quadratures and we verify the epi-convergence of the resulting approximations. The earlier epi-convergence results do not seem applicable in these instances. In particular, in the third example, the feasible set depends on the measure in an essential way.

When using empirical measures instead of integration quadratures in approximating our example problems, one gets *almost sure* epi-convergence from the general result of Artstein and Wets (1995). In numerical tests, integration quadratures seem to result in faster convergence than empirical measures thus allowing for smaller values of ν and cheaper computations. A rigorous quantitative analysis of such phenomena would require an epigraphical analysis along the lines of Attouch and Wets (1991). Unfortunately, quantitative results for approximations of optimization problems often rely on strong convexity properties which are missing from many important problems in practice, and in particular, in two of our test problems. Some quantitative results for approximations of stochastic programs can be found in Römisch and Schultz (1993), Shapiro (1994), Rachev and Römisch (2002). Our focus here is rather on deriving as weak conditions as possible that will allow us to deduce asymptotic epi-convergence of discretizations of as general models as possible. Even such asymptotic epi-convergence results for discretizations of stochastic programs are nontrivial, which can be seen from the lack of them for most discretization methods besides empirical approximations. This paper presents a simple deterministic approximation scheme with a theoretical justification for a rather general class of practically interesting problems.

The rest of this paper is organized as follows. Section 2 gives a brief review of modern integration quadratures and their use in generation of weakly convergent probability measures. In Section 3, we derive an epi-convergence result for $E^{P^\nu}f$. In Section 4 we combine results from Sections 2 and 3 to construct epi-convergent discretizations of some particular problems, and we study the stability of the corresponding optimal values numerically.

2 Constructing weakly convergent probability measures

Based on the importance of weak convergence in studying epi-convergence of stochastic programs, it is natural to try to choose the measures P^ν in (SP^ν) so that they converge weakly to P as $\nu \nearrow \infty$.

Weak convergence, denoted by $P^\nu \rightarrow P$, means that

$$E^{P^\nu} \varphi \rightarrow E^P \varphi, \quad (1)$$

for all bounded and continuous functions φ ; see Billingsley (1999). The literature of numerical integration contains many methods for generating such sequences. These methods often perform much better in numerical integration than crude Monte Carlo. For very low-dimensional integrals, Gaussian quadratures are usually most effective, but in higher dimensions, low discrepancy sequences and point sets (quasi Monte Carlo methods) often give better results. Both classes of methods are briefly reviewed below.

2.1 Low-dimensional spaces: Gaussian quadratures

Gaussian quadratures are usually very efficient in one-dimensional integration Press et al. (1992). For different choices of integration limits a and b and a weight (density) function w , they yield approximations

$$\int_a^b \varphi(\eta) w(\eta) d\eta \approx \sum_{i=1}^\nu w_i^\nu \varphi(\eta_i^\nu), \quad (2)$$

where the quadrature points η_i^ν and weights w_i^ν are chosen so that the quadrature has an optimal order of accuracy: *a ν point Gaussian quadrature is exact for all polynomials of degree $2\nu - 1$ or less*. Given a , b and w , the values of η_i and w_i can be computed numerically. For certain choices of a , b and w , the computation of the quadrature points and weights is particularly easy.

For $a = -\infty$, $b = \infty$, $w(\eta) = \exp(-\eta^2)$, (2) is known as *Gauss-Hermite quadrature*. See Press et al. (1992) for a *C*-routine for computing the points η_i^ν and the weights w_i^ν of the Gauss-Hermite quadrature. Gauss-Hermite quadrature can be used to approximate the expectation under the normal distribution P as

$$E^P \varphi = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{\xi^2}{2}} \varphi(\xi) d\xi = \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} e^{-\eta^2} \varphi(\sqrt{2}\eta) d\eta \approx \sum_{i=1}^\nu \frac{1}{\sqrt{\pi}} w_i^\nu \varphi(\sqrt{2}\eta_i^\nu).$$

It turns out that $w_i^\nu > 0$ and $\sum_{i=1}^\nu \frac{w_i}{\sqrt{\pi}} = 1$, the latter following from the fact that the quadrature is exact for constant functions. Thus,

$$P^\nu := \sum_{i=1}^\nu \frac{w_i^\nu}{\sqrt{\pi}} \delta_{\sqrt{2}\eta_i^\nu}$$

defines a probability measure which satisfies $E^{P^\nu} \varphi = E^P \varphi$ for all polynomials φ of order $2\nu - 1$ or less. (In particular, P^ν matches the first $2\nu - 1$ moments of the normal distribution; compare with Høyland et al. (2003)). The Weierstrass approximation theorem then implies that the measures P^ν converge weakly to P as $\nu \nearrow \infty$.

Gaussian quadratures do not directly extend to higher dimensions. The generation of integration quadratures of optimal order in arbitrary dimension is an open problem Press et al. (1992). If a random variable can be expressed as $\xi = G(\zeta)$, where G is almost everywhere continuous, and $\zeta = (\zeta_1, \dots, \zeta_d)$ for independent random variables ζ_i with densities w_i , we can approximate the distribution of each ζ_i with a Gaussian quadrature to get discrete measures Q_i^ν , and then construct the measure

$$P^\nu = (Q_1^\nu, \dots, Q_d^\nu) G^{-1}; \quad (3)$$

see Theorem 1 below. Such approximations work well in dimensions 1, 2 or 3, but in higher dimensions they suffer from the “curse of dimensionality”: if one approximates each Q_i with an k -point quadrature, a 10-dimensional approximation of the above form would have k^{10} quadrature points. Fortunately, there are better integration quadratures for high-dimensional spaces.

2.2 Higher dimensions: low discrepancy point sets and the method of inversion

In the scalar case, it is typical to approximate the uniform distribution on $[0, 1]$ and to transform each point with the inverse of the distribution function of the desired distribution. This is known as the *method of inversion*. The same idea works whenever $P = QG^{-1}$, where Q is the multivariate uniform distribution and G is Q -a.s. continuous, in other words, whenever

$$\xi = G(u),$$

where u is uniformly distributed in the unit cube $[0, 1]^d$, and $G : [0, 1]^d \rightarrow \Xi$ is almost everywhere continuous. This is based on the following very useful result from Billingsley (1999) where U is any metric space with Borel algebra $\mathcal{B}$.

Theorem 1 (Billingsley) *Let $G : (U, \mathcal{B}) \rightarrow (\Xi, \Sigma)$ be a measurable function and Q a probability distribution on $(U, \mathcal{B})$. Then $QG^{-1}(A) := Q(G^{-1}A)$ defines a probability measure on (Ξ, Σ) , and if G is Q -a.s. continuous, then*

$$Q^\nu \rightarrow Q \implies Q^\nu G^{-1} \rightarrow QG^{-1}.$$

Given a Q -a.s. continuous G and a discrete approximation $Q^\nu = \sum_{i=1}^\nu p_i^\nu \delta_{u_i^\nu}$ of Q , Theorem 1 says that the discrete measures

$$P^\nu := Q^\nu G^{-1} = \sum_{i=1}^\nu p_i^\nu \delta_{G(u_i^\nu)}$$

converge weakly to $P = QG^{-1}$ whenever $Q^\nu \rightarrow Q$. It is natural to try to choose discrete approximations Q^ν which are as close as possible to the uniform distribution Q . Modern methods of numerical integration do exactly this; see the books of Niederreiter (1992) and Sloan and Joe (1994). Much of this theory has evolved around the following notion of distance from Q .

Definition 2 *The star-discrepancy of a point set $\{u_1, \dots, u_\nu\} \subset [0, 1]^d$ is defined as*

$$D^*(u_1, \dots, u_\nu) = \sup_{C \in \mathcal{C}_0} |Q^\nu(C) - Q(C)|, \quad (4)$$

where

$$Q^\nu = \sum_{i=1}^\nu \frac{1}{\nu} \delta_{u_i},$$

and $\mathcal{C}_0$ is the set of rectangles $C \subset [0, 1]^d$ with $0 \in C$.

The following is a direct consequence of Corollary 11 in Lucchetti et al. (1994).

Proposition 3 For each ν , let $\{u_1^\nu, \dots, u_\nu^\nu\}$ be point sets in the unit cube. The measures

$$Q^\nu = \sum_{i=1}^{\nu} \frac{1}{\nu} \delta_{u_i^\nu}$$

converge weakly to the uniform distribution if and only if $D^*(u_1^\nu, \dots, u_\nu^\nu) \rightarrow 0$.

Thus, if we can find point sets whose star-discrepancy approaches zero as $\nu \nearrow \infty$, we obtain weakly convergent discrete approximations of the uniform distribution. If $P = QG^{-1}$, we can then use the method of inversion to get weakly convergent discretizations of P . In the literature of numerical integration, many methods have been proposed that are aimed at producing point sets that have as low star-discrepancy as possible. It is thus natural to employ them in the construction of discrete measures P^ν and the corresponding approximations (SP^ν) . This is what the present paper is about.

This approach to discretization of stochastic programs is close in spirit to the method proposed in Pflug (2001), where the aim is to find discrete measures P^ν that are as close as possible to P in the sense of the so called *Wasserstein-distance*. In general, the problem of finding a discrete measure that minimizes a distance from a given measure can be very hard. Fortunately, in the case of star-discrepancy, many efficient methods are already available.

Example 4 (low discrepancy sequences) Low discrepancy sequences are infinite sequences whose first ν points have low discrepancy for all ν . Examples are

1. *Faure sequence* (Faure, 1982). A FORTRAN 77-routine for Faure sequence has been implemented by Fox (1986) as ACM Algorithm 647.
2. *Sobol sequence* (Sobol', 1967). A C-routine for Sobol sequence is available in GSL (Gnu Scientific Library, www.gnu.org/software/gsl/gsl.html).
3. *Niederreiter sequence* (Bratley et al., 1992). This is also available in GSL.

These satisfy

$$D^*(u_1, \dots, u_\nu) \leq C \frac{(\log \nu)^d}{\nu} \quad \forall \nu, \quad (5)$$

for a constant C independent of ν . These examples fall in the general class of (t, s) -sequences; see (Niederreiter, 1992, Chapter 4). Figure 1 displays the first 15 and 127 points for Faure and Sobol sequences in the 2-dimensional unit cube.

In direct numerical integration, infinite low discrepancy sequences are useful in that after evaluating a ν -point quadrature, one can continue to compute the next $(\nu + 1)$ -point quadrature simply by evaluating the function at one new point. In stochastic programming this advantage is lost since, in general, the solution x and thus the integrand $f(x, \cdot)$ changes every time a new point (scenario) is added to the problem. This raises the question whether it is possible to obtain more accurate quadratures if it is not required that ν points of a $(\nu + 1)$ -point quadrature are the points of the ν -point quadrature. This is indeed possible.

Example 5 (low discrepancy point sets) A set of points $\{u_1, \dots, u_\nu\}$ in the unit cube is called a low discrepancy point set if it has low discrepancy. Examples are

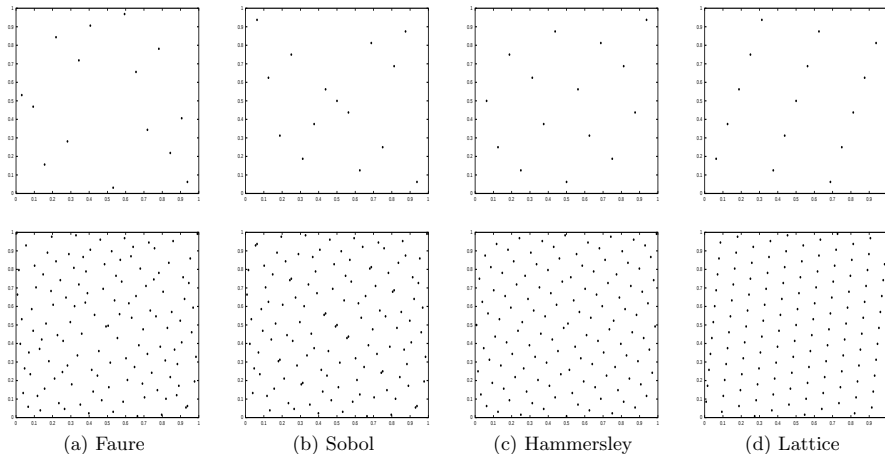


Figure 1: Discretizations of the uniform distribution by 15 and 127 points

1. *Hammersley point sets* (Hammersley, 1960). Hammersley point sets can be obtained quite easily from the Halton sequence (Halton, 1960) that has been implemented in Fox (1986).
2. *(t, m, s)-nets* are a general class of low discrepancy point sets that are described in detail in (Niederreiter, 1992, Chapter 4).

These satisfy

$$D^*(u_1, \dots, u_\nu) \leq C \frac{(\log \nu)^{d-1}}{\nu}, \quad (6)$$

for a constant C independent of ν . Figure 1 displays 15 and 127 Hammersley points in a 2-dimensional unit cube.

There is another class of quadratures designed to take advantage of additional regularity properties of integrands.

Example 6 (lattice rules) Lattice rules are a general family of methods for generating point sets with low discrepancy; see for example (Niederreiter, 1992, Chapter 5), Sloan and Joe (1994) and L'Ecuyer and Lemieux (2000). For each d and ν there exist lattice rules that give point sets satisfying (5); see (Niederreiter, 1992, page 115). This is not as good as (6), but for certain classes of functions much tighter error bounds can be derived.

Figure 1 displays lattices of 15 and 127 points produced by the so called Korobov lattice rule (Korobov, 1959). Parameters required by the method were provided by Pierre L'Ecuyer and Christiane Lemieux (personal communication).

It can be shown that if $Q^\nu = \sum_{i=1}^\nu \frac{1}{\nu} \delta_{u_i^\nu}$, then

$$|E^{Q^\nu} \varphi - E^Q \varphi| \leq V(\varphi) D^*(u_1^\nu, \dots, u_\nu^\nu),$$

where $V(\varphi)$ is the variation of φ in the sense of Hardy and Krause; see (Niederreiter, 1992, Section 2.2). In direct integration, the above methods achieve the asymptotic convergence rate of ν^{-1} , whereas for pure Monte Carlo methods the rate is $\nu^{-\frac{1}{2}}$; more precisely, in Monte Carlo integration, the standard deviation of the integration error is $\sigma(\varphi)\nu^{-\frac{1}{2}}$, where $\sigma(\varphi)$ is the standard deviation of φ . In Monte Carlo, the error bound is independent of the dimension of the space, whereas the bounds in the above examples may depend on the dimension so that the actual error bound achieved in practice is much greater than $\sigma(\varphi)\nu^{-\frac{1}{2}}$. In numerical tests, however, low discrepancy point sets and sequences are often much more efficient than pure Monte Carlo; see for example Sloan and Joe (1994). In the tests of Section 4, one can see a similar effect in discretizations of stochastic programs.

3 Epi-convergence of the objectives

Given efficient procedures for constructing finitely supported measures P^ν that converge weakly to P , our next step is to find conditions that guarantee the epi-convergence of $E^{P^\nu}f$ to $E^P f$. Recall that the *domain* of an extended real-valued function g is the set $\text{dom } g = \{x \mid g(x) < \infty\}$, and its *lower closure* is the function

$$(\text{cl } g)(x) = \liminf_{y \rightarrow x} g(y).$$

A function is called *lower semicontinuous* (lsc) if it is equal to its lower closure. The *lower epi-limit* of a sequence $\{F^\nu\}$ of functions is the lsc function given by

$$(\text{e-lim inf } F^\nu)(x) = \inf_{x^\nu \rightarrow x} \liminf_{\nu \rightarrow \infty} F^\nu(x^\nu)$$

and the *upper epi-limit* is the lsc function given by

$$(\text{e-lim sup } F^\nu)(x) = \inf_{x^\nu \rightarrow x} \limsup_{\nu \rightarrow \infty} F^\nu(x^\nu).$$

If $\text{e-lim inf } F^\nu = \text{e-lim sup } F^\nu$, then the common limit, denoted $\text{e-lim } F^\nu$, is called the *epi-limit* of $\{F^\nu\}$ and the sequence is said to *epi-converge* to it.

Epi-convergence has many important implications in studying approximations of minimization problems. The following is one of them; see (Attouch, 1984, Section 2.2).

Theorem 7 *If $\text{e-lim } F^\nu = F$, then*

$$\limsup_{\nu \rightarrow \infty} \inf F^\nu \leq \inf F,$$

and if there is a convergent sequence $x^k \rightarrow x$ such that $x^k \in \text{argmin } F^{\nu^k}$ for some subsequence $\{\nu^k\}_{k=1}^\infty$, then $x \in \text{argmin } F$ and $\inf F^{\nu^k} \rightarrow \inf F$. In particular, if there is a compact set C such that $\text{argmin } F^\nu \cap C \neq \emptyset$ for all ν , then $\inf F^\nu \rightarrow \inf F$.

Our proof of epi-convergence for (SP^ν) is based on ideas from Artstein and Wets (1995), where $\{P^\nu\}_{\nu=1}^\infty$ was a sequence of empirical measures, and the main tools were the strong law of large numbers and an approximation algorithm due to Beer (1987). In our case, $\{P^\nu\}_{\nu=1}^\infty$ is a weakly convergent non-random sequence, and our main tools are the algorithm of Beer and Theorem 8

below. Recall that a sequence of functions $\{\varphi^\nu\}_{\nu=1}^\infty$ converges continuously to a function φ^0 at $\xi \in \Xi$ if

$$\varphi^\nu(\xi^\nu) \rightarrow \varphi^0(\xi),$$

whenever $\xi^\nu \rightarrow \xi$. The following is based on (Artstein and Wets, 1994, Remark 4.3) and (Schultz, 2000, pp. 67–68).

Theorem 8 *If $P^\nu \rightarrow P^0$, $\varphi^\nu \rightarrow \varphi^0$ continuously at P^0 -almost every $\xi \in \Xi$, and if for each $\epsilon > 0$ there exists a measurable set $K_\epsilon \subset \Xi$ and a bound $b_\epsilon \in \mathbb{R}$, such that for each $\nu = 0, 1, 2, \dots$*

1. $|\varphi^\nu(\xi)| \leq b_\epsilon$ for P^ν -almost every $\xi \in K_\epsilon$,
2. $\int_{\Xi \setminus K_\epsilon} |\varphi^\nu(\xi)| P^\nu(d\xi) < \epsilon$,

then

$$\lim_{\nu \rightarrow \infty} E^{P^\nu} \varphi^\nu = E^{P^0} \varphi^0.$$

Proof. Choose an $\epsilon > 0$ and let $b_\epsilon \in \mathbb{R}$ and $K_\epsilon \subset \Xi$ be the bound and the measurable set, respectively, provided by the last condition. For each ν , let φ'_ϵ be the bounded function whose value at a point ξ is the projection of $\varphi^\nu(\xi)$ onto the interval $[-b_\epsilon, b_\epsilon]$. Then $|\varphi'_\epsilon| \leq |\varphi^\nu|$, $\varphi'_\epsilon(\xi) = \varphi^\nu(\xi)$ for P^ν -almost every $\xi \in K_\epsilon$, and $\varphi'_\epsilon \rightarrow \varphi^0_\epsilon$ continuously P^0 -a.s. We have

$$|E^{P^\nu} \varphi^\nu - E^{P^0} \varphi^0| \leq |E^{P^\nu} \varphi^\nu - E^{P^\nu} \varphi'_\epsilon| + |E^{P^\nu} \varphi'_\epsilon - E^{P^0} \varphi'_\epsilon| + |E^{P^0} \varphi'_\epsilon - E^{P^0} \varphi^0|,$$

where the second term on the right converges to zero by (Billingsley, 1968, Theorem 5.5), and for $\nu = 0, 1, \dots$,

$$|E^{P^\nu} \varphi^\nu - E^{P^\nu} \varphi'_\epsilon| = \left| \int_{\Xi \setminus K_\epsilon} [\varphi^\nu(\xi) - \varphi'_\epsilon(\xi)] P^\nu(d\xi) \right| \leq 2 \int_{\Xi \setminus K_\epsilon} |\varphi^\nu(\xi)| P^\nu(d\xi) \leq 2\epsilon.$$

Since $\epsilon > 0$ was arbitrary, the result follows. $\square$

In particular, if φ is P^0 -a.s. continuous and bounded, then

$$E^{P^\nu} \varphi \rightarrow E^{P^0} \varphi.$$

Note that this is also implied directly by Theorem 1.

We can now state our epi-convergence result for $E^{P^\nu} f$.

Theorem 9 *Let $P^\nu \rightarrow P^0$. If for each $x \in X$,*

1. *there is an open set $N \ni x$ such that f is bounded from below on $N \times \Xi$,*
2. *$(\text{cl } f)(x, \cdot) = f(x, \cdot)$ P^0 -a.s.,*

then

$$\text{e-lim inf } E^{P^\nu} f \geq E^P f.$$

If for each $x \in \text{dom } E^{P^0} f$,

3. *there is a sequence $x^\nu \rightarrow x$ such that P^ν and $\varphi^\nu := f(x^\nu, \cdot)$ satisfy the conditions of Theorem 8,*

then

$$\text{e-lim sup } E^{P^\nu} f \leq E^P f.$$

Proof. To verify the first claim, fix an $x \in X$ and let $x^\nu \rightarrow x$. According to Beer (1987) (see also the proof of (Bertsekas and Shreve, 1978, Theorem 7.14)), the first condition implies that there exists a sequence $\{f^k\}$ of bounded from below Lipschitz functions on $N \times \Xi$ such that $f^k \nearrow \text{cl } f$. The functions $g^k = \min\{f^k, k\}$ are also Lipschitz with $g^k \nearrow \text{cl } f$ but, in addition, they are bounded. Then

$$E^{P^\nu} f(x^\nu) \geq E^{P^\nu} g^k(x^\nu) \geq E^{P^\nu} g^k(x) - L^k d(x^\nu, x) \quad \forall \nu, k$$

where L^k is the Lipschitz constant of g^k and d is the metric on X . Since $P^\nu \rightarrow P$, we get

$$\liminf_{\nu \rightarrow \infty} E^{P^\nu} f(x^\nu) \geq \liminf_{\nu \rightarrow \infty} E^{P^\nu} g^k(x) = E^{P^0} g^k(x) \quad \forall k,$$

where $E^{P^0} g^k(x) \nearrow E^{P^0}(\text{cl } f)(x)$ as $k \nearrow \infty$ by the monotone convergence theorem. Since by the second condition, $E^{P^0}(\text{cl } f)(x) = E^{P^0} f(x)$, and since $x \in X$ and $x^\nu \rightarrow x$ were arbitrary, the claim follows.

For the second claim, it suffices to consider points $x \in \text{dom } E^{P^0} f$. The result then follows from the third condition and Theorem 8. $\square$

Note that choosing $P^\nu = P^0$ for all ν , the first claim shows that under conditions 1 and 2 $E^{P^0} f$ is lsc.

Theorem 9 applies to any discretization method that generates weakly convergent sequences of measures. In particular, it yields epi-convergence results for the method proposed in Pflug (2001) where the measures P^ν are chosen by minimizing the distance of P^ν from P in the sense of the Wasserstein metric; see also Pflug and Hochreiter (2003).

Theorem 9 is close to the epi-convergence results in Lucchetti and Wets (1993), Artstein and Wets (1994), Zervos (1999), but it has some advantages. The results of Zervos (1999) concerned real-valued functions which do not allow modeling constraints as infinite penalties. In Lucchetti and Wets (1993), Artstein and Wets (1994), the tightness-like conditions are much stronger than condition 3 above. In particular, condition 3 only asks for a measurable set K_ϵ instead of a compact one, and instead of all sequences $x^\nu \rightarrow x$, it only involves *one* sequence for each $x \in \text{dom } E^{P^0} f$. This is important since it is the tightness-like conditions that are usually hard to check for discretizations in practice. Also, requiring condition 3 to hold at all $x \in X$ would imply $\text{dom } E^{P^0} f = X$. In a sense, we have traded the stronger tightness-like conditions for conditions 1 and 2, which are often much easier to check; see Section 4. The lower-boundedness property in condition 1 holds in many applications arising in practice. Condition 2 holds in particular if f is lsc, which was assumed in (Bertsekas and Shreve, 1978, Section 8.3) and Lucchetti and Wets (1993). According to the remark after Theorem 7, we have the following simplified version, which is often sufficient in applications.

Corollary 10 *Let $P^\nu \rightarrow P^0$ and assume that f is lsc. If*

1. *for each $x \in X$, there is an open set $N \ni x$ such that f is bounded from below on $N \times \Xi$,*
2. *for each $x \in \text{dom } E^{P^0} f$, $f(x, \cdot)$ is P^0 -a.s. continuous and bounded,*

then the functions $E^{P^\nu} f$ both pointwise and epi-converge to $E^P f$.

4 Numerical tests

4.1 Mean-variance portfolio optimization

We start the testing with a model which can be solved exactly. Discretization is unnecessary in such cases but here we get to compare the approximate solutions with the exact one. We will study the mean-variance model

$$\begin{aligned}
 & \underset{x \in \mathbb{R}^n}{\text{minimize}} && E^{P^0}(r \cdot x - \bar{r} \cdot x)^2 \\
 & \text{subject to} && \bar{r} \cdot x \geq w, \\
 & && \sum_{i=1}^n x_i \leq 1, \\
 & && x \in C,
 \end{aligned} \tag{MP}$$

where $x = (x_1, \dots, x_n)$ is a portfolio of assets, $r = (r_1, \dots, r_n)$ is the vector of returns (that is, r_i is the ratio of the final and initial price of asset i), $r \cdot x = \sum_{i=1}^n r_i x_i$ is the terminal wealth, w is the required level of expected wealth and C is the set of feasible portfolios. The components of the return vector r are random variables with joint distribution P^0 and expectation $\bar{r}$. The expectation in (MP) can be computed explicitly as

$$E^{P^0}(r \cdot x - \bar{r} \cdot x)^2 = E^{P^0}[(r - \bar{r}) \cdot x]^2 = E^{P^0}[x \cdot (r - \bar{r})(r - \bar{r})^T x] = x \cdot V x,$$

where $V = E^{P^0}[(r - \bar{r})(r - \bar{r})^T]$ is the covariance matrix of r . If V and $\bar{r}$ are known, (MP) can then be solved without discretization with standard solvers yielding the optimal value and optimal solution.

To test the performance of integration quadratures in discretization, we approximate problem (MP) by the discretizations

$$\begin{aligned}
 & \underset{x \in \mathbb{R}^n}{\text{minimize}} && \sum_{i=1}^{\nu} p_i^{\nu} (r_i^{\nu} \cdot x - \bar{r} \cdot x)^2 \\
 & \text{subject to} && \bar{r} \cdot x \geq w, \\
 & && \sum_{i=1}^n x_i \leq 1, \\
 & && x \in C.
 \end{aligned} \tag{MP}^{\nu}$$

Under mild conditions, convergence of optimal values and solutions can be guaranteed. Recall that the *support*, $\text{supp } P$, of a measure P is the intersection of all closed sets of full measure. For a Borel probability measure $\text{supp } P$ is well defined and unique with $P(\text{supp } P) = 1$.

Proposition 11 *Assume that $\text{supp } P^0$ is bounded, C is closed, and that the measures*

$$P^{\nu} = \sum_{i=1}^{\nu} p_i^{\nu} \delta_{r_i^{\nu}}$$

converge weakly to P^0 and satisfy $\text{supp } P^{\nu} \subset \text{supp } P^0$. If the feasible set is bounded, then the optimal values of $(MP)^{\nu}$ converge to that of (MP) and the cluster points of the solutions of $(MP)^{\nu}$ are solutions of (MP).

Proof. This fits the format of (SP) with $\xi = r$ and

$$f(x, r) = (r \cdot x - \bar{r} \cdot x)^2 + \delta_{C'}(x),$$

where $C' = \{x \in C \mid \bar{r} \cdot x \geq w, \sum_{i=1}^n x_i \leq 1\}$. By Theorem 7, it suffices to verify the conditions of Corollary 10. Lower semicontinuity and condition 1 are clear. Since $\text{supp } P^\nu \subset \text{supp } P^0$, we can assume that $\Xi = \text{supp } P^0$, and then condition 2 holds by boundedness of $\text{supp } P^0$. $\square$

In our test, the number of assets $n = 10$ and

$$r = \bar{r} + 12L(u - \frac{1}{2}e),$$

where u is uniformly distributed in the 10-dimensional unit cube, L is a 10×10 matrix and e is a vector of ones. Then $\text{supp } P^0$ is bounded, r has mean $\bar{r}$ and variance $V = LL^T$. We can then solve (MP) exactly by standard QP-solvers and the discretizations (MP^ν) are easily generated by the integration quadratures described in Section 2.2. Note that the objective of (MP^ν) can be written as $x \cdot V^\nu x$, where $V^\nu = E^{P^\nu}[(r - \bar{r})(r - \bar{r})^T] = \sum_{i=1}^\nu p_i^\nu (r_i^\nu - \bar{r})(r_i^\nu - \bar{r})^T$, so (MP^ν) can also be solved with a QP-solver. We chose $C = \mathbb{R}_+^n$, which means that “short selling” is prohibited.

With our choices of $\bar{r}$ and V , the optimal value in the original problem (MP) turned out to be 1.9221. Figure 2 shows the development of the optimal value of (MP^ν) as a function of the number of quadrature points $\nu = 100, 200, \dots, 10000$ for six quadratures. In our implementation, the number of points in the lattice rule is restricted to powers of 2. Each quadrature produces discretizations whose optimal value seems to converge toward the exact value 1.9221. The objective values corresponding to Halton sequence seem to behave most stably whereas Hammersley exhibits slowest convergence.

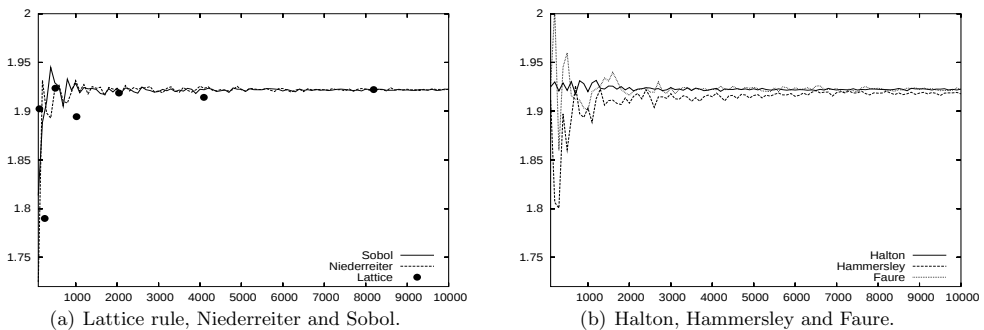


Figure 2: Optimal values of (MP^ν) as a function of ν .

For comparison, we discretized the problem also with Monte Carlo sampling. Almost sure epi-convergence of such discretizations have been established under quite general conditions in Artstein and Wets (1995). These conditions are strictly weaker than those in Theorem 9, which guarantees *sure* (not just almost sure) epi-convergence. For each $\nu = 100, 200, \dots, 10000$, we generated 250

discretizations and computed the average and the 90% confidence interval of the corresponding optimal values. In other words, 25 out of the 250 discretizations obtained with Monte Carlo fell outside this interval. The results are displayed in Figure 3. The optimal values obtained with Sobol sequence are repeated for reference. The average of the Monte Carlo values seems to converge towards the correct value but the convergence of the confidence interval seems slow.

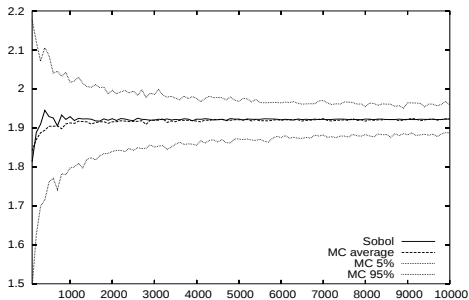


Figure 3: Average and 90% confidence intervals for Monte Carlo.

Figure 4(a) displays the logarithmic error in the optimal value for Sobol discretizations as a function of $\ln \nu$. Figure 4(b) does the same for Monte Carlo. Interestingly, in both cases, there seems to be an upper bound on the log-error which is linear in $\ln \nu$. For Sobol, the slope is roughly -1 whereas for MC it seems to be close to -1/2. This kind of quantitative behavior is not explained by the general results of Section 3, but since, in this particular example, the objective is strongly convex, the quantitative results of Römisich and Schultz (1993), Shapiro (1994), Rachev and Römisich (2002) may apply. However, it is not at all clear how one should quantitatively compare the performance of a deterministic discretization method with a random one. One may be able to give convergence rate results for a sequence of deterministic approximations of a specially structured stochastic program as the number of quadrature points is increased, but the same cannot be done for Monte Carlo. Indeed, no matter how large a sample is taken, Monte Carlo can lead to an arbitrarily bad approximation of the original problem. For Monte Carlo, one may be able to estimate confidence intervals (or “statistical bounds”) for the approximate optimal values, but such intervals have to do only with the randomness of the approximation method, not with the actual problem to be solved. The above way of computing confidence intervals and averages from 250 independent sample approximations does not seem very practical since it involves the solution of 250 optimization problems.

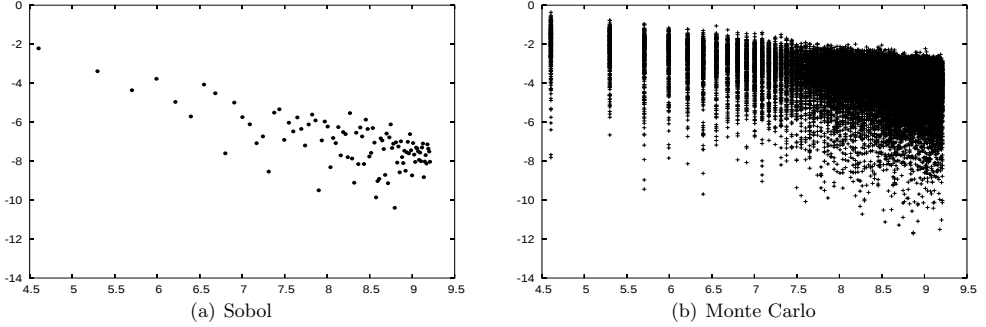


Figure 4: Log of objective errors as a function of $\ln \nu$.

4.2 Utility maximization

The objective in the Markowitz model penalizes for exceeding the expected wealth $\bar{r} \cdot x$. When the distribution of r is symmetric, this does not matter, but in practice, the distribution of r is usually nonsymmetric since $r \geq 0$. The following utility maximization problem still makes sense

$$\begin{aligned}
 & \underset{x \in \mathbb{R}^n}{\text{maximize}} && E^{P^0} u(r \cdot x) && (UP) \\
 & \text{subject to} && \sum_{i=1}^n x_i \leq w_0, \\
 & && x \in C.
 \end{aligned}$$

Here x , r and C are as in the previous example and u measures the utility from terminal wealth. The components of the return vector r are nonnegative random variables with joint distribution P^0 .

In general, (UP) cannot be solved analytically, so we will consider the discretizations

$$\begin{aligned}
 & \underset{x \in \mathbb{R}^n}{\text{maximize}} && \sum_{i=1}^{\nu} p_i^{\nu} u(r_i^{\nu} \cdot x) && (UP^{\nu}) \\
 & \text{subject to} && \sum_{i=1}^n x_i \leq w_0, \\
 & && x \in C.
 \end{aligned}$$

Proposition 12 Assume $\text{supp } P^0 \subset \mathbb{R}_+^n$, u is continuous and bounded on $\mathbb{R}_+$, C is closed and contained in $\mathbb{R}_+^n$ (short selling is not allowed) and that the measures

$$P^{\nu} = \sum_{i=1}^{\nu} p_i^{\nu} \delta_{r_i^{\nu}}$$

converge weakly to P^0 and satisfy $\text{supp } P^{\nu} \subset \mathbb{R}_+^n$. Then the optimal values of (UP^{ν}) converge to that of (UP) and the cluster points of the solutions of (UP^{ν}) are solutions of (UP) .

Proof. This fits the format of (SP) with $\Xi = \mathbb{R}_+^n$, $\xi = r$, and

$$f(x, r) = -u(r \cdot x) + \delta_{C'}(x),$$

where $C' = \{x \in C \mid \sum_{i=1}^n x_i \leq w_0\}$. By Theorem 7, it suffices to verify the conditions of Corollary 10. Since u is continuous and C is closed, f is lsc. Condition 1 follows from the boundedness of u on $\mathbb{R}_+$, and the fact that $r \cdot x \in \mathbb{R}_+$ for all $r \geq 0$ and $x \in C \subset \mathbb{R}_+$. Condition 2 follows from the boundedness and continuity of u on $\mathbb{R}_+$. $\square$

Many familiar utility functions, like the exponential utility, are bounded on $\mathbb{R}_+$. More general utility functions are easily modified to be bounded on $\mathbb{R}_+$ in a way that does not affect computations in practice.

In our test, the number of assets $n = 10$, r is log-normally distributed, $u(w) = -\exp(-w)$ and $C = \mathbb{R}_+^n$. Figure 5 shows the development of the optimal value of (UP^ν) as a function of the number of quadrature points $\nu = 100, 110, \dots, 2500$ for six quadratures. Again, the quadratures seem to converge to a common value, but this time, Halton and Faure exhibit slowest convergence whereas Sobol, Niederreiter and lattice rule seem to work best. Figure 6 depicts the development of the average and the 90% confidence interval for the optimal values obtained with 250 Monte Carlo samples for each value of ν . The values obtained with Sobol are shown for reference. The average converges to the same value as the optimal values obtained by the quadratures but the confidence interval narrows down very slowly.

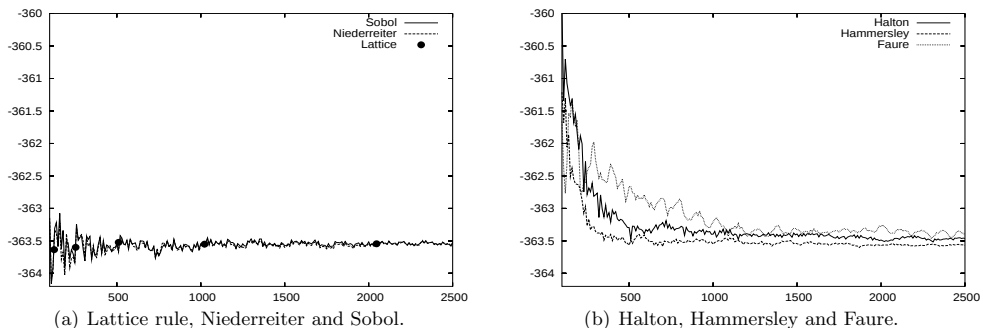


Figure 5: Optimal values of (UP^ν) as a function of ν .

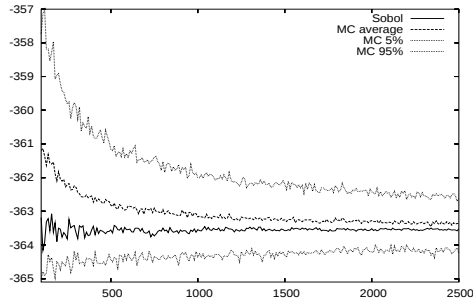


Figure 6: Average and 90% confidence intervals for Monte Carlo.

4.3 Super-replication of contingent claims

Consider the problem

$$\begin{aligned}
 & \underset{V, \theta}{\text{minimize}} && V && (PP) \\
 & \text{subject to} && S_0 \cdot \theta \leq V, \\
 & && S \cdot \theta \geq F, \quad P^0\text{-a.s.} \\
 & && \theta \in C,
 \end{aligned}$$

where V is the wealth invested in a portfolio $\theta = (\theta_1, \dots, \theta_J)$ of assets that have prices $S_0 = (S_0^1, \dots, S_0^J)$ at the beginning and $S = (S^1, \dots, S^J)$ at the end of a holding period and F is a cash-flow at the end of the holding period. S and F are random variables with joint distribution P^0 . (PP) can be interpreted as a pricing problem of a seller of the cash-flow F ; see for example King (2002) and (Korf, 2002, Section 7). The seller tries to find the least amount of initial wealth that can be used to buy a portfolio that is almost surely worth at least F at the end of the holding period. The optimum value of (PP) is called the seller's price of F and an optimal portfolio θ is called a seller's hedge.

(PP) is a semi-infinite linear programming problem and, in general, impossible to solve analytically. Replacing P^0 by a discrete measure $P^\nu = \sum_{i=1}^\nu p_i^\nu \delta_{(S_i^\nu, F_i^\nu)}$ with $p_i^\nu > 0$, for all $i = 1, \dots, \nu$ gives the problem

$$\begin{aligned}
 & \underset{V, \theta}{\text{minimize}} && V && (PP^\nu) \\
 & \text{subject to} && S_0 \cdot \theta \leq V, \\
 & && S_i^\nu \cdot \theta \geq F_i^\nu, \quad i = 1, \dots, \nu, \\
 & && \theta \in C,
 \end{aligned}$$

which is an LP problem for which many solvers are available.

In this example, the feasible region depends on the measure, and $\text{dom } E^{P^0} f$ is impossible to characterize explicitly in general. This problem does not fit the frameworks of Lucchetti and Wets (1993), Artstein and Wets (1994), Zervos (1999) and Shapiro (2000).

Proposition 13 Assume that the points $\{(S_i^\nu, F_i^\nu)\}_{i=1}^\nu$ are all contained in $\text{supp } P^0$ and that for some $\{p_i^\nu\}_{i=1}^\nu$, $\nu = 0, 1, 2, \dots$, with $p_i^\nu > 0$, for all $i = 1, \dots, \nu$, the measures

$$P^\nu = \sum_{i=1}^\nu p_i^\nu \delta_{(S_i^\nu, F_i^\nu)}$$

converge weakly to P^0 . If the feasible set is bounded, then the optimal values of (PP^ν) converge to the seller's price of F and the cluster points of the solutions of (PP^ν) are seller's hedges for F .

Proof. This can be written as (SP) with $x = (V, \theta)$, $\xi = (S, F)$ and

$$f(V, \theta, S, F) = V + \delta_C(\theta) + \delta_{C_0}(V, \theta) + \delta_{C_1}(\theta, S, F),$$

where

$$C_0 = \{(V, \theta) \mid S_0 \cdot \theta \leq V\}$$

and

$$C_1 = \{(\theta, S, F) \mid S \cdot \theta \geq F\}.$$

Since C_0 and C_1 are closed, f is lsc. It is also clear that condition 1 of Corollary 10 holds. To verify condition 2, note first that for each $(V, \theta) \in \text{dom } E^{P^0} f$, $f(V, \theta, \cdot)$ is the constant function V on the set

$$C_1(\theta) = \{(S, F) \mid S \cdot \theta \geq F\}$$

which is of full measure. Since $C_1(\theta)$ is closed, we must have $\text{supp } P^0 \subset C_1(\theta)$ for every $(V, \theta) \in \text{dom } E^{P^0} f$. Thus, condition 2 holds if we let $\Xi = \text{supp } P^0$, which is legitimate since $\text{supp } P^\nu \subset \text{supp } P^0$. $\square$

In our test, the set of assets consists of cash, SP500 index and 28 European call and put options on the index with maturity of 17 calendar days. The value of S is fully determined by the value of the index at the maturity which is assumed to be log-normally distributed. The cash-flow F is taken to be that of a call option with the same maturity but different strike than any other call included in S .

Figure 7(a) displays the objective values obtained with Sobol along with the averages and 90% confidence intervals obtained with Monte Carlo from 250 samples for each value of $\nu = 100, 110, \dots, 2500$.

Our random variable being one-dimensional in the current problem suggests using Gaussian quadratures. The use of Gauss-Hermite quadrature for discretizing the normal distribution has been already described in Subsection 2.1. We will also utilize Gauss-Legendre quadrature that gives convergent discretizations of the one-dimensional uniform distribution on $[0, 1]$. From this we obtain discretizations of the normal distribution by mapping each point through the inverse of the normal distribution function. The results are shown in Figure 7(b). With $\nu = 60$, the optimal values obtained with Gauss-Legendre quadrature have converged to the same value as the optimal values obtained with Sobol after 2500 points. Gauss-Hermite is almost as good but it results in slight oscillations.

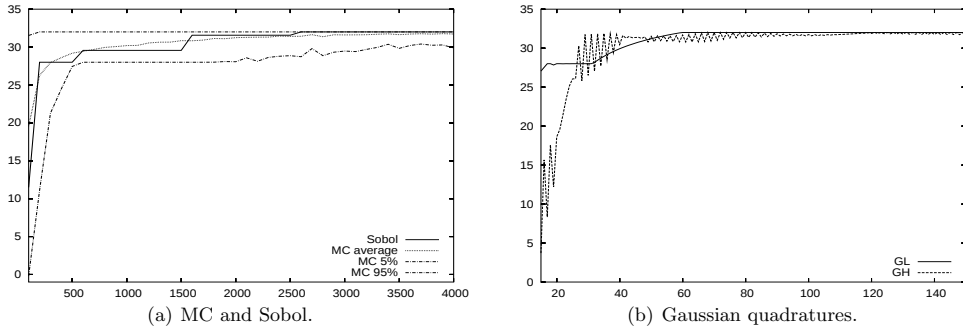


Figure 7: Optimal values of (PP^ν) as a function of ν .

References

- Z. Artstein and R. J.-B. Wets. Stability results for stochastic programs and sensors, allowing for discontinuous objective functions. *SIAM J. Optim.*, 4(3):537–550, 1994.
- Z. Artstein and R. J.-B. Wets. Consistency of minimizers and the SLLN for stochastic programs. *J. Convex Anal.*, 2(1-2):1–17, 1995.
- H. Attouch. *Variational convergence for functions and operators*. Pitman (Advanced Publishing Program), Boston, MA, 1984.
- H. Attouch and R. J.-B. Wets. Quantitative stability of variational systems. I. The epigraphical distance. *Trans. Amer. Math. Soc.*, 328(2):695–729, 1991.
- G. Beer. A geometric algorithm for approximating semicontinuous function. *J. Approx. Theory*, 49(1):31–40, 1987.
- D. P. Bertsekas and S. E. Shreve. *Stochastic optimal control, The discrete time case*, volume 139 of *Mathematics in Science and Engineering*. Academic Press Inc., New York, 1978.
- P. Billingsley. *Convergence of probability measures*. John Wiley & Sons Inc., New York, 1968.
- P. Billingsley. *Convergence of probability measures*. John Wiley & Sons Inc., New York, 2nd edition, 1999.
- J. R. Birge and R. J.-B. Wets. Designing approximation schemes for stochastic optimization problems, in particular for stochastic programs with recourse. *Math. Programming Stud.*, (27):54–102, 1986.
- P. Bratley, B. L. Fox, and H. Niederreiter. Implementation and tests of low-discrepancy sequences. *ACM Trans. Model. Comput. Simul.*, 2(3):195–213, 1992.
- I. Deák. Multidimensional integration and stochastic programming. In *Numerical techniques for stochastic optimization*, volume 10 of *Springer Ser. Comput. Math.*, pages 187–200. Springer, Berlin, 1988.

- M. A. H. Dempster and R. T. Thompson. EVPI-based importance sampling solution procedures for multistage stochastic linear programmes on parallel MIMD architectures. *Ann. Oper. Res.*, 90:161–184, 1999.
- J. Dupačová and R. J.-B. Wets. Asymptotic behavior of statistical estimators and of optimal solutions of stochastic optimization problems. *Ann. Statist.*, 16(4):1517–1549, 1988.
- H. Faure. Discrepance de suites associées à un système de numération (en dimension s). *Acta Arith.*, 41(4):337–351, 1982.
- B. L. Fox. Algorithm 647: Implementation and relative efficiency of quasirandom sequence generators. *ACM Transactions on Mathematical Software*, 12(4):362–376, 1986.
- K. Frauendorfer. *Stochastic two-stage programming*, volume 392 of *Lecture Notes in Economics and Mathematical Systems*. Springer-Verlag, Berlin, 1992.
- J. H. Halton. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numer. Math.*, 2:84–90, 1960.
- J. M. Hammersley. Monte Carlo methods for solving multivariable problems. *Ann. New York Acad. Sci.*, 86:844–874 (1960), 1960.
- K. Høyland, M. Kaut, and S. Wallace. A heuristic for moment-matching scenario generation. *Computational Optimization and Applications*, 24:169–186, 2003.
- K. Høyland and S. W. Wallace. Generating scenario trees for multistage decision problems. *Management Science*, 47(2):295–307, 2001.
- G. Infanger. Monte Carlo (importance) sampling within a Benders decomposition algorithm for stochastic linear programs. *Ann. Oper. Res.*, 39:69–95, 1992.
- P. Kall, A. Ruszczyński, and K. Frauendorfer. Approximation techniques in stochastic programming. In *Numerical techniques for stochastic optimization*, volume 10 of *Springer Ser. Comput. Math.*, pages 33–64. Springer, Berlin, 1988.
- A. J. King. Duality and martingales: a stochastic programming perspective on contingent claims. *Math. Program.*, 91(3, Ser. B):543–562, 2002.
- L. Korf. Stochastic programming duality: L^∞ multipliers for unbounded constraints with an application to mathematical finance. *submitted*, 2002.
- N. M. Korobov. Approximate evaluation of repeated integrals. *Dokl. Akad. Nauk SSSR*, 124:1207–1210, 1959.
- P. L’Ecuyer and C. Lemieux. Variance reduction via lattice rules. *Management Science*, 46(2):1214–1235, 2000.
- R. Lepp. Approximations to stochastic programs with complete recourse. *SIAM J. Control Optim.*, 28(2):382–394, 1990.

- R. Lucchetti, G. Salinetti, and R. J.-B. Wets. Uniform convergence of probability measures: topological criteria. *J. Multivariate Anal.*, 51(2):252–264, 1994.
- R. Lucchetti and R. J.-B. Wets. Convergence of minima of integral functionals, with applications to optimal control and stochastic optimization. *Statist. Decisions*, 11(1):69–84, 1993.
- H. Niederreiter. *Random number generation and quasi-Monte Carlo methods*, volume 63 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.
- T. Pennanen and M. Koivu. Integration quadratures in discretization of stochastic programs. *Stochastic Programming E-print Series*, 2002.
- G. Pflug. Scenario tree generation for multiperiod financial optimization by optimal discretization. *Math. Program.*, 89(2, Ser. B):251–271, 2001.
- G. Pflug and R. Hochreiter. Scenario generation for stochastic multi-stage decision processes as facility location problems. Technical report, Department of Statistics and Decision Support Systems, University of Vienna, 2003.
- W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical recipes in C, The art of scientific computing*. Cambridge University Press, Cambridge, 2nd edition, 1992.
- Svetlozar T. Rachev and Werner Römisch. Quantitative stability in stochastic programming: the method of probability metrics. *Math. Oper. Res.*, 27(4):792–818, 2002.
- S. M. Robinson and R. J.-B. Wets. Stability in two-stage stochastic programming. *SIAM J. Control Optim.*, 25(6):1409–1416, 1987.
- R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998.
- Werner Römisch and Rüdiger Schultz. Stability of solutions for stochastic programs with complete recourse. *Math. Oper. Res.*, 18(3):590–609, 1993.
- R. Schultz. Some aspects of stability in stochastic programming. *Ann. Oper. Res.*, 100:55–84, 2000.
- A. Shapiro. Quantitative stability in stochastic programming. *Math. Programming*, 67(1, Ser. A): 99–108, 1994.
- A. Shapiro. Stochastic programming by monte carlo simulation methods. *Stochastic Programming E-Print Series*, 2000.
- I. H. Sloan and S. Joe. *Lattice methods for multiple integration*. Oxford Science Publications. The Clarendon Press Oxford University Press, New York, 1994.
- I. M. Sobol'. The distribution of points in a cube and the approximate evaluation of integrals. *U.S.S.R. Computational Math. And Math. Phys.*, (4):86–112, 1967.

- S. Vogel and P. Lachout. On continuous convergence and epi-convergence of random functions. II. Sufficient conditions and applications. *Kybernetika (Prague)*, 39(1):99–118, 2003.
- M. Zervos. On the epiconvergence of stochastic optimization problems. *Math. Oper. Res.*, 24(2): 495–508, 1999.

Variance reduction in sample approximations of stochastic programs

Matti Koivu

Department of Management Science,
Helsinki School of Economics,
PL1210 00101 Helsinki, Finland

Abstract

This paper studies the use of randomized Quasi-Monte Carlo methods (RQMC) in sample approximations of stochastic programs. In high dimensional numerical integration, RQMC methods often substantially reduce the variance of sample approximations compared to MC. It seems thus natural to use RQMC methods in sample approximations of stochastic programs. It is shown, that RQMC methods produce epi-convergent approximations of the original problem. RQMC and MC methods are compared numerically in five different portfolio management models. In the tests, RQMC methods outperform MC sampling substantially reducing the sample variance and bias of optimal values in all the considered problems.

Keywords: Stochastic optimization, discretization, variance reduction techniques, randomized quasi-monte carlo methods, antithetic variates.

1 Introduction

Let Ξ be Borel subset of $\mathbb{R}^d$, and Σ the Borel σ -algebra on Ξ . Let P be a probability measure on (Ξ, Σ) , and f an extended real-valued function on $\mathbb{R}^n \times \Xi$, such that $f(x, \cdot)$ is measurable for every $x \in \mathbb{R}^n$. This paper studies numerical solution through discretization of stochastic programs of the form

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad E^P f(x) := \int_{\Xi} f(x, \xi) P(d\xi), \quad (SP)$$

where the integral is interpreted as $+\infty$ when $f(x, \cdot) \notin L^1(\Xi, \Sigma, P)$. The decision variable x is not a function of ξ , so (SP) represents a *static* (one-stage) stochastic program. By allowing f to take on the value $+\infty$ we can incorporate constraints into the objective, which makes (SP) a very general model for optimal static decision making under uncertainty. Unlike most studies of stochastic programs, we do not assume the feasible set

$$\text{dom } E^P f(x) = \{x \in \mathbb{R}^n \mid f(x, \cdot) \in L^1(\Xi, \Sigma, P)\}$$

to be known a priori. This is essential e.g. in stochastic programs without relatively complete recourse and in certain financial applications, where the determination of the feasible set is part of the problem rather than its statement; see Subsection 4.2.

A common approach to solving (SP) , is to replace P by a finitely supported measure of the form

$$P^\nu = \sum_{i=1}^{\nu} p_i^\nu \delta_{\xi_i^\nu},$$

where $\delta_{\xi_i^\nu}$ denotes the unit mass located at ξ_i^ν . This yields

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad E^{P^\nu} f(x) := \sum_{i=1}^{\nu} p_i^\nu f(x, \xi_i^\nu), \quad (SP^\nu)$$

which, is often easier to solve than (SP) . In general, the aim is to choose P^ν so that (SP^ν) is a good approximation of (SP) and that the number ν of support points of P^ν is small enough to allow for numerical solution of (SP^ν) . The simplest and the best-known method for numerical approximation of high-dimensional integrals is the Monte Carlo method (MC), i.e. random sampling. MC has also become the most popular method for constructing sample approximations of stochastic programs. However, in the literature of numerical integration there are many methods that usually perform better than MC in high-dimensional integration; see e.g. (Boyle et al., 1997, Jäckel, 2002). Quasi-Monte Carlo (QMC) methods can be seen as a deterministic counterpart to the MC method. They are designed to produce point sets that cover the d -dimensional unit hypercube as uniformly as possible. By suitable transformations QMC methods can be used to discretize many other probability distributions as well. They are just as easy to use as MC but they often result in faster convergence of the approximations thus allowing for smaller values of ν and cheaper computations.

L'Ecuyer and Lemieux (2002) review several QMC constructions and their randomizations that have been proposed to provide unbiased estimators and for error estimation. Randomizing QMC methods allows us to view them as variance reduction techniques. Randomized Quasi-Monte Carlo (RQMC) methods can be used just like MC in estimating confidence intervals and variances for sample approximations in numerical integration. RQMC often result in significant variance reduction with respect to MC. In this paper, we apply RQMC to stochastic optimization and obtain similar results. RQMC methods can be viewed as an alternative to MC in computing statistical bounds, as e.g. in Shapiro (2003). In our tests, the bounds for the optimal values obtained with RQMC are consistently tighter than those obtained with MC.

Other variance reduction techniques, like antithetic variates, importance - and latin hypercube sampling have been used in stochastic optimization e.g. in Kouwenberg (2001), Infanger (1992), Higle (1998) and Linderoth et al. (2002). These studies show that variance reduction techniques can significantly improve the accuracy of the sample approximations over MC. It was found in Linderoth et al. (2002) that latin hypercube sampling provides tighter confidence intervals for optimal values than MC. In our tests, the best performing RQMC methods consistently outperform latin hypercube sampling.

Since we are dealing with minimization problems, a natural framework for analyzing approximations is *epi-convergence*; see Attouch (1984) or Rockafellar and Wets (1998) for introduction to epi-convergence. Epi-convergence of the objectives is a minimal property that should be satisfied by any approximation scheme for optimization problems in order to get asymptotic convergence of optimal values and solutions. Epi-convergence for sample approximations of stochastic programs have been proved in Artstein and Wets (1995) for MC, and in Pennanen and Koivu (2003) for QMC. In MC $\{P^\nu\}_{\nu=1}^\infty$ is a sequence of empirical measures, whereas in QMC it is a weakly convergent non-random sequence. In this paper we will show that the epi-convergence result derived in Pennanen and Koivu (2003) for QMC also applies to RQMC methods.

The rest of this paper is organized as follows. Section 2 gives a brief review of the epi-convergence results that will be utilized in this paper. Section 3 reviews the used randomization technique for QMC point sets. It is shown in Section 3 that RQMC methods produce weakly convergent probability measures, thus allowing us to utilize the epi-convergence results derived in Pennanen and

Koivu (2003). In Section 4 we use RQMC methods to construct epi-convergent sample approximations of stochastic programs in various test problems, and compare the behaviour of optimal values numerically with MC.

2 Epi-convergence of sample approximations

Epi-convergence results for sample approximations of stochastic optimization problems have been given in Artstein and Wets (1995) for MC, and in Pennanen and Koivu (2003) for QMC. In MC $\{P^\nu\}_{\nu=1}^\infty$ is a sequence of empirical measures, whereas in QMC it is a weakly convergent non-random sequence, that is

$$E^{P^\nu} \varphi \rightarrow E^P \varphi, \quad (1)$$

for all bounded and continuous functions φ ; see Billingsley (1999). Epi-convergence has many important implications in studying approximations of minimization problems; see e.g. Rockafellar and Wets (1998). The following is one of them; see (Attouch, 1984, Section 2.2).

Theorem 1 *If a sequence of functions F^ν epi-converges to F , then*

$$\limsup_{\nu \rightarrow \infty} \inf F^\nu \leq \inf F,$$

and if there is a convergent sequence $x^k \rightarrow x$ such that $x^k \in \operatorname{argmin} F^{\nu^k}$ for some subsequence $\{\nu^k\}_{k=1}^\infty$, then $x \in \operatorname{argmin} F$ and $\inf F^{\nu^k} \rightarrow \inf F$. In particular, if there is a compact set C such that $\operatorname{argmin} F^\nu \cap C \neq \emptyset$ for all ν , then $\inf F^\nu \rightarrow \inf F$.

Recall that a function g is called *lower semicontinuous* (lsc) if for every x

$$\liminf_{y \rightarrow x} g(y) \geq g(x).$$

Theorem 2 (Artstein and Wets (1995)) *Let $\xi_1, \xi_2, \dots$ be a sequence of i.i.d P -distributed drawings from Ξ and let*

$$P^\nu = \sum_{i=1}^{\nu} \frac{1}{\nu} \delta_{\xi_i}.$$

If

1. $f(x, \xi) : \mathbb{R}^n \times \Xi \rightarrow (-\infty, \infty]$ *is measurable on $\mathbb{R}^n \times \Xi$, and $f(\cdot, \xi)$ for ξ fixed is lsc in x ,*
2. *for each $x_0 \in \mathbb{R}^n$ there exists an open set $N \ni x_0$ and an integrable function $\alpha(\xi) : \Xi \rightarrow (-\infty, \infty)$, such that for almost all $\xi \in \Xi$ the inequality $f(x, \xi) \geq \alpha(\xi)$ holds for all $x \in N$,*

then the functions $E^{P^\nu} f$ almost surely epi-converge to $E^P f$.

The following is a simplified version of the epi-convergence result in Pennanen and Koivu (2003), which is sufficient in the applications of this paper.

Theorem 3 (Pennanen and Koivu (2003)) *Let $P^\nu \rightarrow P^0$ and assume that f is lsc. If*

1. *for each $x \in \mathbb{R}^n$, there is an open set $N \ni x$ such that f is bounded from below on $N \times \Xi$,*
2. *for each $x \in \operatorname{dom} E^{P^0} f$, $f(x, \cdot)$ is P^0 -a.s. continuous and bounded,*

then the functions $E^{P^\nu} f$ both pointwise and epi-converge to $E^P f$.

Note that the conditions of Theorem 3 imply the conditions of Theorem 2.

3 Randomized quasi-monte carlo and weak convergence

A discrete approximation P^ν of P is usually generated as follows: In the scalar case, approximate the uniform distribution on $[0, 1]$ and transform each point with the inverse of the distribution function of the desired distribution. This is known as the *method of inversion*. The same idea works whenever $P = QG^{-1}$, where Q is the multivariate uniform distribution and G is Q -a.s. continuous, in other words, whenever

$$\xi = G(u),$$

where u is uniformly distributed in the unit cube $[0, 1]^d$, and $G : [0, 1]^d \rightarrow \Xi$ is almost everywhere continuous. This is based on the following very useful result from Billingsley (1999) where U is any metric space with Borel algebra $\mathcal{B}$.

Theorem 4 (Billingsley) *Let $G : (U, \mathcal{B}) \rightarrow (\Xi, \Sigma)$ be a measurable function and Q a probability distribution on $(U, \mathcal{B})$. Then $QG^{-1}(A) := Q(G^{-1}A)$ defines a probability measure on (Ξ, Σ) , and if G is Q -a.s. continuous, then*

$$Q^\nu \rightarrow Q \implies Q^\nu G^{-1} \rightarrow QG^{-1}.$$

Given a Q -a.s. continuous G and a discrete approximation $Q^\nu = \sum_{i=1}^\nu p_i^\nu \delta_{u_i^\nu}$ of Q , Theorem 4 says that the discrete measures

$$P^\nu := Q^\nu G^{-1} = \sum_{i=1}^\nu p_i^\nu \delta_{G(u_i^\nu)}$$

converge weakly to $P = QG^{-1}$ whenever $Q^\nu \rightarrow Q$. It is then natural to try to choose discrete approximations Q^ν which are as close as possible to the uniform distribution Q . Quasi-monte carlo methods are designed to do exactly this; see the books of Niederreiter (1992) and Sloan and Joe (1994). Much of this theory has evolved around the following notion of distance from Q .

Definition 5 *The star-discrepancy of a point set $\mathcal{U}_\nu = \{u_1, \dots, u_\nu\} \subset [0, 1]^d$ is defined as*

$$D^*(\mathcal{U}_\nu) = \sup_{C \in \mathcal{C}_0} |Q^\nu(C) - Q(C)|, \quad (2)$$

where

$$Q^\nu = \sum_{i=1}^\nu \frac{1}{\nu} \delta_{u_i},$$

and $\mathcal{C}_0$ is the set of rectangles $C \subset [0, 1]^d$ with $0 \in C$.

The following is a direct consequence of Corollary 11 in Lucchetti et al. (1994).

Proposition 6 *For each ν , let $\mathcal{U}_\nu^\nu = \{u_1^\nu, \dots, u_\nu^\nu\}$ be point sets in the unit cube. The measures*

$$Q^\nu = \sum_{i=1}^\nu \frac{1}{\nu} \delta_{u_i^\nu}$$

converge weakly to the uniform distribution if and only if $D^(\mathcal{U}_\nu^\nu) \rightarrow 0$.*

Thus, if we can find point sets whose star-discrepancy approaches zero as $\nu \nearrow \infty$, we obtain weakly convergent discrete approximations of the uniform distribution. If $P = QG^{-1}$, we can then use the method of inversion to get weakly convergent discretizations of P .

In the literature of numerical integration, many methods have been developed for producing infinite sequences, which satisfy the property, that $D^*(\mathcal{U}_\nu) = O(\nu^{-1}(\ln \nu)^d)$, for all ν . Such sequences are called *low-discrepancy sequences*. The main constructions of low discrepancy sequences are due to Halton (1960), Sobol' (1967), Faure (1982) and Niederreiter (1988). The last three methods fall in the general class of (t, s) -sequences; see Niederreiter (1992). If it is not required that ν points of a $(\nu + 1)$ -point quadrature are the points of the ν -point quadrature, it is possible to obtain more accurate quadratures called *low discrepancy point sets*, which satisfy $D^*(\mathcal{U}_\nu) = O(\nu^{-1}(\ln \nu)^{d-1})$. Examples of low discrepancy point sets include Hammersley point sets (Hammersley, 1960), which are easily obtained from the Halton sequence and so called (t, m, s) -nets, which are obtained by using certain parts of the points in (t, s) -sequences; see (Niederreiter, 1992, Chapter 4). Another general family of methods for generating point sets with low discrepancy are lattice rules, which are designed to take advantage of additional regularity properties of integrands; see for example Niederreiter (1992), Sloan and Joe (1994) and L'Ecuyer and Lemieux (2000).

To enable practical error estimation for QMC methods a number of randomization techniques have been proposed in the literature; see L'Ecuyer and Lemieux (2002) for an excellent survey. An easy way of randomizing any QMC point set without destroying its regular structure, suggested by Cranley and Patterson (1976), is to shift it randomly, modulo 1, with respect to all of the coordinates.

Let $\mathcal{U}_\nu = \{u_1, \dots, u_\nu\} \subset [0, 1]^d$ be a low discrepancy point set in a d -dimensional unit hypercube. Generate a point u uniformly distributed in $[0, 1]^d$ and replace every u_i in $\mathcal{U}_\nu$ with $\tilde{u}_i = (u_i + u) \bmod 1$, where $i = 1, \dots, \nu$. Now $\tilde{\mathcal{U}}_\nu = \{\tilde{u}_1, \dots, \tilde{u}_\nu\}$ is a randomized point set used to approximate $[0, 1]^d$ uniform distribution. This can be repeated m times, independently, with the same $\mathcal{U}_\nu$. We thus obtain m i.i.d copies of the random variable $E^{P^\nu} \varphi$, which we denote by $E^{P^1_\nu} \varphi, \dots, E^{P^m_\nu} \varphi$. Let $\hat{\sigma}^2 = \sum_{j=1}^m (E^{P^j_\nu} \varphi - \hat{\mu})^2 / (m - 1)$, where $\hat{\mu} = (E^{P^1_\nu} \varphi + \dots + E^{P^m_\nu} \varphi) / m$.

Proposition 7 (L'Ecuyer and Lemieux (2000))

$$E[E^{P^j_\nu} \varphi] = E^P \varphi \quad \text{and} \quad E[\hat{\sigma}^2] = \text{Var}[E^{P^j_\nu} \varphi].$$

Hence, $E^{P^j_\nu} \varphi$ is an unbiased estimator of $E^P \varphi$ and $\hat{\sigma}^2$ is an unbiased estimator of its variance. Proposition 7 holds for an arbitrary point set $\mathcal{U}_\nu$; see (L'Ecuyer and Lemieux, 2002, Tuffin, 1996).

In direct numerical integration, Monte Carlo methods achieve a convergence rate of $\nu^{-\frac{1}{2}}$; more precisely, in Monte Carlo, the standard deviation of the integration error is $\text{Std}(\varphi)\nu^{-\frac{1}{2}}$, where $\text{Std}(\varphi)$ is the standard deviation of φ . The following estimates the convergence speed for the variance of a randomized QMC estimator obtained from a low discrepancy sequence.

Theorem 8 (Tuffin (1996)) *For any low discrepancy sequence $\mathcal{U}_\nu \subset [0, 1]^d$ and almost everywhere continuous and bounded function φ over $[0, 1]^d$, we have*

$$\text{Var} \left(\frac{1}{\nu} \sum_{i=1}^{\nu} \varphi(\tilde{u}_i) \right) = O(\nu^{-2}(\ln \nu)^{2d}).$$

In Monte Carlo, the convergence speed is independent of the dimension of the space, whereas the above convergence speed depends on the dimension, so that the actual error estimates obtained in practice with RQMC may be greater than $Std(\varphi)\nu^{-\frac{1}{2}}$. In many practical applications, however, RQMC methods considerably improve the accuracy over MC. One explanation offered for the success of QMC and RQMC methods on high dimensional problems is that the integrands may have effective dimensions much smaller than d . Effective dimension is roughly the number of important dimensions of the problem, which account for most of the variability of the estimator; see Caffisch et al. (1997) and Wang and Fang (2002) for details. Asymptotically the variance reduction factor obtained with RQMC over MC is proportional to ν . The same effect can be observed in the test problems of Section 4, for sample variances of optimal values already with moderate values of ν .

It is well known, that for MC

$$\inf_{x \in \mathbb{R}^n} E [E^{P^\nu} f(x)] \geq E \left[\inf_{x \in \mathbb{R}^n} E^{P^\nu} f(x) \right],$$

i.e. $v^* \geq E[\bar{v}^*]$ where v^* denotes the optimal value of the true problem (SP). That is, $\bar{v}^*$ is a biased estimator of v^* . This property also holds for RQMC methods. The value $\bar{v}^*$ is called a valid statistical lower bound of the true optimal value v^* if $v^* \geq E[\bar{v}^*]$ and $\bar{v}^*$ epi-converges to v^* as $\nu \rightarrow \infty$; see e.g. Shapiro (2003).

For obtaining epi-converge of the sample approximations of stochastic programs generated via RQMC methods we need to show that RQMC methods generate weakly convergent probability measures.

Lemma 9 *Let $\mathcal{U}_\nu$ and $\tilde{\mathcal{U}}_\nu$ be low discrepancy and randomized low discrepancy point sets, respectively. Discrepancy of a randomized low discrepancy point set $D(\tilde{\mathcal{U}}_\nu)$ satisfies*

$$D(\tilde{\mathcal{U}}_\nu) \leq 2^{2d} D^*(\mathcal{U}_\nu).$$

If $D^(\mathcal{U}_\nu) \rightarrow 0$, the measures*

$$Q^\nu = \sum_{i=1}^{\nu} \frac{1}{\nu} \delta_{\tilde{u}_i^\nu}$$

converge weakly to the uniform distribution.

Proof. From Niederreiter (1992) we get

$$D^*(\mathcal{U}_\nu) \leq D(\mathcal{U}_\nu) \leq 2^d D^*(\mathcal{U}_\nu),$$

where $D(\mathcal{U}_\nu)$ is a discrepancy measure. Tuffin (1996) showed that

$$D(\tilde{\mathcal{U}}_\nu) \leq 2^d D(\mathcal{U}_\nu),$$

which yields

$$D(\tilde{\mathcal{U}}_\nu) \leq 2^{2d} D^*(\mathcal{U}_\nu).$$

The weak convergence of the probability measures $Q^\nu = \sum_{i=1}^{\nu} \frac{1}{\nu} \delta_{\tilde{u}_i^\nu}$ follows from Proposition 6 by noting that $D^*(\mathcal{U}_\nu) \rightarrow 0$. $\square$

Hence, we can use the results of Theorem 3 for obtaining epi-convergence of $E^{P^\nu} f$ to $E^P f$. In sample approximations of stochastic programs a natural goal is to try to generate the samples so that the bias $v^* - E[\bar{v}^*]$ and the sample variance of the optimal values are as small as possible. In the next Section we use RQMC methods as variance reduction techniques alone and in combination with other variance reduction techniques to improve the accuracy of sample approximations with respect to MC in various test problems.

4 Numerical tests

In the numerical tests we compare MC with variance reduction techniques: Antithetic Variates (AV), Latin Hypercube sampling (LH), randomized Lattice Rules (LR), Sobol (SOB), Faure (FAU), Hammersley (HAM), Niederreiter (NIE) and Halton (HAL) point sets in discretization of five portfolio optimization problems. We will also test the efficiency of the best performing RQMC methods in combination with AV, namely Sobol sequence (SOB+AV) and lattice rules (LR+AV). For the MC method and randomization of the QMC point sets we use the Mersenne Twister generator (MT19937) by Matsumoto and Nishimura (1998). The LIBSEQ¹ library based on Friedel and Keller (2002) is used for Latin Hypercube sampling. Rank-1 lattice rules are used to generate the lattice point sets; see e.g. L'Ecuyer and Lemieux (2000)². Our implementation of the Sobol sequence is based on the implementation in Press et al. (1992). For Niederreiter sequence the routine in GSL (Gnu Scientific Library) is used. Routines by Fox (1986) are used for Faure and Halton sequences and the Hammersley point sets are easily obtained from the Halton sequence; see Hammersley (1960).

We consider one-stage problems with $\nu = 2^i$ scenarios, where $i = 5, \dots, 14$. For every i we generate 250 independent discretizations, solve the resulting problems and record the obtained optimum value and other relevant statistics. The same procedure is repeated for each test problem, except in Section 4.2.1, where the random variable is one-dimensional and $i = 5, \dots, 9$.

The test problems are divided into two categories. In Section 4.1 we consider problems without implicit constraints, i.e. $\text{dom } E^P f$ is known and does not depend on P . In Section 4.2 we consider problems with implicit constraints, i.e. $\text{dom } E^P f$ may not be known and may depend on P .

4.1 Problems without implicit constraints

4.1.1 Mean-variance portfolio optimization

We start the numerical tests with a model which can be solved exactly. Of course, sample approximations are unnecessary in such cases but here we get to compare the approximate solutions with

¹www.multires.caltech.edu/software/libseq

²The parameters required by the method were provided by Professor L'Ecuyer.

the exact one. Consider the mean-variance model

$$\begin{aligned}
& \underset{x \in \mathbb{R}^n}{\text{minimize}} && E^{P^0}(r \cdot x - \bar{r} \cdot x)^2 \\
& \text{subject to} && \bar{r} \cdot x \geq w, \\
& && \sum_{i=1}^n x_i \leq 1, \\
& && x \in C,
\end{aligned} \tag{MP}$$

where $x = (x_1, \dots, x_n)$ is a portfolio of assets, $r = (r_1, \dots, r_n)$ is the vector of returns, $r \cdot x = \sum_{i=1}^n r_i x_i$ is the terminal wealth, w is the required level of expected wealth and C is the set of feasible portfolios. The components of the return vector r are random variables with joint distribution P^0 and expectation $\bar{r}$. As is well-known, the expectation in (MP) can be computed explicitly as

$$E^{P^0}(r \cdot x - \bar{r} \cdot x)^2 = E^{P^0}[(r - \bar{r}) \cdot x]^2 = E^{P^0}[x \cdot (r - \bar{r})(r - \bar{r})^T x] = x \cdot V x,$$

where $V = E^{P^0}[(r - \bar{r})(r - \bar{r})^T]$ is the variance matrix of r . If V and $\bar{r}$ are known, (MP) can then be solved without discretization with standard solvers yielding the optimal value and optimal solution.

To test the performance of the proposed variance reduction techniques, we approximate problem (MP) by the discretizations

$$\begin{aligned}
& \underset{x \in \mathbb{R}^n}{\text{minimize}} && \sum_{i=1}^{\nu} p_i^{\nu} (r_i^{\nu} \cdot x - \bar{r} \cdot x)^2 \\
& \text{subject to} && \bar{r} \cdot x \geq w, \\
& && \sum_{i=1}^n x_i \leq 1, \\
& && x \in C.
\end{aligned} \tag{MP}^{\nu}$$

Under mild conditions, convergence of optimal values and solutions can be guaranteed. The proof of the following Proposition can be found in Pennanen and Koivu (2003).

Proposition 10 (Pennanen and Koivu (2003)) *Assume that $\text{supp } P^0$ is bounded, C is closed, and that the measures*

$$P^{\nu} = \sum_{i=1}^{\nu} p_i^{\nu} \delta_{r_i^{\nu}}$$

converge weakly to P^0 and satisfy $\text{supp } P^{\nu} \subset \text{supp } P^0$. If the feasible set is bounded, then the optimal values of $(MP)^{\nu}$ converge to that of (MP) and the cluster points of the solutions of $(MP)^{\nu}$ are solutions of (MP).

In our test, the number of assets $n = 10$ and

$$r = \bar{r} + 12L(u - \frac{1}{2}e),$$

where u is uniformly distributed in the 10-dimensional unit cube, L is a 10×10 matrix and e is a vector of ones. Then $\text{supp } P^0$ is bounded, r has mean $\bar{r}$ and variance $V = LL^T$. We chose $C = \mathbb{R}_+^n$, which means that “short selling” is prohibited. With our choices of $\bar{r}$ and V , the optimal value in the original problem (MP) is 1.9221.

The numerical test results are displayed in Table 1, where $\hat{\mu}$ and $\hat{\sigma}$ denote the sample mean and standard deviation computed from 250 optimal values of (MP^ν) for different values of ν . The value $vr = \hat{\sigma}_{MC}^2 / \hat{\sigma}_q^2$, denotes the variance reduction factors for optimal values obtained with sampling method q with respect to the variance of MC, for all the considered methods and reported values of ν . The best performing methods are LR and Sobol, Halton and Niederreiter sequences, with variance reduction factors increasing with ν . These methods clearly outperform MC, AV and LH sampling. The results with AV are presented to point out the fact, that the use of AV doubles the variance with respect to MC because the objective function is quadratic and it is well known, that AV reduces the variance compared to MC only when the integrand is a monotonically increasing function of the random variables; see (Bratley et al., 1987). Figure 1 shows the sample mean and 90% confidence intervals for the optimal values obtained with LR and MC. Lattice rules produce much tighter confidence intervals and reduces the sample bias for the optimal value, compared to MC.

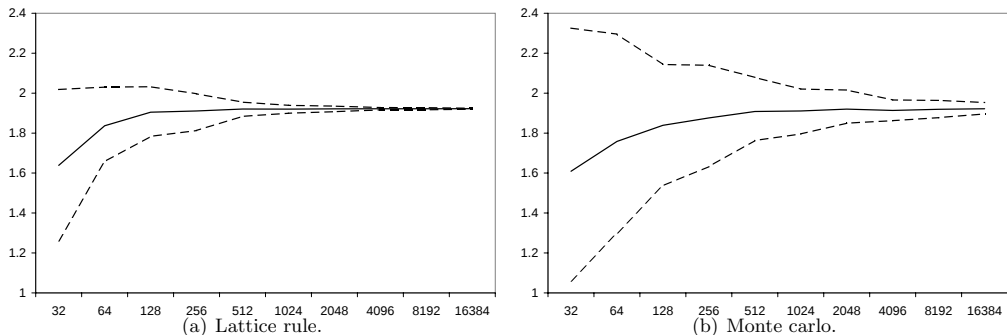


Figure 1: Mean and 90% confidence interval for the markowitz problem.

Table 1: Statistics for MP^ν as a function of ν .

ν		MC	AV	LH	LR	SOB	FAU	HAM	NIE	HAL
32	$\hat{\mu}$	1.609	1.413	1.668	1.639	1.662	1.461	1.567	1.708	1.704
	$\hat{\sigma}$	5.00E-1	6.99E-1	4.06E-1	3.68E-1	3.83E-1	5.76E-1	4.92E-1	3.78E-1	3.31E-1
	vr	1.0	0.5	1.5	1.8	1.7	0.8	1.0	1.7	2.3
64	$\hat{\mu}$	1.758	1.689	1.752	1.837	1.855	1.742	1.796	1.840	1.818
	$\hat{\sigma}$	3.38E-1	4.61E-1	2.98E-1	1.48E-1	2.05E-1	2.56E-1	2.20E-1	1.82E-1	1.72E-1
	vr	1.0	0.5	1.3	5.2	2.7	1.7	2.4	3.4	3.9
128	$\hat{\mu}$	1.839	1.803	1.846	1.905	1.875	1.888	1.890	1.883	1.889
	$\hat{\sigma}$	2.08E-1	2.93E-1	1.74E-1	7.73E-2	1.20E-1	1.06E-1	1.22E-1	1.16E-1	9.43E-2
	vr	1.0	0.5	1.4	7.3	3.0	3.9	2.9	3.2	4.9
256	$\hat{\mu}$	1.876	1.820	1.887	1.911	1.906	1.904	1.909	1.916	1.913
	$\hat{\sigma}$	1.53E-1	2.22E-1	1.15E-1	5.78E-2	6.93E-2	5.78E-2	6.33E-2	6.19E-2	5.52E-2
	vr	1.0	0.5	1.8	7.1	4.9	7.1	5.9	6.1	7.7
512	$\hat{\mu}$	1.908	1.877	1.899	1.921	1.920	1.909	1.916	1.914	1.914
	$\hat{\sigma}$	1.04E-1	1.38E-1	8.48E-2	2.15E-2	3.45E-2	4.48E-2	3.75E-2	3.53E-2	2.84E-2
	vr	1.0	0.6	1.5	23.5	9.1	5.4	7.8	8.7	13.5
1024	$\hat{\mu}$	1.911	1.902	1.917	1.920	1.920	1.920	1.921	1.923	1.922
	$\hat{\sigma}$	7.09E-2	1.02E-1	5.87E-2	1.24E-2	1.81E-2	2.10E-2	1.97E-2	1.89E-2	1.67E-2
	vr	1.0	0.5	1.5	32.7	15.3	11.4	12.9	14.0	18.1
2048	$\hat{\mu}$	1.920	1.906	1.920	1.922	1.921	1.921	1.923	1.922	1.922
	$\hat{\sigma}$	4.98E-2	7.02E-2	4.15E-2	8.17E-3	8.87E-3	1.32E-2	1.01E-2	1.02E-2	9.34E-3
	vr	1.0	0.5	1.4	37.1	31.5	14.2	24.1	23.8	28.4
4096	$\hat{\mu}$	1.914	1.917	1.920	1.922	1.922	1.922	1.922	1.922	1.922
	$\hat{\sigma}$	3.29E-2	5.25E-2	2.89E-2	3.65E-3	4.69E-3	6.97E-3	5.29E-3	6.43E-3	5.90E-3
	vr	1.0	0.4	1.3	80.9	49.0	22.2	38.6	26.1	31.1
8192	$\hat{\mu}$	1.919	1.921	1.923	1.922	1.922	1.922	1.922	1.922	1.922
	$\hat{\sigma}$	2.61E-2	3.57E-2	2.02E-2	3.33E-3	3.32E-3	3.55E-3	2.83E-3	3.59E-3	2.94E-3
	vr	1.0	0.5	1.7	61.4	61.7	54.1	85.3	52.8	78.8
16384	$\hat{\mu}$	1.922	1.920	1.923	1.922	1.922	1.922	1.922	1.922	1.922
	$\hat{\sigma}$	1.79E-2	2.46E-2	1.42E-2	1.49E-3	1.37E-3	1.84E-3	1.53E-3	1.18E-3	1.78E-3
	vr	1.0	0.5	1.6	145	170	94.5	137	229	101

4.1.2 Utility maximization

Consider the problem

$$\begin{aligned}
& \underset{x \in \mathbb{R}^n}{\text{maximize}} && E^{P^0} u(r \cdot x) \\
& \text{subject to} && \sum_{i=1}^n x_i \leq 1, \\
& && x \in C.
\end{aligned} \tag{UP}$$

Here x and C are as in the previous example, u measures the utility from terminal wealth, and the components of the return vector r are nonnegative random variables with joint distribution P^0 .

In general, (UP) cannot be solved analytically, so we consider the discretizations

$$\begin{aligned}
& \underset{x \in \mathbb{R}^n}{\text{maximize}} && \sum_{i=1}^{\nu} p_i^{\nu} u(r_i^{\nu} \cdot x) \\
& \text{subject to} && \sum_{i=1}^n x_i \leq 1, \\
& && x \in C.
\end{aligned} \tag{UP}^{\nu}$$

The same type of problem was analyzed in Pennanen and Koivu (2003), so we can use their Proposition to show the epi-convergence of $(UP)^{\nu}$ to (UP) .

Proposition 11 (Pennanen and Koivu (2003)) *Assume $\text{supp } P^0 \subset \mathbb{R}_+^n$, u is continuous and bounded on $\mathbb{R}_+$, C is closed and contained in $\mathbb{R}_+^n$ (short selling is not allowed) and that the measures*

$$P^{\nu} = \sum_{i=1}^{\nu} p_i^{\nu} \delta_{r_i^{\nu}}$$

converge weakly to P^0 and satisfy $\text{supp } P^{\nu} \subset \mathbb{R}_+^n$. Then the optimal values of $(UP)^{\nu}$ converge to that of (UP) and the cluster points of the solutions of $(UP)^{\nu}$ are solutions of (UP) .

In the test, the number of assets $n = 10$, r is log-normally distributed, $u(w) = -\exp(-w)$ and $C = \mathbb{R}_+^n$. Table 2 summarizes the test results. AV reduces the bias and variance of the optimal values significantly compared to MC. Among the RQMC methods LR perform the best, with all the other quadratures, except Faure sequence, performing almost as well. Since the use of AV reduced the variance of optimal values considerably, we tested them in combination with LR and Sobol sequence, see Table 3. The combination of these methods produce the most significant variance reduction factors compared to MC. Figure 2 displays the sample mean and 90% confidence interval for the optimal values obtained with LR and MC. Again the variance reduction factors with RQMC methods increase almost linearly with ν .

Table 2: Statistics UP^ν as a function of ν .

ν		MC	AV	LH	LR	SOB	FAU	HAM	HAL	NIE
32	$\hat{\mu}$	-298.543	-312.304	-312.014	-312.372	-310.990	-306.122	-308.435	-306.970	-310.261
	$\hat{\sigma}$	1.94E+1	3.84E+0	7.10E+0	5.18E+0	6.90E+0	1.20E+1	5.89E+0	5.82E+0	5.85E+0
	vr	1.0	25.6	7.5	14.0	7.9	2.6	10.8	11.1	11.0
64	$\hat{\mu}$	-304.718	-313.774	-313.855	-314.307	-313.418	-311.252	-312.172	-311.824	-313.258
	$\hat{\sigma}$	1.26E+1	2.67E+0	3.55E+0	2.77E+0	3.78E+0	7.26E+0	3.56E+0	3.33E+0	3.13E+0
	vr	1.0	22.3	12.6	20.8	11.2	3.0	12.6	14.4	16.3
128	$\hat{\mu}$	-309.980	-314.782	-314.865	-314.945	-314.853	-314.599	-314.174	-313.945	-314.681
	$\hat{\sigma}$	9.07E+0	1.72E+0	1.92E+0	1.38E+0	2.02E+0	2.88E+0	1.81E+0	1.69E+0	1.54E+0
	vr	1.0	27.7	22.4	43.1	20.2	9.9	25.2	29.0	34.8
256	$\hat{\mu}$	-312.546	-315.173	-315.171	-315.303	-315.201	-315.295	-315.136	-314.915	-315.236
	$\hat{\sigma}$	6.78E+0	1.09E+0	1.04E+0	7.66E-1	9.55E-1	1.86E+0	9.40E-1	8.99E-1	8.28E-1
	vr	1.0	38.9	42.4	78.4	50.4	13.2	52.0	56.9	67.1
512	$\hat{\mu}$	-313.600	-315.151	-315.368	-315.448	-315.475	-315.289	-315.350	-315.357	-315.389
	$\hat{\sigma}$	4.40E+0	9.27E-1	5.33E-1	4.17E-1	4.77E-1	1.23E+0	4.78E-1	4.79E-1	4.07E-1
	vr	1.0	22.5	68.1	111	84.9	12.8	84.6	84.2	116
1024	$\hat{\mu}$	-314.658	-315.443	-315.414	-315.489	-315.502	-315.411	-315.440	-315.471	-315.457
	$\hat{\sigma}$	3.07E+0	6.18E-1	3.06E-1	2.02E-1	2.60E-1	5.66E-1	2.68E-1	2.44E-1	2.11E-1
	vr	1.0	24.7	100	232	139	29.4	131	158	211
2048	$\hat{\mu}$	-314.858	-315.440	-315.465	-315.505	-315.512	-315.473	-315.492	-315.486	-315.498
	$\hat{\sigma}$	2.15E+0	4.58E-1	1.74E-1	1.07E-1	1.37E-1	3.07E-1	1.32E-1	1.40E-1	1.10E-1
	vr	1.0	22.1	153	404	245	49.0	264	235	381
4096	$\hat{\mu}$	-315.270	-315.475	-315.481	-315.506	-315.506	-315.482	-315.509	-315.498	-315.496
	$\hat{\sigma}$	1.70E+0	2.94E-1	1.16E-1	5.89E-2	7.61E-2	1.57E-1	7.60E-2	7.01E-2	6.13E-2
	vr	1.0	33.3	216	832	498	118	500	587	767
8192	$\hat{\mu}$	-315.421	-315.483	-315.495	-315.505	-315.502	-315.493	-315.508	-315.503	-315.504
	$\hat{\sigma}$	1.23E+0	2.02E-1	6.62E-2	3.16E-2	3.78E-2	9.33E-2	4.05E-2	4.25E-2	3.20E-2
	vr	1.0	37.5	348	1527	1062	175	926	844	1482
16384	$\hat{\mu}$	-315.381	-315.495	-315.495	-315.504	-315.505	-315.511	-315.503	-315.506	-315.504
	$\hat{\sigma}$	8.07E-1	1.58E-1	4.89E-2	2.00E-2	1.91E-2	4.66E-2	2.08E-2	2.11E-2	1.90E-2
	vr	1.0	25.9	272	1630	1791	299	1510	1462	1804

Table 3: Statistics for UP^ν as a function of ν , Lattice rule and Sobol with AV.

ν		LR+AV	SOB+AV	ν		LR+AV	SOB+AV
32	$\hat{\mu}$	-313.078	-312.744	1024	$\hat{\mu}$	-315.485	-315.488
	$\hat{\sigma}$	1.61E+0	2.26E+0		$\hat{\sigma}$	1.13E-1	1.33E-1
	vr	145	74		vr	737	528
64	$\hat{\mu}$	-314.278	-314.295	2048	$\hat{\mu}$	-315.498	-315.496
	$\hat{\sigma}$	1.15E+0	1.31E+0		$\hat{\sigma}$	5.94E-2	7.43E-2
	vr	121	93		vr	1312	837
128	$\hat{\mu}$	-314.962	-315.040	4096	$\hat{\mu}$	-315.499	-315.503
	$\hat{\sigma}$	5.79E-1	6.35E-1		$\hat{\sigma}$	3.91E-2	4.71E-2
	vr	246	204		vr	1886	1304
256	$\hat{\mu}$	-315.324	-315.317	8192	$\hat{\mu}$	-315.505	-315.507
	$\hat{\sigma}$	3.16E-1	3.75E-1		$\hat{\sigma}$	2.75E-2	2.45E-2
	vr	460	327		vr	2007	2532
512	$\hat{\mu}$	-315.415	-315.429	16384	$\hat{\mu}$	-315.504	-315.505
	$\hat{\sigma}$	2.37E-1	2.16E-1		$\hat{\sigma}$	1.41E-2	1.50E-2
	vr	343	416		vr	3295	2889

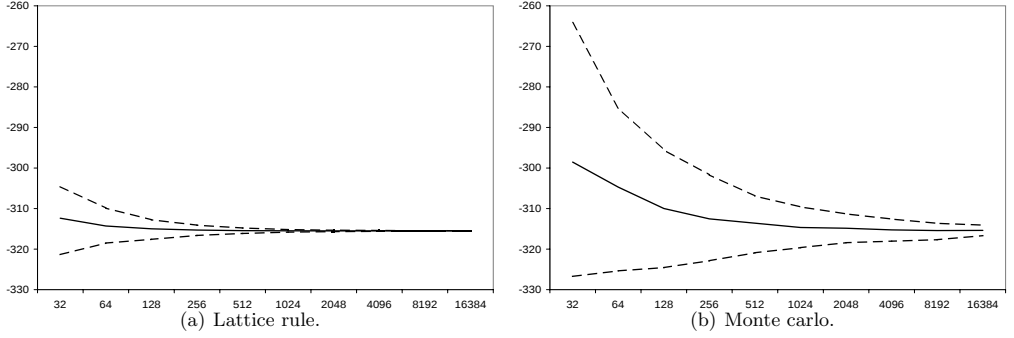


Figure 2: Mean and 90% confidence interval for the utility maximization problem.

4.1.3 Hedging with contingent claims

Assume that a company's operating revenue at time $t = 0, \dots, T$ can be expressed as a function $\pi_t(\xi)$, where $\xi = (\xi_0, \dots, \xi_T)$ is a stochastic process with joint distribution P^0 . The company wishes to hedge its operating revenue against unfavorable outcomes of ξ using contingent claims with pay-outs $F_t(\xi)$. Let $\theta_+ = (\theta_1, \dots, \theta_J)$ and $\theta_- = (\theta_1, \dots, \theta_J)$ denote the amounts of contingent claims bought and sold with prices P_a and P_b , respectively, at time $t = 0$. The company faces the hedging problem

$$\begin{aligned}
 & \underset{\theta_+, \theta_-}{\text{maximize}} && E^{P^0} \left[u(\pi_0(\xi_0) - tc_a \cdot \theta_+ - tc_b \cdot \theta_-) + \sum_{t=1}^T u(\pi_t(\xi) + F_t(\xi) \cdot (\theta_+ - \theta_-)) \right] && (HP) \\
 & \text{subject to} && P_a \cdot \theta_+ - P_b \cdot \theta_- \leq \pi_0(\xi_0) \\
 & && \theta_+, \theta_- \geq 0,
 \end{aligned}$$

where u is a utility function, $\pi_0(\xi_0)$ is fixed and tc_a and tc_b denote the transaction costs of bought and sold assets, respectively. Since (HP) is impossible to solve analytically we consider the discretizations

$$\begin{aligned}
 & \underset{\theta_+, \theta_-}{\text{maximize}} && \sum_{i=1}^{\nu} p_i^{\nu} \left[u(\pi_0(\xi_0) - tc_a \cdot \theta_+ - tc_b \cdot \theta_-) + \sum_{t=1}^T u(\pi_t(\xi_i^{\nu}) + F_t(\xi_i^{\nu}) \cdot (\theta_+ - \theta_-)) \right] && (HP^{\nu}) \\
 & \text{subject to} && P_a \cdot \theta_+ - P_b \cdot \theta_- \leq \pi_0(\xi_0) \\
 & && \theta_+, \theta_- \geq 0.
 \end{aligned}$$

Proposition 12 Assume that u is continuous and concave, the first moments of the random variables $\pi_t(\xi)$ and $F_t(\xi)$ exist and

$$P^\nu = \sum_{i=1}^{\nu} p_i^\nu \delta_{(\xi_{t,i}^\nu)}^T$$

is a sequence of empirical measures. Then with probability one the optimal values of (HP^ν) converge to that of (HP) and the cluster points of the solutions of (HP^ν) are solutions of (HP) .

Proof. This can be written as (SP) with $x = (\theta_+, \theta_-)$ and

$$f(x, \xi) = -u(\pi_0(\xi_0) - tc_a \cdot \theta_+ - tc_b \cdot \theta_-) - \sum_{t=1}^T u(\pi_t(\xi) + F_t(\xi) \cdot (\theta_+ - \theta_-)) + \delta_{C'}(\theta_+, \theta_-),$$

where $C' = \{(\theta_+, \theta_-) \in \mathbb{R}_+^n \mid P_a \cdot \theta_+ - P_b \cdot \theta_- \leq \pi_0(\xi_0)\}$. By Theorem 1 it suffices to verify the conditions of Theorem 2. Since u is continuous and $\pi_t(\xi)$ and $F_t(\xi)$ are measurable f is measurable and lsc in x . To verify condition 2 let (x^0, ξ^0) be such that $f(x^0, \xi^0) < \infty$. By convexity of $-u$ we have

$$f(x, \xi) \geq f(x^0, \xi^0) + \gamma_0^0(tc_a \cdot (\theta_+^0 - \theta_+) + tc_b \cdot (\theta_-^0 - \theta_-)) + \sum_{t=1}^T \gamma_t^0(\pi_t(\xi) + F_t(\xi) \cdot (\theta_+ - \theta_-) - \pi(\xi^0) - F_t(\xi^0) \cdot (\theta_+^0 - \theta_-^0)),$$

where γ_t^0 denote subgradients of $-u$. Using the Cauchy-Schwarz inequality we get that for any bounded $N \ni x^0$

$$f(x, \xi) \geq \psi^0 - \gamma_0^0(tc_a \cdot \theta_+ + tc_b \cdot \theta_-) + \sum_{t=1}^T \gamma_t^0(\pi_t(\xi) + F_t(\xi) \cdot (\theta_+ - \theta_-)) \geq a + \sum_{t=1}^T \gamma_t^0 \pi_t(\xi) + b \sum_{t=1}^T |F_t(\xi)|, \forall x \in N,$$

where a and b are constants. Since it was assumed that the first moments of the random variables $\pi_t(\xi)$ and $F_t(\xi)$ exist condition 2 is satisfied. $\square$

By assuming that $\pi_t(\xi)$ and $F_t(\xi)$ are almost everywhere continuous and bounded, the conditions of Theorem 3 would be satisfied and we would obtain epi-convergence for RQMC methods. However, it is interesting to study the behavior of RQMC methods in this problem numerically. In the test $u(w) = -\exp(-w)$, $T = 12$, ξ_0 is deterministic and ξ_t is a three dimensional log-normally distributed random variable, which means that the dimension of the probability space, $d = 36$. The stochastic factors affecting the company's operating revenue are the Euro-U.S. dollar (USD), Norwegian krone-USD exchange rates and the USD price of zinc. The set of contingent claims consists of zero coupon bonds and futures contracts for the underlying stochastic factors, with maturities $1, 2, \dots, T$ months.

The results are displayed in Table 4. The use of AV increased the variance of optimal values compared to MC, because the profit function $\pi_t(\xi)$ is not a monotonically increasing function of the

random variables, these results are not reported. The results for Niederreiter sequence are missing, because in our implementation of the sequence the maximum dimension for the probability space is 12. Again RQMC methods, except Faure sequence, clearly beat MC and LR seem to perform slightly better than the other RQMC methods. In this problem Faure sequence performs poorly and even loses to MC for low values of ν . Latin hypercube sampling substantially improves the performance over MC. Compared to LH, LR and Sobol sequence produce more accurate estimates for optimal values, for all values of ν . Figure 3 displays the sample mean and 90% confidence interval for the optimal values obtained with LR and MC, which also shows that LR clearly outperform MC. The sample means of optimal values seem to converge toward a common value with all the methods, even though we were able to proof the epi-convergence only for MC.

Table 4: Statistics for HP as a function of ν .

ν		MC	LH	LR	SOB	FAU	HAM	HAL
32	$\hat{\mu}$	-210.165	-214.596	-216.088	-215.000	-202.903	-205.388	-203.808
	$\hat{\sigma}$	1.77E+1	8.87E+0	6.69E+0	8.17E+0	2.09E+1	1.26E+1	1.31E+1
	vr	1.0	4.0	7.0	4.7	0.7	2.0	1.8
64	$\hat{\mu}$	-213.241	-216.771	-217.430	-217.106	-207.698	-215.237	-215.076
	$\hat{\sigma}$	1.22E+1	4.69E+0	3.87E+0	3.82E+0	1.71E+1	5.11E+0	5.04E+0
	vr	1.0	6.8	10.0	10.2	0.5	5.7	5.9
128	$\hat{\mu}$	-216.275	-217.842	-218.051	-217.945	-211.040	-217.492	-217.844
	$\hat{\sigma}$	7.35E+0	3.04E+0	2.38E+0	2.57E+0	1.42E+1	2.52E+0	3.10E+0
	vr	1.0	5.8	9.6	8.2	0.3	8.5	5.6
256	$\hat{\mu}$	-217.459	-218.298	-218.332	-218.198	-211.715	-218.202	-218.196
	$\hat{\sigma}$	5.28E+0	1.90E+0	1.19E+0	1.43E+0	1.16E+1	1.53E+0	1.47E+0
	vr	1.0	7.8	19.7	13.6	0.2	11.8	12.9
512	$\hat{\mu}$	-218.143	-218.461	-218.482	-218.433	-215.089	-218.372	-218.398
	$\hat{\sigma}$	4.06E+0	1.15E+0	8.14E-1	7.79E-1	7.46E+0	8.59E-1	9.04E-1
	vr	1.0	12.5	25	27.2	0.3	22.4	20.2
1024	$\hat{\mu}$	-218.101	-218.402	-218.496	-218.455	-218.264	-218.455	-218.423
	$\hat{\sigma}$	2.60E+0	7.59E-1	4.42E-1	5.01E-1	2.19E+0	4.79E-1	5.77E-1
	vr	1.0	11.8	35	27	1.4	30	20
2048	$\hat{\mu}$	-218.272	-218.484	-218.505	-218.494	-218.212	-218.490	-218.501
	$\hat{\sigma}$	1.76E+0	4.81E-1	2.79E-1	2.82E-1	1.51E+0	2.98E-1	3.06E-1
	vr	1.0	13.4	40	39	1.4	35	33
4096	$\hat{\mu}$	-218.535	-218.484	-218.513	-218.496	-218.419	-218.512	-218.516
	$\hat{\sigma}$	1.41E+0	3.80E-1	1.26E-1	1.59E-1	6.31E-1	1.70E-1	1.86E-1
	vr	1.0	13.8	125	79	5	69	58
8192	$\hat{\mu}$	-218.553	-218.504	-218.508	-218.507	-218.450	-218.507	-218.509
	$\hat{\sigma}$	9.54E-1	2.46E-1	8.34E-2	8.79E-2	5.60E-1	8.89E-2	1.01E-1
	vr	1.0	15.1	131	118	3	115	90
16384	$\hat{\mu}$	-218.455	-218.505	-218.514	-218.510	-218.489	-218.508	-218.511
	$\hat{\sigma}$	6.72E-1	1.70E-1	6.63E-2	5.29E-2	4.84E-1	5.26E-2	6.27E-2
	vr	1.0	15.6	103	161	2	164	115

4.2 Problems with implicit constraints

In the remaining examples, the feasible regions depend on the probability measure. These problems do not fit the frameworks of Lucchetti and Wets (1993), Artstein and Wets (1994), Zervos (1999) or Shapiro (2000).

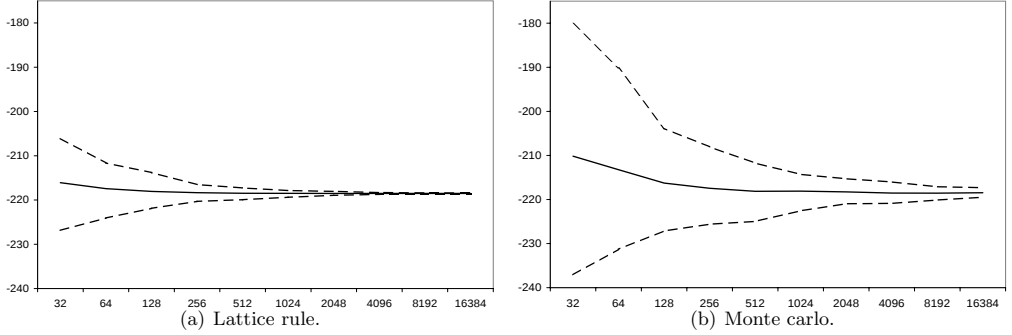


Figure 3: Mean and 90% confidence interval for the hedging problem.

4.2.1 Super-replication of contingent claims

Consider the problem

$$\begin{aligned}
 & \underset{V, \theta}{\text{minimize}} && V && (PP) \\
 & \text{subject to} && S_0 \cdot \theta \leq V, \\
 & && S \cdot \theta \geq F, \quad P^0\text{-a.s.} \\
 & && \theta \in C,
 \end{aligned}$$

where V is the wealth invested in a portfolio $\theta = (\theta_1, \dots, \theta_J)$ of assets that have prices $S_0 = (S_0^1, \dots, S_0^J)$ at the beginning and $S = (S^1, \dots, S^J)$ at the end of a holding period and F is a cash-flow at the end of the holding period. S and F are random variables with joint distribution P^0 . (PP) is a semi-infinite linear programming problem and, in general, impossible to solve analytically. Replacing P^0 by a discrete measure $P^\nu = \sum_{i=1}^\nu p_i^\nu \delta_{(S_i^\nu, F_i^\nu)}$ with $p_i^\nu > 0$, for all $i = 1, \dots, \nu$ yields

$$\begin{aligned}
 & \underset{V, \theta}{\text{minimize}} && V && (PP^\nu) \\
 & \text{subject to} && S_0 \cdot \theta \leq V, \\
 & && S_i^\nu \cdot \theta \geq F_i^\nu, \quad i = 1, \dots, \nu, \\
 & && \theta \in C,
 \end{aligned}$$

which is an LP problem for which many solvers are available.

Proposition 13 (Pennanen and Koivu (2003)) *Assume that the points $\{(S_i^\nu, F_i^\nu)\}_{i=1}^\nu$ are all contained in $\text{supp } P^0$ and that for some $\{p_i^\nu\}_{i=1}^\nu$, $\nu = 0, 1, 2, \dots$, with $p_i^\nu > 0$, for all $i = 1, \dots, \nu$, the measures*

$$P^\nu = \sum_{i=1}^\nu p_i^\nu \delta_{(S_i^\nu, F_i^\nu)}$$

converge weakly to P^0 . If the feasible set is bounded, then the optimal values of (PP^ν) converge to that of (PP) and the cluster points of the solutions of (PP^ν) are solutions of (PP) .

In our test, the set of assets consists of cash, SP500 index and 28 European call and put options on the index with maturity of 17 calendar days. The value of S is fully determined by the value of the index at the maturity which is assumed to be log-normally distributed. The cash-flow F is taken to be that of a call option with the same maturity but different strike than any other call included in S .

Since the random variable in this problem is one dimensional all the QMC methods produce identical discretizations. As a result we consider discretizations only with LR, AV and MC. Table 5 displays the test results. The use of AV does not improve the performance over MC. Lattice rules reduce the variance of optimal values considerably and with 256 scenarios the optimal values have converged. Figure 4 displays the average and 90% confidence interval for optimum values of (PP^ν) obtained with LR and MC, for each value of $\nu = 2^i, i = 5, 6, \dots, 9$. With LR the confidence interval is much tighter and the optimal value converges faster than with MC.

Table 5: Statistics for PP as a function of ν .

ν		MC	AV	LR
32	$\hat{\mu}$	19,598	17,320	28,417
	$\hat{\sigma}$	1,71E+1	2,37E+1	2,25E+0
	vr	1,0	0,5	57,4
64	$\hat{\mu}$	27,072	26,739	29,682
	$\hat{\sigma}$	6,27E+0	6,87E+0	1,74E+0
	vr	1,0	0,8	13,0
128	$\hat{\mu}$	29,844	30,261	31,287
	$\hat{\sigma}$	3,18E+0	2,12E+0	9,05E-1
	vr	1,0	2,2	12,3
256	$\hat{\mu}$	31,194	31,177	32,004
	$\hat{\sigma}$	1,39E+0	1,37E+0	0,00E+0
	vr	1,0	1,0	∞
512	$\hat{\mu}$	31,786	31,841	32,004
	$\hat{\sigma}$	7,44E-1	6,00E-1	0,00E+0
	vr	1,0	1,5	∞

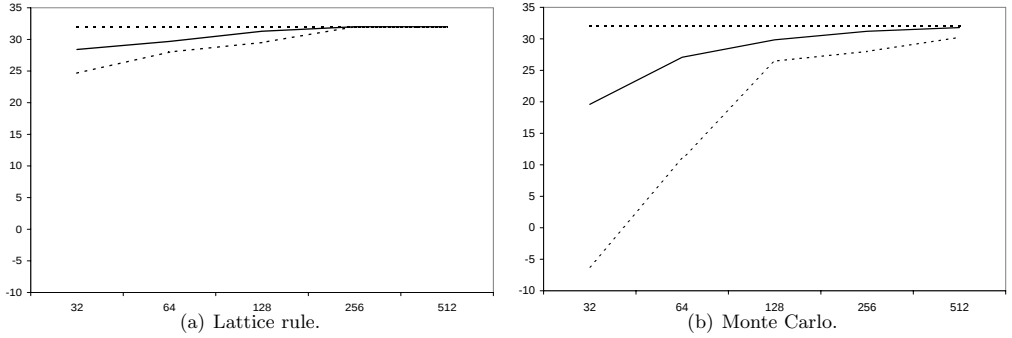


Figure 4: Mean and 90% confidence interval for the hedging problem.

4.2.2 Utility maximization with wealth constraint

Consider the following problem

$$\begin{aligned}
 & \underset{x \in \mathbb{R}^n}{\text{maximize}} && E^{P^0} u(r \cdot x) \\
 & \text{subject to} && \sum_{i=1}^n x_i \leq 1, \\
 & && x \in C \\
 & && r \cdot x \geq 0, \quad P^0\text{-a.s.}
 \end{aligned} \tag{WP}$$

This problem is a modification of the utility maximization problem of Section 4.1.2. Here $C \subset \mathbb{R}^n$, so short selling is allowed but we have added a constraint, which requires the final wealth to be almost surely non-negative. Here we are interested in studying how the short selling affects the behavior of the optimal values and solutions. The function u measures the utility from terminal wealth and the components of the return vector r are random variables with joint distribution P^0 . Discretization of (WP) yields

$$\begin{aligned}
 & \underset{x \in \mathbb{R}^n}{\text{maximize}} && \sum_{i=1}^{\nu} p_i^{\nu} u(r_i^{\nu} \cdot x) \\
 & \text{subject to} && \sum_{i=1}^n x_i \leq 1, \\
 & && x \in C, \\
 & && r_i^{\nu} \cdot x \geq 0, \quad i = 1, \dots, \nu.
 \end{aligned} \tag{WP^{\nu}}$$

Proposition 14 *Assume u is continuous, nondecreasing and bounded on $\mathbb{R}_+$, C is closed and that the measures*

$$P^{\nu} = \sum_{i=1}^{\nu} p_i^{\nu} \delta_{r_i^{\nu}}$$

converge weakly to P^0 and satisfy $\text{supp } P^{\nu} \subset \text{supp } P^0$. Then the optimal values of (WP $^{\nu}$) converge to that of (WP) and the cluster points of the solutions of (WP $^{\nu}$) are solutions of (WP).

Proof. This fits the format of (SP) with $\Xi = \text{supp } P^0$, $\xi = r$, and

$$f(x, r) = -u(r \cdot x) + \delta_{C_1}(x) + \delta_{C_2}(x, r),$$

where

$$C_1 = \left\{ x \in C \mid \sum_{i=1}^n x_i \leq 1 \right\}$$

and

$$C_2 = \{(x, r) \mid r \cdot x \geq 0\}.$$

We need to verify the conditions of Theorem 3. Since u is continuous, C_1 and C_2 are closed, f is lsc. Condition 1 follows from the fact that u is nondecreasing and bounded on $\mathbb{R}_+$. To verify condition 2 note that for each $x \in \text{dom } E^P f$ by continuity of the inner product the requirement $r \cdot x \geq 0, P^0$ -a.s. is equivalent to $r \cdot x \geq 0, \forall r \in \text{supp } P^0$. Since $\text{supp } P^\nu \subset \text{supp } P^0$ Condition 2 follows from the boundedness and continuity of u on $\mathbb{R}_+$. $\square$

In the test $\text{supp } r = \mathbb{R}_+^n$, which together with the wealth constraint implies that, $\text{dom } E^P f$ is $\mathbb{R}_+^n$, so this problem differs from the utility maximization problem of Section 4.1.2 only in finite samples. The numerical test results are presented in Tables 6 and 7. The results are similar to those of Section 4.1.2. The use of antithetic variates reduces the variance considerably. When no other variance reduction technique is used, LR, Sobol and Niederreiter sequences perform the best and they reduce the variance by a factor as large as 2000. The combination of AV with LR and Sobol sequence are again the most efficient techniques; see Table 7. As expected, the sample average of optimal values converges to the same value as in the utility maximization problem of Section 4.1.2. Expected value and 90% confidence interval for the optimal values obtained with LR and MC are shown in Figure 5. In this problem, LR reduce the sample bias by a large factor and produce very tight confidence intervals for the optimal value. We characterize the infeasibility of the optimal solutions with implicit constraints by the amount of short selling in each discretized problem. The sample mean and 90% confidence interval for the maximum amount of short selling in the optimal portfolios for LR and MC are shown in Figure 6. With LR the minimum investment proportion converges towards zero much faster than with MC.

Table 6: Statistics for (WP^ν) as a function of ν .

ν		MC	AV	LH	LR	SOB	FAU	HAM	HAL	NIE
32	$\hat{\mu}$	-249.881	-305.066	-306.359	-308.130	-305.050	-283.060	-290.676	-289.656	-302.463
	$\hat{\sigma}$	4.64E+1	8.81E+0	9.47E+0	7.56E+0	9.20E+0	2.55E+1	2.05E+1	1.97E+1	1.01E+1
	vr	1.0	27.7	24.0	37.6	25.5	3.3	5.1	5.6	21.1
64	$\hat{\mu}$	-289.891	-312.191	-312.484	-313.656	-312.167	-306.806	-308.114	-307.586	-311.766
	$\hat{\sigma}$	1.84E+1	3.63E+0	4.08E+0	2.99E+0	4.10E+0	8.74E+0	5.80E+0	5.23E+0	3.74E+0
	vr	1.0	25.7	20.3	37.8	20.1	4.4	10.0	12.3	24.1
128	$\hat{\mu}$	-304.675	-314.157	-314.529	-314.818	-314.600	-314.295	-312.871	-312.464	-314.280
	$\hat{\sigma}$	1.13E+1	2.03E+0	2.05E+0	1.60E+0	2.09E+0	2.95E+0	2.50E+0	2.46E+0	1.68E+0
	vr	1.0	31.1	30.2	50.0	29.1	14.6	20.4	21.1	45.0
256	$\hat{\mu}$	-310.518	-314.955	-315.057	-315.253	-315.143	-315.198	-314.895	-314.597	-315.161
	$\hat{\sigma}$	7.25E+0	1.20E+0	1.08E+0	7.77E-1	9.65E-1	1.88E+0	1.03E+0	1.03E+0	8.51E-1
	vr	1.0	36.7	44.8	86.9	56.4	14.8	49.6	49.2	72.5
512	$\hat{\mu}$	-312.667	-315.069	-315.333	-315.435	-315.462	-315.238	-315.311	-315.278	-315.373
	$\hat{\sigma}$	4.62E+0	9.56E-1	5.44E-1	4.20E-1	4.80E-1	1.23E+0	4.93E-1	5.06E-1	4.12E-1
	vr	1.0	23.4	72.2	121	92.6	14.1	87.7	83.5	126
1024	$\hat{\mu}$	-314.353	-315.408	-315.401	-315.486	-315.498	-315.400	-315.425	-315.455	-315.452
	$\hat{\sigma}$	3.13E+0	6.34E-1	3.12E-1	2.02E-1	2.60E-1	5.67E-1	2.74E-1	2.53E-1	2.12E-1
	vr	1.0	24.3	101	239	144	30.4	130	153	217
2048	$\hat{\mu}$	-314.729	-315.424	-315.459	-315.504	-315.511	-315.470	-315.489	-315.480	-315.496
	$\hat{\sigma}$	2.18E+0	4.66E-1	1.77E-1	1.07E-1	1.38E-1	3.07E-1	1.34E-1	1.43E-1	1.11E-1
	vr	1.0	21.9	152	412	250	50.4	266	233	389
4096	$\hat{\mu}$	-315.208	-315.469	-315.479	-315.505	-315.505	-315.481	-315.508	-315.497	-315.502
	$\hat{\sigma}$	1.71E+0	2.96E-1	1.17E-1	5.93E-2	7.62E-2	1.57E-1	7.65E-2	7.08E-2	5.69E-2
	vr	1.0	33.5	214	835	505	119	500	584	905
8192	$\hat{\mu}$	-315.394	-315.480	-315.494	-315.502	-315.501	-315.492	-315.507	-315.502	-315.503
	$\hat{\sigma}$	1.23E+0	2.02E-1	6.66E-2	3.17E-2	3.80E-2	9.34E-2	4.09E-2	4.28E-2	3.24E-2
	vr	1.0	37.3	344	1512	1057	175	912	833	1454
16384	$\hat{\mu}$	-315.492	-315.494	-315.495	-315.506	-315.505	-315.511	-315.503	-315.506	-315.506
	$\hat{\sigma}$	8.68E-1	1.59E-1	4.93E-2	1.95E-2	1.91E-2	4.67E-2	2.09E-2	2.12E-2	1.89E-2
	vr	1.0	29.9	310	1991	2066	346	1733	1673	2103

Table 7: Statistics for (WP^ν) as a function of ν , Lattice rule and Sobol with AV.

ν		LR+AV	SOB+AV	ν		LR+AV	SOB+AV
32	$\hat{\mu}$	-305.578	-306.069	1024	$\hat{\mu}$	-315.452	-315.482
	$\hat{\sigma}$	8.92E+0	7.35E+0		$\hat{\sigma}$	1.39E-1	1.37E-1
	vr	27	40		vr	510	518
64	$\hat{\mu}$	-312.915	-313.070	2048	$\hat{\mu}$	-315.496	-315.494
	$\hat{\sigma}$	2.82E+0	2.12E+0		$\hat{\sigma}$	5.46E-2	7.54E-2
	vr	42	75		vr	1597	837
128	$\hat{\mu}$	-314.719	-314.737	4096	$\hat{\mu}$	-315.498	-315.502
	$\hat{\sigma}$	7.23E-1	7.99E-1		$\hat{\sigma}$	3.95E-2	4.78E-2
	vr	244	200		vr	1876	1285
256	$\hat{\mu}$	-315.192	-315.211	8192	$\hat{\mu}$	-315.505	-315.506
	$\hat{\sigma}$	4.17E-1	4.50E-1		$\hat{\sigma}$	2.66E-2	2.47E-2
	vr	302	259		vr	2156	2489
512	$\hat{\mu}$	-315.366	-315.405	16384	$\hat{\mu}$	-315.505	-315.504
	$\hat{\sigma}$	2.49E-1	2.30E-1		$\hat{\sigma}$	1.36E-2	1.67E-2
	vr	344	404		vr	4096	2719

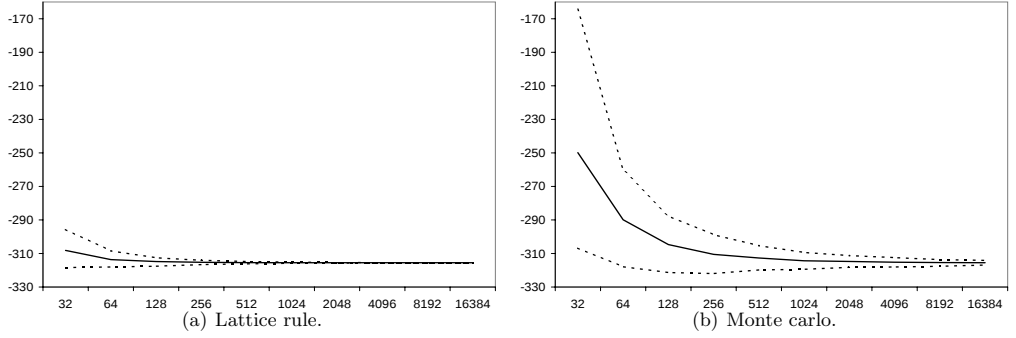


Figure 5: Mean and 90% confidence interval for the optimal value in utility maximization problem with implicit constraints.

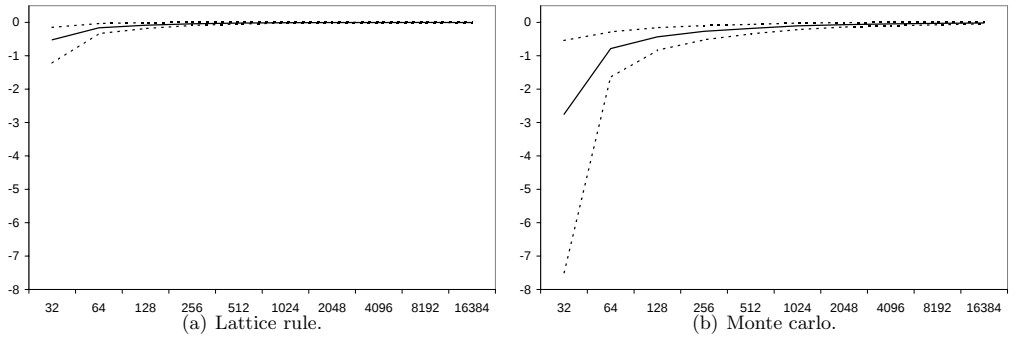


Figure 6: Mean and 90% confidence interval for infeasibility $\min_i x_i$.

References

- Z. Artstein and R. J.-B. Wets. Stability results for stochastic programs and sensors, allowing for discontinuous objective functions. *SIAM J. Optim.*, 4(3):537–550, 1994.
- Z. Artstein and R. J.-B. Wets. Consistency of minimizers and the SLLN for stochastic programs. *J. Convex Anal.*, 2(1-2):1–17, 1995.

- H. Attouch. *Variational convergence for functions and operators*. Pitman (Advanced Publishing Program), Boston, MA, 1984.
- P. Billingsley. *Convergence of probability measures*. John Wiley & Sons Inc., New York, 2nd edition, 1999.
- P. Boyle, M. Broadie, and P. Glasserman. Monte carlo methods for security pricing. *Journal of Economic Dynamics and Control*, 21:1267–1321, 1997.
- P. Bratley, B. L. Fox, and L. E. Schrage. *A guide to simulation*. Springer-Verlag, 2 edition, 1987.
- R. Caflisch, W. Morokoff, and A. Owen. Valuation of mortgage backed securities using brownian bridges to reduce effective dimension. *Journal of Computational Finance*, 1:27–46, 1997.
- R. Cranley and T. N. L. Patterson. Randomization of number theoretic methods for multiple integration. *SIAM J. Numer. Anal.*, 13(6):904–914, 1976.
- H. Faure. Discrepance de suites associées à un système de numération (en dimension s). *Acta Arith.*, 41(4):337–351, 1982.
- B. L. Fox. Algorithm 647: Implementation and relative efficiency of quasirandom sequence generators. *ACM Transactions on Mathematical Software*, 12(4):362–376, 1986.
- I. Friedel and A. Keller. Fast generation of randomized low discrepancy point sets. In K.-T. Fang, F. J. Hickernell, and H. Niederreiter, editors, *Monte Carlo and Quasi-Monte Carlo Methods 2000*, pages 257–273. Springer Verlag, 2002.
- J. H. Halton. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numer. Math.*, 2:84–90, 1960.
- J. M. Hammersley. Monte Carlo methods for solving multivariable problems. *Ann. New York Acad. Sci.*, 86:844–874 (1960), 1960.
- J. L. Hingle. Variance reduction and objective function evaluation in stochastic linear programs. *Journal on Computing*, 10:236–247, 1998.
- G. Infanger. Monte Carlo (importance) sampling within a Benders decomposition algorithm for stochastic linear programs. *Ann. Oper. Res.*, 39:69–95, 1992.
- P. Jäckel. *Monte Carlo Methods in Finance*. John Wiley & Sons, 2002.
- R. Kouwenberg. Scenario generation and stochastic programming models for asset liability management. *European J. Oper. Res.*, 134(2):279–292, 2001.
- P. L’Ecuyer and C. Lemieux. Variance reduction via lattice rules. *Management Science*, 46(2): 1214–1235, 2000.
- P. L’Ecuyer and C. Lemieux. Recent advances in randomized quasi-monte carlo methods. In M. Dror, P. L’Ecuyer, and F. Szidarovszki, editors, *Modeling Uncertainty: An Examination of Stochastic Theory, Methods, and Applications*, pages 419–474. Kluwer Academic Publishers, 2002.

- J. Linderoth, A. Shapiro, and S. Wright. The empirical behavior of sampling methods for stochastic programming. Optimization technical report 02-01, Computer Sciences Department, University of Wisconsin-Madison, 2002.
- R. Lucchetti, G. Salinetti, and R. J.-B. Wets. Uniform convergence of probability measures: topological criteria. *J. Multivariate Anal.*, 51(2):252–264, 1994.
- R. Lucchetti and R. J.-B. Wets. Convergence of minima of integral functionals, with applications to optimal control and stochastic optimization. *Statist. Decisions*, 11(1):69–84, 1993.
- M. Matsumoto and T. Nishimura. Mersenne twister: A 623-dimensionally equidistributed uniform pseudorandom number generator. *ACM Transactions Modeling Computer Simulation*, 8(1):3–30, 1998.
- H. Niederreiter. Low-discrepancy and low-dispersion sequences. *J. Number Th.*, 30:51–70, 1988.
- H. Niederreiter. *Random number generation and quasi-Monte Carlo methods*, volume 63 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.
- T. Pennanen and M. Koivu. Epi-convergent discretizations of stochastic programs via integration quadratures. *Stochastic Programming E-print Series*, 2003.
- W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical recipes in C, The art of scientific computing*. Cambridge University Press, Cambridge, 2nd edition, 1992.
- R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998.
- A. Shapiro. Stochastic programming by monte carlo simulation methods. *Stochastic Programming E-Print Series*, 2000.
- A. Shapiro. Inference of statistical bounds for multistage stochastic programming problems. *Mathematical Methods of Operations Research*, 58:57–68, 2003.
- I. H. Sloan and S. Joe. *Lattice methods for multiple integration*. Oxford Science Publications. The Clarendon Press Oxford University Press, New York, 1994.
- I. M. Sobol'. The distribution of points in a cube and the approximate evaluation of integrals. *U.S.S.R. Computational Math. And Math. Phys.*, (4):86–112, 1967.
- B. Tuffin. On the use of low discrepancy sequences in Monte Carlo methods. *Monte Carlo Methods and Appl.*, 2(4):295–320, 1996.
- X. Wang and K. T. Fang. The effective dimension and quasi-monte carlo integration. *Journal of Complexity*, 19:101–124, 2002.
- M. Zervos. On the epiconvergence of stochastic optimization problems. *Math. Oper. Res.*, 24(2):495–508, 1999.

A-SARJA: VÄITÖSKIRJOJA - DOCTORAL DISSERTATIONS. ISSN 1237-556X.

- A:197. ANSSI ÖÖRNI: Consumer Search in Electronic Markets. 2002. ISBN 951-791-680-9.
- A:198. ARI MATIKKA: Measuring the Performance of Owner-Managed Firms: A systems approach. 2002. ISBN 951-791-685-X.
- A:199. RIITTA KOSONEN: Governance, the Local Regulation Process, and Enterprise Adaptation in Post-Socialism. The Case of Vyborg. 2002. ISBN 951-791-692-2.
- A:200. SUSANNE SUHONEN: Industry Evolution and Shakeout Mechanisms: The Case of the Internet Service Provider Industry. 2002. ISBN 951-791-693-0.
- A:201. MATTI TUOMINEN: Market-Driven Capabilities and Operational Performance. Theoretical Foundations and Managerial Practices. 2002. ISBN 95-791-694-9.
- A:202. JUSSI KARHUNEN: Essays on Tender Offers and Share Repurchases. 2002. ISBN 951-791-696-5.
- A:203. HENNAMARI MIKKOLA: Empirical Studies on Finnish Hospital Pricing Methods. 2002. ISBN 951-791-714-7.
- A:204. MIKA KORTELAJINEN: EDGE: a Model of the Euro Area with Applications to Monetary Policy. 2002. ISBN 951-791-715-5.
- A:205. TOMI LAAMANEN: Essays on Technology Investments and Valuation. 2002. ISBN 951-791-716-3.
- A:206. MINNA SÖDERQVIST: Internationalisation and its Management at Higher-Education Institutions. Applying Conceptual, Content and Discourse Analysis. 2002. ISBN 951-791-718-X.
- A:207. TARJA PIETILÄINEN: Moninainen yrittäminen. Sukupuoli ja yrittäjänäisten toimintatila tietoteollisuudessa. 2002. ISBN 951-791-719-8.
- A:208. BIRGIT KLEYMANN: The Development of Multilateral Alliances. The Case of the Airline Industry. 2002. ISBN 951-791-720-1.
- A:209. MIKAEL EPSTEIN: Risk Management of Innovative R&D Project. Development of Analysis Model. A Systematic Approach for the Early Detection of Complex Problems (EDCP) in R&D Projects in Order to Increase Success in Enterprises. 2002. ISBN 951-791-717-9.
- A:210. SAMI KAJALO: Deregulation of Retail Hours in Finland: Historical and Empirical Perspectives. 2002. ISBN 951-791-734-1.
- A:211. TOMMI KASURINEN: Exploring Management Accounting Change in the Balanced Scorecard Context. Three Perspectives. 2003. ISBN 951-791-736-8.

- A:212. LASSE NIEMI: Essays on Audit Pricing. 2003. ISBN 951-791-751-1.
- A:213. MARKKU KAUSTIA: Essays on Investor Behavior and Psychological Reference Prices. 2003. ISBN 951-791-754-6.
- A:214. TEEMU YLIKOSKI: Access Denied: Patterns of Consumer Internet Information Search and the Effects of Internet Search Expertise. 2003. ISBN 951-791-755-4.
- A:215. PETRI HALLIKAINEN: Evaluation of Information System Investments. 2003. ISBN 951-791-758-9.
- A:216. PETRI BÖCKERMAN: Empirical Studies on Working Hours and Labour Market Flows. 2003. ISBN 951-791-760-0.
- A:217. JORMA PIETALA: Päivittäistavarakaupan dynamiikka ja ostoskäyttäytyminen Pääkaupunkiseudulla. 2003. ISBN 951-791-761-9.
- A:218. TUOMAS VÄLIMÄKI: Central Bank Tenders: Three Essays on Money Market Liquidity Auctions. 2003. ISBN 951-791-762-7.
- A:219. JUHANI LINNAINMAA: Essays on the Interface of Market Microstructure and Behavioral Finance. 2003. ISBN 951-791-783-X.
- A:220. MARKKU SALIMÄKI: Suomalaisen design-teollisuuden kansainvälinen kilpailukyky ja kansainvälistyminen. Strateginen ryhmä –tutkimus design-aloilta. 2003 ISBN 951-791-786-4.
- A:221. HANNU KAHRA: Consumption, Liquidity and Strategic Asset Allocation. 2003. ISBN 951-791-791-0.
- A:222. TONI RIIPINEN: The Interaction of Environmental and Trade Policies. 2003. ISBN 951-791-797-X.
- A:223. MIKKO SYRJÄNEN: Data Envelopment Analysis in Planning and Heterogeneous Environments. 2003. ISBN 951-791-806-2.
- A:224. ERKKI HÄMÄLÄINEN: Evolving Logistic Roles of Steel Distributors. 2003. ISBN 951-791-807-0.
- A:225. SILJA SIITONEN: Impact of Globalisation and Regionalisation Strategies on the Performance of the World's Pulp and Paper Companies. 2003. ISBN 951-791-808-9.
- A:226. EIREN TUUSJÄRVI: Multifaceted Norms in SMC Export Cooperation: A Discourse Analysis of Normative Expectations. 2003. ISBN 951-791-812-7.
- A:227. MIKA MALIRANTA: Micro Level Dynamics of Productivity Growth. An Empirical Analysis of the Great Leap in Finnish Manufacturing Productivity in 1975-2000. 2003. ISBN 951-791-815-1.
- A:228. NINA KOISO-KANTTILA: Essays on Consumers and Digital Content. 2003. ISBN 951-791-816-X.
- A:229. PETER GABRIELSSON: Globalising Internationals: Product Strategies of ICT Companies. 2004. ISBN 951-791-825-9, ISBN 951-791-826-7 (Electronic dissertation).

- A:230. SATU NURMI: Essays on Plant Size, Employment Dynamics and Survival. 2004. ISBN 951-791-829-1, ISBN 951-791-830-5 (Electronic dissertation).
- A:231. MARJA-LIISA KURONEN: Vakuutusehtotekstin uudistamisprosessi, matkalla alamaisestä asiakkaaksi. 2004. ISBN 951-791-833-X, ISBN 951-791-834-8 (Electronic dissertation).
- A:232. MIKA KUISMA: Erilaistuminen vai samanlaistuminen? Vertaileva tutkimus paperiteollisuusyhtiöiden ympäristöjohtamisesta. 2004. ISBN 951-791-835-6, ISBN 951-791-836-4 (Elektronic dissertation).
- A:233. ANTON HELANDER: Customer Care in System Business. 2004. ISBN 951-791-838-0.
- A:234. MATTI KOIVU: A Stochastic Optimization Approach to Financial Decision Making. 2004. ISBN 951-791-841-0, ISBN 951-791-842-9 (Elektronic dissertation).

A-SARJA: MUITA JULKAISUJA - OTHER PUBLICATIONS

ANNE HERBERT: The Paradoxes of Action Learning: An Interpretive and Critical Inquiry into Vocational Educators' Professional Development. 2002. ISBN 951-791-684-1.

B-SARJA: TUTKIMUKSIA - RESEARCH REPORTS. ISSN 0356-889X.

- B:38. KRISTIINA KORHONEN(ed.): Current Reflections on the Pacific Rim. 2002. ISBN 951-791-661-2.
- B:39. RISTO Y. JUURMAA: Performance and International Competitiveness of Listed Metal and Telecommunication Industry Groups 1992 - 2000. Finland vs Sweden and Germany. 2002. ISBN 951-791-668-X.
- B:40. KAIJA TUOMI – SINIKKA VANHALA (toim.): Yrityksen toiminta, menestyminen ja henkilöstön hyvinvointi. Seurantatutkimus metalliteollisuudessa ja vähittäiskaupan alalla. 2002. ISBN 951-791-674-4.
- B:41. ANNE ÄYVÄRI: Verkottuneen pienyrityksen markkinointikyvykyys. 2002. ISBN 951-791-682-5.
- B:42. RIKU OKSMAN: Intohimoa ja ammattitaitoa: puheenvuoroja tuottajan työstä. 2002. ISBN 951-791-700-7.
- B:43. RISTO TAINIO – KARI LILJA – TIMO SANTALAINEN: Organizational Learning in the Context of Corporate Growth and Decline: A Case Study of a Major Finnish Bank. 2002. ISBN 951-791-717-1
- B:44. ELINA HENTTONEN – PÄIVI ERIKSSON – SUSAN MERILÄINEN: Teknologiayrittämisen sukupuoli. Naiset miesten maailmassa. 2003. ISBN 951-791-737-6.
- B:45. KIRSI KORPIAHO: "Kyllä siinä pitää elää mukana!" Kirjanpitäjien tarinoita työstä, osaamisesta ja oppimisesta työyhteisönäkökulmasta analysoituna. 2003. ISBN 951-791-742-2.

- B:46. NIILLO HOME (toim.): Puheenvuoroja ECR-toiminnasta. Discussions on ECR – Summaries. 2003. ISBN 951-791-749-X.
- B:47. PÄIVI KARHUNEN – RIITTA KOSONEN – MALLA PAAJANEN: Gateway-käsitteen elinkaari Venäjän-matkailussa. Etelä-Suomi Pietarin-matkailun väylänä. 2003. ISBN 951-791-756-2.
- B:48. ANNELI KAUPPINEN – ANNE ARANTO – SATU RÄMÖ (toim.): Myyttiset markkinat. 2003. ISBN 951-791-771-6.
- B:49. MIKKO SAARIKIVI – SIMO RIIHONEN: Suomen puuteollisuuden kilpailukyvyn parantaminen ja kansainvälistyminen piha- ja ympäristörakentamisessa. 2003. ISBN 951-791-779-1.
- B:50. KATARIINA KEMPPAINEN – ARI P.J. VEPSÄLÄINEN – JUKKA KALLIO – TIMO SAARINEN – MARKKU TINNILÄ: From Supply Chain to Networks: A Study of SCM Practices in Finnish Industrial Companies. 2003. ISBN 951-791-780-5.
- B:51. SAMI SARPOLA: Enterprise Resource Planning (ERP) Software Selection and Success of Acquisition Process in Wholesale Companies. 2003. ISBN 951-791-802-X.
- B:52. MATTI TUOMINEN (ed.): Essays on Capabilities Based Marketing and Competitive Superiority. Fimac II - Research: Mai Anttila, Saara Hyvönen, Kristian Möller, Arto Rajala, Matti Tuominen. 2003. ISBN 951-791-814-3.
- B:53. PÄIVI KARHUNEN – RIITTA KOSONEN – ANTTI LEIVONEN: Osaamisen siirtyminen Suomalais-venäläisissä tuotantoalliansseissa. Tapaustutkimuksia pietarista ja leningradin alueelta. 2003. ISBN 951-791-820-8.
- B:54. JARMO ERONEN: Kielten välinen kilpailu: Taloustieteellis-sosiolingvistinen tarkastelu. 2004. ISBN 951-791-828-3.

CKIR-SARJA: HELSINKI SCHOOL OF ECONOMICS. CENTER FOR KNOWLEDGE AND INNOVATION RESEARCH. CKIR WORKING PAPERS. ISSN 1458-5189.

- CKIR:1. SATINDER P. GILL: The Engagement Space and Parallel Coordinated Movement: Case of a Conceptual Drawing Task. 2002. ISBN 951-791-660-4.
- CKIR:2. PEKKA ISOTALUS – HANNI MUUKKONEN: How Do Users of Pda's React to an Animated Human Character in Online News? 2002. ISBN 951-791-664-7.

E-SARJA: SELVITYKSIÄ - REPORTS AND CATALOGUES. ISSN 1237-5330.

- E:100. JUHA KINNUNEN: Opiskelijoiden valikoituminen pääaineisiin Helsingin kauppakorkeakoulussa. Pääainetoiveita ja niihin vaikuttavia tekijöitä kartoittava kyselytutkimus vuosina 1995-2000 opintonsa aloittaneista. 2002. ISBN 951-791-669-8.
- E:101. Research Catalogue 2000 – 2002. Projects and Publications. 2002. ISBN 951-791-670-1.

E:102. DAN STEINBOCK: The U.S. CIBER Experience: The Centers for International Business Education and Research (CIBERs). 2003. ISBN 951-791-781-3.

N-SARJA: HELSINKI SCHOOL OF ECONOMICS. MIKKELI BUSINESS CAMPUS PUBLICATIONS.
ISSN 1458-5383

N:6 JUHA SIIKAVUO: Taloushallinnon opas alkavalle yrittäjälle. 2002. ISBN 951-791-686-8.

N:7. JOHANNA NISKANEN: Etelä-Savon pk-yritysten vienti Tanskaan: ulkomaankaupan erityisraportti 2001. 2002. ISBN 951-791-687-6.

N:8. MIKKO NUMMI: Etelä-Savon pk-yritysten vienti Saksaan: ulkomaankaupan erityisraportti 2001. 2002. ISBN 951-791-688-4.

N:9. NOORA RUOHONEN – RIIKKA OLLI: Etelä-Savon pk-yritysten vienti Tsekiin: ulkomaankaupan erityisraportti 2001. 2002. ISBN 951-791-689-2.

N:10. ANNA HÄKKINEN – ESKO LÄIKKÖ: Etelä-Savon pk-yritysten vientikohteena USA: ulkomaankaupan erityisraportti 2001. 2002. ISBN 951-791-690-6.

N:11. JUHA SIIKAVUO: Verkko-oppimisympäristön kehittäminen. Esimerkkinä HA Boctok Venäjän-kaupan erikoistumisopintojen yksi moduuli, vuosi: 2002. 2002. ISBN 951-791-695-7.

N:12. JUHO PETTER PUHAKAINEN: German Venture Capitalists' Decision Criteria in New Venture Evaluation. 2002. ISBN 951-791-650-7.

N:13. MILJA LEMMETYINEN: Suomalaisyrityksen etabloituminen Saksaan. Ulkomaankaupan erityisraportti 2002. 2002. ISBN 951-791-730-9.

N:14. TAPIO PALLASVIRTA: Pk-yritysten vienti Espanjaan. Ulkomaankaupan erityisraportti 2002. 2002. ISBN 951-791-731-7.

N:15. ELINA HAVERINEN: Etelä-Savon pk-yritysten Viron kauppa. Ulkomaankaupan erityisraportti 2003. ISBN 951-791-732-5.

N:16. REETA RÖNKKÖ: Latinalainen Amerikka markkina-alueena Argentiina ja Brasilia. Ulkomaankaupan erityisraportti 2003. ISBN 951-791-733-3.

N:17. JAAKKO VARVIKKO – JUHA SIIKAVUO: Koulutus, oppiminen ja akateeminen yrittäjyys. 2003. ISBN 951-791-745-7.

N:18. ANNE GUSTAFSSON-PESONEN – SATU SIKANEN: Yrittäjäkoulutuksesta yrittäjäksi. 2003 ISBN 951-791-763-5.

N:19. TOIVO KOSKI: Impact of a venture capitalists' value added on value of a venture. 2003. ISBN 951-791-764-3.

N:20. LAURA HIRVONEN: Itävalta suomalaisyritysten markkina-alueena. 2003. ISBN 951-791-765-1.

N:21. LAURA MALIN: Etelä-Savon pk-yritysten vienti Belgiaan. 2003. ISBN 951-791-766-X.

- N:22. JUKKA PREPULA: Ranska suomalaisten pk-yritysten vientikohteena. 2003. ISBN: 951-791-767-8.
- N:23. HENNA HUCZKOWSKI: Pk-yritysten perustaminen Puolaan. 2003. ISBN 951-791-768-6.
- N:24. HENNA KATAJA – LEENA MÄÄTTÄ: Kiina suomalaisen pk-yrityksen vientikohteena. 2003. ISBN: 951-791-769-4.
- N:25. KAROLIINA IJÄS: Etelä-Savon pk-yritysten vienti Puolaan. 2003. ISBN: 951-791-770-8.
- N:26. MARJO VAHLSTEN: Matkailupalvelujen markkinoinnin kehittäminen verkkoyhteistyön avulla. 2003. ISBN: 951-791-792-9.
- N:27. TUULI SAVOLAINEN: Slovakia suomalaisten pk-yritysten markkina-alueena. 2003. ISBN: 951-791-793-7.
- N:28. HARRY MAASTOVAARA: Etelä-Savon yritysten ulkomaankauppa 2001. 2003. ISBN: 951-791-794-5.
- N:31. HANNA PERÄLÄ: Etelä-Savon pk-yritysten vienti Ruotsiin. 2003. ISBN: 951-791-799-6.
- N:34. TOIVO KOSKI – ANTTI EKLÖF: Uudenmaan yrityshautomoista irtaantuneiden yritysten menestyminen, Yrittäjien näkemyksiä yrityshautomotoiminnasta sekä selvitys ”yrittämisestä Työtä 2000” –projektin asiakkaiden yritystoiminnasta. 2003. ISBN 951-791-805-4.

W-SARJA: TYÖPAPEREITA - WORKING PAPERS . ISSN 1235-5674.

- W:304. PETRI BÖCKERMAN – KARI HÄMÄLÄINEN – MIKA MALIRANTA: Explaining Regional Job and Worker Flows. 2002. ISBN 951-791.662-0.
- W:305. PEKKA KORHONEN – MIKKO SYRJÄNEN: Evaluation of Cost Efficiency in Finnish Electricity Distribution. 2002. ISBN 951-791-663-9.
- W:306. SATU NURMI: The Determinants of Plant Survival in Finnish Manufacturing. 2002. ISBN 951-791-665-5.
- W:307. JUSSI KARHUNEN: Taking Stock of Themselves. An Analysis of the Motives and the Market Reaction in Finnish Share Repurchase Programs 2002. ISBN 951-791-666-3.
- W:308. PEKKA ILMAKUNNAS – HANNA PESOLA: Matching Functions and Efficiency Analysis. 2002. ISBN 951-791-671-X.
- W:309. MARKKU SÄÄKSJÄRVI: Software Application Platforms: From Product Architecture to Integrated Application Strategy. 2002. ISBN 951-791-672-8.
- W:310. MILLA HUURROS – HANNU SERISTÖ: Alliancing for Mobile Commerce: Convergence of Financial Institutions and Mobile Operators. 2002. ISBN 951-791-673-6.
- W:311. ANSSI ÖÖRNI: Objectives of Search and Combination of Information Channels in Electronic Consumer Markets: An Explorative Study. 2002. ISBN 951-791-675-2.

- W:312. ANSSI ÖÖRNI: Consumer Search in Electronic Markets: Experimental Analysis of Travel Services. 2002. ISBN 951-791-676-0.
- W:313. ANSSI ÖÖRNI: Dominant Search Pattern in Electronic Markets: Simultaneous or Sequential Search. 2002. ISBN 951-791-677-9.
- W:314. ANSSI ÖÖRNI: The Amount of Search in Electronic Consumer Markets. 2002. ISBN 951-791-678-7.
- W:315. KLAUS KULTTI – TUOMAS TAKALO – TANJA TANAYAMA: R&d Spillovers and Information Exchange: A Case Study. 2002. ISBN 951-791-681-7.
- W:316. OLLI TAHVONEN: Timber Production v.s. Old Growth Conservation with Endogenous Prices and Forest Age Classes. 2002. ISBN 951-791-691-4.
- W:317. KLAUS KULTTI – JUHA VIRRANKOSKI: Price Distribution in a Symmetric Economy. 2002. ISBN 951-791-697-3.
- W:318. KLAUS KULTTI – TONI RIIPINEN: Multilateral and Bilateral Meetings with Production Heterogeneity. 2002. ISBN 951-791-698-1.
- W:319. MARKKU KAUSTIA: Psychological Reference Levels and IPO Stock Returns. 2002. ISBN 951-791-699-X.
- W:320. MERVI LINDQVIST: Possible Research Issues in Management and Control of New Economy Companies. 2002. ISBN 951-791-701-5.
- W:321. MARKO LINDROOS: Coalitions in Fisheries. 2002. ISBN 951-791-702-3.
- W:322. MIKKO SYRJÄNEN: Non-discretionary and Discretionary Factors and Scale in Data Envelopment Analysis. 2002. ISBN 951-791-705-8.
- W:323. KLAUS KULTTI – HANNU VARTIAINEN: VonNeumann-Morgenstern Solution to the Cake Division Problem. 2002. ISBN 951-791-708-2.
- W:324. TOM LAHTI: A Review of the Principal-agent Theory and the Theory of Incomplete Contracts: An Examination of the Venture Capital Context. 2002. ISBN 951-791-709-0.
- W:325. KRISTIAN MÖLLER – PEKKA TÖRRÖNEN: Business Suppliers' Value-Creation Potential: A Capability-based Analysis. 2002. ISBN 951-791-710-4.
- W:326. KRISTIAN MÖLLER – ARTO RAJALA – SENJA SVAHN: Strategic Business Nets – Their Types and Management. 2002. ISBN 951-791-711-2.
- W:327. KRISTIAN MÖLLER – SENJA SVAHN – ARTO RAJALA: Network Management as a Set of Dynamic Capabilities. 2002. ISBN 951-791-712-0.
- W:328. PANU KALMI: Employee Ownership and Degeneration. Evidence from Estonian case studies. 2002. ISBN 951-791-713-9.
- W:329. ANNELI NORDBERG: Yrittäjyys, johtajuus ja johtaminen – uuden talouden innovatiivisia haasteita. 2002. ISBN 951-791-721-X.
- W:330. LEENA LOUHIHALA-SALMINEN: Communication and language use in merged corporations: Cases Stora Enso and Nordea. 2002. ISBN 951-791-722-8.

- W:331. TOMMI KASURINEN: Conceptualising the Encoding Process Related to Institutionalisation in Organisations. From Key Performance Indicator Scorecard to a Strategic Balanced Scorecard. 2002. ISBN 951-791-723-6.
- W:332. PEKKA KORHONEN – HELENA TOPDAGI: Performance of the AHP in Comparison of Gains and Losses. 2002. ISBN 951-791-724-4.
- W:333. TARJA JORO – PEKKA KORHONEN – STANLEY ZIONTS: An Interactive Approach to Improve Estimates of Value Efficiency in Data Envelopment Analysis. 2002. ISBN 951-791-725-2.
- W:334. JUHA-PEKKA TOLVANEN – JEFF GRAY – MATTI ROSSI (edit.): Proceedings of the Second-Domain Specific Modeling Languages Workshop. 2002. ISBN 951-791-726-0.
- W:335. SATU NURMI: Sectoral Differences In Plant Start-up Size. 2003. ISBN 951-791-738-4.
- W:336. SATU NURMI: Plant Size, Age And Growth In Finnish Manufacturing. 2003. ISBN 951-791-739-2.
- W:337. PETRI HALLIKAINEN – HANNU KIVIJÄRVI: Appraisal of Strategic it Investments: Payoffs And Tradeoffs. 2003. ISBN 951-791-740-6.
- W:338. SENJA SVAHN: Knowledge Creation in Business Networks – A Dynamic-capability Perspective. 2003. ISBN 951-791-743-0.
- W:339. KRISTIAN MÖLLER – SENJA SVAHN: Role of Knowledge in the Value Creation in Business Nets. 2003. ISBN 951-791-744-9.
- W:340. ELI MOEN – KARI LILJA: European Works Councils in M-Real and Norske Skog: The Impact of National Traditions in Industrial Relations. 2003. ISBN 951-791-750-3.
- W:341. KJELD MÖLLER: Salatulla "arvopaperistamisella" tuhottiin yrittäjyyttä. 2003. ISBN 951 791-752-X
- W:342. ATSO ANDERSEN: Competition Between European Stock Exchanges. 2003. ISBN 951-791-753-8.
- W:343. MARKO MERISAVO: The Effects of Digital Marketing on Customer Relationships. 2003. ISBN 951-791-757-0.
- W:344. KLAUS KULTTI – JUHA VIRRANKOSKI: Price Distribution in a Random Matching Model. 2003. ISBN 951-791-759-7.
- W:345. PANU KALMI: The Rise and Fall of Employee Ownership in Estonia, 1987-2001. 2003. ISBN 951-791-772-4.
- W:346. SENJA SVAHN: Managing in Networks: Case Study of Different Types of Strategic Nets. 2003. ISBN 951-791-774-0.
- W:347. KRISTIAN MÖLLER – SENJA SVAHN: Crossing East-West Boundaries: Knowledge Sharing in Intercultural Business Networks. 2003. ISBN 951-791-775-9.
- W:348. KRISTIAN MÖLLER – SENJA SVAHN: Managing in Emergence: Capabilities for Influencing the Birth of New Business Fields. 2003. ISBN 951-791-776-7.

- W:349. TOM RAILIO: The Taxation Consequences of Scandinavian Mutual Fund Investments and After-Tax Performance Evaluation. 2003. ISBN 951-791-777-5.
- W:350. KIRSI LAPOINTE: Subjekttiivinen ura työurien tutkimuksessa ja teorioissa. 2003. ISBN 951-791-778-3.
- W:351. PANU KALMI: The Study of Co-operatives in Modern Economics: A Methodological Essay. 2003. ISBN 951-791-783-X.
- W:352. MARJA TAHVANAINEN: Short-term International Assignments: Popular Yet Largely Unknown Way Of Working Abroad. 2003. ISBN 951-791-784-8.
- W:353. MARKKU KUULA – ANTOINE STAM: An Interior Point Method for Multi-party Negotiation Support. 2003. ISBN 951-791-787-2.
- W:354. JOUKO KINNUNEN: Quantification of Ordered-level Business Sentiment Survey Forecasts by Means of External Validation Data. 2003. ISBN 951-791-790-2.
- W:355. TOM RAILIO: The Nature of Disagreements and Comparability Between Finnish Accumulating Mutual Funds and Voluntary Pension Insurances. 2003. ISBN 951-791-798-8.
- W:356. JUKKA JALAVA: 'Has Our Country the Patience Needed to Become Wealthy?' Productivity in the Finnish Manufacturing Industry, 1960-2000. 2003. ISBN 951-791-803-8.
- W:357. JARI VESANEN: Breaking Down Barriers for Personalization – A Process View. 2003. ISBN 951-791-804-6.
- W:358. JUHA VIRRANKOSKI: Search Intensities, Returns to Scale, and Uniqueness of Unemployment Equilibrium. 2003. ISBN 951-791-809-7.
- W:359. JUHA VIRRANKOSKI: Search, Entry, and Unique Equilibrium. 2003. ISBN 951-791-810-0.
- W:360. HANNA KALLA: Exploration of the Relationship Between Knowledge Creation, Organisational Learning, and Social Capital: Role of Communication. 2003. ISBN 951-791-813-5.
- W:361. PEKKA SÄÄSKILAHTI: Strategic R&D and Network Compatibility. 2003. ISBN 951-791-817-8.
- W:362. MAIJU PERÄLÄ: Allyn Young and the Early Development Theory. 2003. ISBN 951-791-818-6.
- W:363. OSSI LINDSTRÖM – ALMAS HESHMATI: Interaction of Real and Financial Flexibility: An Empirical Analysis. 2004. ISBN 951-791-827-5 (Electronic working paper).
- W:364. RAIMO VOUTILAINEN: Quantitative Methods in Economics and Management Science. 2004. ISBN 951-791-832-1 (Electronic working paper).
- W:365. MATTIKELOHARJU – SAMULI KNÜPFER – SAMI TORSTILA: Retail Incentives in Privatizations: Anti-Flipping Devices or Money Left on the Table? 2004. ISBN 951-791-839-9 (Elektronic working paper).

Y-SARJA: HELSINKI SCHOOL OF ECONOMICS.
CENTRE FOR INTERNATIONAL BUSINESS RESEARCH. CIBR RESEARCH PAPERS.
ISBN 1237-394X.

Y:7. ZUHAIR AL-OBAIDI – MIKA GABRIELSSON: Multiple Sales Channel Strategies in Export Marketing of High Tech SMEs. 2002. ISBN 951-791-703-1.

Z-SARJA: HELSINKI SCHOOL OF ECONOMICS.
CENTRE FOR INTERNATIONAL BUSINESS RESEARCH. CIBR WORKING PAPERS. ISSN 1235-3931.

Z:9. V.H. MANEK KIRPALANI – MIKA GABRIELSSON: Worldwide Evolution of Channels Policy. 2002. ISBN 951-791-704-X.

Kaikkia Helsingin kauppakorkeakoulun julkaisusarjassa ilmestyneitä julkaisuja voi tilata osoitteella:

KY-Palvelu Oy
Kirjakauppa
Runeberginkatu 14-16
00100 Helsinki
Puh. (09) 4313 8310, fax (09) 495 617
Sähköposti: kirjak@kyypari.hkkk.fi

Helsingin kauppakorkeakoulu
Julkaisutoimittaja
PL 1210
00101 Helsinki
Puh. (09) 4313 8579, fax (09) 4313 8305
Sähköposti: julkaisu@hkkk.fi

All the publications can be ordered from

Helsinki School of Economics
Publications officer
P.O.Box 1210
FIN-00101 Helsinki
Phone +358-9-4313 8579, fax +358-9-4313 8305
E-mail: julkaisu@hkkk.fi