Funding optimization for a bnak: Integrating credit and liquidity risk

Petrus Strydom

October 19, 2016

Abstract

In this paper we apply two optimization frameworks to determine the optimal wholesale funding mix of a bank given uncertainty in both credit and liquidity risk. A stochastic linear programming method is used to find the optimal strategy to be maintained across all scenarios. A recursive learning method is developed to provide the bank with a trading signal to dynamically adjust the wholesale funding mix as the macroeconomic environment changes. The performance of the two methodologies is compared in the final section.

Contents

1 Introduction **1**
1.1 SLP: Literature study 2
1.2 RRL : Literature study 4

2 Model setup **7**
2.1 Sub-models 9

3 Stochastic Linear Programming **12**
3.1 Eventtree 12
3.2 Methodology 13
3.3 Results 15

4 Recurrent Reinforcement Learning **18**
4.1 Methodology 18
4.2 Results 21

5 Results **23**

Bibliography **26**

List of Figures

2.1 Diagram of the model framework to apply the optimization methods 10

4.1 Steps in the RRL optimization methodology 22
4.2 Portion of funding gap filled with short dated debt as credit losses change 22

Chapter 1

Introduction

Banks provide loans to both retail and corporate counterparties. These loans are assets on the balance sheet that yield a certain interest rate. The bank requires funding (a liability on the balance sheet) to support this lending activity. The main types of funding available to a bank are:

- Deposits from both retail and wholesale customers.
- Debt instruments of varying term issued directly to the market (wholesale funding).

This exposes the bank to the risk of counterparties failing to repay the loans, which is termed *credit events*. The deposit and debt instruments used to fund the loans are usually short term in nature creating a mismatch compared to the long term nature of the asset profile (i.e. a 20 year mortgage loan funded via 3 month debt instruments). This mismatch exposes the bank to interest rate risk (assets and liabilities re-price at different durations) and liquidity risk (the uncertainty of the cost of funding at future dates). The extreme and novel macroeconomic realities observed over the last couple of years exposed a number of weaknesses in the risk management methodologies used by banks. This includes much higher credit losses than expected, higher liquidity premiums on wholesale funding during times of distress and the volatility of the deposit base during a flight to safety. A major weakness in the current risk management methodology is the understanding of the relationship of credit, liquidity and interest rate risk.

To ensure profitability the interest earned on the assets should exceed the cost of funding. The bank needs to continuously fund the balance sheet as the existing funding mature and the level of the deposits change with the economic environment. Wholesale funding is an important funding source for South African banks. Bank's issue debt at various durations, ranging from overnight to 60 month instruments. In a positive interest rate environment short dated debt is usually cheaper compared to longer dated instruments however funding with short dated instruments exposes the bank to more roll over risk events, where the cost of rolling debt is uncertain (i.e. liquidity risk). The optimization methodologies attempt to balance the cost of wholesale funding with the liquidity and interest rate risk.

This paper integrates the sub-components underlying the banks' balance sheet to facilitate the projection of the net interest income allowing for both liquidity, interest and credit risk. The sub-components include retail and wholesale loans, retail and wholesale deposits and bank issued debt instruments. Stochastic linear program ("SLP") and recursive learning ("RRL") models are developed to determine the optimal duration mixes for the wholesale funding.

The calibration of the sub-components is a research topic in its own right. Only a simplified representation was assumed to empirically test the optimization models developed in this paper.

The SLP method is used to determine the optimal duration of the wholesale or debt funding given the uncertainty. This provides the funding duration that should be maintained overtime. The RRL is a dynamic model that provides a trading signal to dynamically adjust the duration of the wholesale funding portfolio as interest rates and the credit losses change. A comparison of the returns of the RRL and SLP is used to test the performance of each method.

1.1 SLP: Literature study

The uncertainty underlying a bank's assets and liabilities has prompted banks to seek greater efficiency in the management of their assets and liabilities. This has led to studies concerned with the structure of the bank's assets and liabilities to achieve some optimal trade-off among the various risks. Chambers and Charnes (1961) [8] wrote one of the first papers based on maximizing profitability within capital and liquidity constraints. Uncertainty is reflected in the credit, liquidity and interest rate risk embedded in the performance of both assets and liabilities. Mathematical programming models that incorporate this uncertainty are known as stochastic programs.

Available stochastic program methodologies include: change constraint programming, dynamic programming, sequential decision theory, stochastic decision trees and linear programming under uncertainty (or stochastic linear programming (SLP)).

The text book by Zenios and Ziemba (2006) [42] set out the practical application of stochastic programming. Kusy and Ziemba (1984) [24] was one of the first practitioners to advocate the used to stochastic linear programming with simple recourse for an asset liability framework, identifying challenges with available computer power to solve these large problems. Guven and Persentili (1997) [18] also put forward the SLP approach to solve the stochastic program presented by the asset liability problem. The evolution of both computational power and more refined search algorithms have promoted this methodology. The method is widely used to support financial decision making, see Kouwenberg and Zenios (2001) [23], Carino *et al.* (1994) [7], Edirisinghe and Patterson (2007) [14], Hill *et al.* (2007) [20] and Yingjie and Cheng-iin (2000) [41] [4]. This methodology allows for a traceable solution when the problem statement extend over multiple periods and support the path dependency of the wholesale funding decisions. The SLP model can be extended to include multiple objectives, such as liquidity constraints and profit maximization. A multi objective approach was not considered as part of this paper however the current methodology can be extended to include this, see Aouni, Colapinto and La Torre (2014) [1] and Kosmidou and Zopounidis (2008) [22].

The solution to solve the stochastic linear programs, including the various forms of recourse rest on the pioneering work by Benders (1962) [3], Dantzig (1963) [9] and Dantzig and Wolfe (1960) [10]. These authors developed various methodologies to decompose a problem using either an inner or outer linearization to solve a large and complex problem. Benders decomposition breaks a large problem into a number of smaller problems that can be solved individually while mining for a global solution through an iterative process. The Dantzig - Wolfe decomposition focus on the duel of the linear problem.

The properties of the linear problem and in particular the properties of the recourse function are key to determine the convergence, feasibility and optimality of the various search algorithms proposed. Van Slyke and Wets (1969) [37] extended Benders decomposition into a solution termed the L-Shape method. This will be the method used to solve the stochastic linear problem in this paper. The text books by Brige and Louveaux (1997) [5] and Kall (1976) [21] provides a good overview of developments in linear programming, including the L-Shape methodology and the various important theoretical consideration to ensure feasibility, optimality and convergence. Murphy (2013) [30], Wets (2000) [39] and Dempster (1980) [11] provides a good review on the L-Shaped methodology. There has been a number of enhancement to the original L-Shape method such as more robust feasibility cuts, using a multi cut approach to speed up convergence and methods such as bunching and realizations, see Brige and Louveaux (1997) [5] for a discussion on these approaches.

1.2 RRL : Literature study

Dynamic programming, and in particular reinforcement learning is widely recognized in financial decision models. This is widely used to develop automated trading rules or portfolio selection models. The setup of the optimization problem, in particular the path dependency and dynamic nature of the decision process aligns well with a dynamic programming methodology. The reward function underlying the reinforcement learning methodology can be non linear providing more flexibility as the SLP method. This flexibility allows for the risk in the form of earnings volatility to be included in the optimization criteria.

The optimization problem share similarities with a Markov decision process ("MDP"). Formulating the optimization problem in this way opens up the field of reinforcement learning. As discussed in Marsland (2009) [27], Goldberg (1989) [15], Busoniu *et al.* (2009) [6] and Sutton (1992) [36] a MDP is a mathematical formulation partitioned over various statuses or time intervals with a transition function to measure the movement across the various statuses and a corresponding reward function to measure the impact of the decision. A MDP has an agent (or multiple agents) that makes policy decisions affecting the transition function. The aim is to train the agent or policy function to optimize the reward, usually based on historic data or real time on-line learning.

An important consideration in specifying the MDP is the path dependency of the reward function. Optimizing the policy decision at time t is dependent on the output of the reward function from time $t = 0$ to time $t - 1$. Dynamic programming is a method used to find an optimal policy for the MDP. Busoniu *et al.* (2009) [6] constructed a Q-function as the cumulative discounted rewards from time 0 to time t to find the optimal policy. A common methodology used to find the optimal solution is based on the Bellman optimal equations based on the Q-function. The Q-function requires each possible state and action pair to be identified to specify an iterative policy search across all these pairs to optimize the cumulative returns.

The action space underlying the optimization problem in this paper is multidimensional and continuous, or even if a more simplified discrete option is constructed consist of a very large number of possible action states. The Q-function optimization requires the evaluation across all or a large portion of possible states. This together with curse of dimensionality requires a fairly large training dataset to support the optimization.

Reinforcement learning differs from supervised learning in that no target outcome is provided. In supervised learning the MDP is trained to historic or on-line data by minimizing the difference of the target and model outcome. For reinforcement learning the system takes actions based on some policy and receives feedback on the performance based on these actions. The parameters driving the policy are adjusted to increase the reward function. There is no target return or outcome for the optimization.

A number of reinforcement learning methodologies have been applied in the context of automated trading decisions and active portfolio management. Neuneier (1996) [31] developed a Q-learning approach to support a portfolio management approach using on-line reinforcement learning.

A recurrent learning algorithm is a recognized methodology applied to train a MDB that is path dependent. Examples of these algorithms are backpropogation through time, see Werbos (1990) [38] and an on-line learning algorithm called real-time recurrent learning (”RTRL”) set out in Rumelhart *et al.* (1985) [33].

Moody *et al.* (1998) [29] and Moody and Saffel (2001) [28] developed a recursive learning algorithm called Recursive Reinforcement Learning (”RRL”) based on the recursive methodologies from Werbos (1990) [38] and Rumelhart *et al.* (1985) [33] using the Shape ratio (defined as the average return divided by the standard deviation of the return) or differential Sharp ratio as the reward function. This methodology was developed to optimize the return of the portfolio selection framework.

The RRL methodology developed has been used in a number of portfolio selection and rule based trading systems. See Dempster and Leemans (2006) [12], Li,Dagli and Enke (2004) [31], Maringer and Ramtohul (2012) [26], Gorse (2011) [16] and Bertoluzzo and Corazza (2014) [4] for application in automated trading rules. The papers extended the RRL to allow for either uncertainty through a stochastic process, an alternative iterative process compared to the gradient rule or more granularity such as transaction costs and non-stationary data.

Chapter 2

Model setup

The bank will have a funding gap each month as existing funding matures. The size of the funding gap to be filled by new wholesale funding will change each month based on the change in the asset and deposit portfolios and the portion of the existing wholesale funding that matures. The size of the wholesale funding portfolio that mature in a particular month is based on the previous funding decisions. The size of the funding gap and thus exposure to cost of funding volatility is impacted by historic funding decisions. The aim of this section is to parametrize the funding gap and wholesale funding decision available to the bank.

A representation of the monthly **net interest income margin** ("NII") is shown below:

$$NII = X^1 * (x^1 - CL) - X^2 * x^2 - X^3 * x^3 - X^4 * x^4 - X^5 * x^5 - X^6 * x^6 \quad (2.1)$$

where X^1 is an asset portfolio consisting of personal, mortgage and corporate loans.
x^1 is the interest rate received on the assets above.
CL is the credit loss on the assets above.
X^2 is a portfolio of retail and corporate deposits.
x^2 is the interest paid on retail and corporate deposits.
X^i, for $i = 3, 4, 5, 6$ represents the size of the wholesale funding across different durations.
x^i, for $i = 3, 4, 5, 6$ represents the interest rate paid on each instrument.

For the purposes of this paper we considered X^i, for $i = 3, 4, 5, 6$ of duration 6,12,18 and 24 months. The interest earned on the asset portfolio (x^1) is net of the credit loss (CL) for the remainder of this paper. A mathematical equation of the bank's balance sheet at month t is:

$$A_t = L_t + E_t \tag{2.2}$$

where E_t is the level of equity, A_t the assets and L_t the liabilities as at month t.

At the end of each projection period t the asset portfolio reduces due to the monthly capital repayment, maturing loans and incurred credit losses. New loans makes up for this natural reduction in the asset portfolio. We assume the asset portfolio stay constant over the projection period.

The balance sheet extends to the following based on the notation above:

$$X_t^1 = X_t^2 + X_t^3 + X_t^4 + X_t^5 + X_t^6 + E, t \in [1, 60] \tag{2.3}$$

where E is fixed over the projection period.

A portion of the wholesale funding base will mature each month based on previous funding decisions. For example the entire portfolio will mature if only funded via monthly instruments. Let Xm_t^i indicate the portion of the portfolio that mature in month t for each $i = 3, 4, 5, 6$. Define Xm_t^3, Xm_t^4, Xm_t^5 and Xm_t^6 as the wholesale funding instruments maturing in month t.

Assuming the equity level is constant (E_t) the funding gap G_t is a function of the change in the asset portfolio ($X_t^1 - X_{t-1}^1$) a change in the deposit portfolio ($X_t^2 - X_{t-1}^2$) and the sum of all the maturing wholesale instruments (Xm_t^i), where $i = 3, 4, 5, 6$.

$$G_t = X_t^1 - X_{t-1}^1 - (X_t^2 - X_{t-1}^2) + Xm_t^3 + Xm_t^4 + Xm_t^5 + Xm_t^6 \tag{2.4}$$

Each month the bank needs to choose between the various wholesale funding instruments to fill the funding gap. The optimization problem tries to identify the best funding mix by optimizing the NII function.

Let F_t be a vector of the funding decision, $\overline{F_t} = \langle F_t^3, F_t^4, F_t^5, F_t^6 \rangle$ such that F_t^3 represent portion of the funding gap (G_t) to be filled by wholesale instruments X_t^3.

2.1 Sub-models

Figure 2.1 highlights the process followed to apply the two optimization methodologies to optimize the NII as set out in equation 2.1. An economic scenarios generator ("ESG") is used to generate a monthly view of prevailing interest rates for a 60 month projection period. A propriety scenario generator using the methodology set out by Sheldon and Smith (2004) [34] was used. The starting point for this exercise is December 2014. The ESG outputs a 60 month projection horizon of prevailing interest rates for each month from December 2014 to December 2019. The ESG model provided 600 unique scenarios, each projected from December 2014 to December 2019.

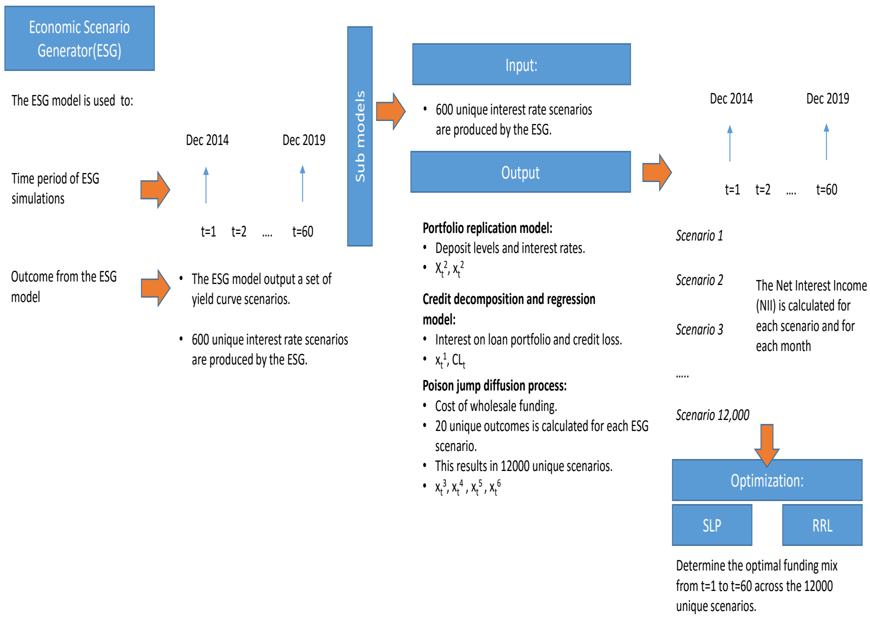
The NII per equation 2.1 is calculated for each of the 600 scenarios, from December 2014 to December 2019. This requires a projection of each of the inputs in equation 2.1 based on the simulated ESG scenario. Various sub-models are used to translate the parameters required per equation 2.1 based on the ESG scenarios. A 5 to 10 year history of data till December 2014 was used to calibrate the various sub-models. The credit loss (CL_t), deposit portfolio behavior (X_t^2, x_t^2) and cost of wholesale funding ($x_t^3, x_t^4, x_t^5, x_t^6$) are projected over the projection period for each of the 600 ESG scenarios. The allows us to calculate the NII per equation 2.1 from December 2014 to December 2019 for each ESG scenario. The optimization models are deployed across the 60 month projection period and scenarios to find the optimal funding decision.

Specifying the sub-models

The sub-models are used to relate the input parameters required to project the NII per equation 2.1 to a yield curve scenario produced by the ESG. The detailed discussion of each sub model is beyond the scope of this paper. The section below provides a brief overview of the models used. The model framework and optimization formulation set out in this paper is agnostic to the sub-model calibrations.

The ESG model per Sheldon and Smith (2004) [34] is arbitrage-free, with calibrations based on the observed or quoted market prices of various instruments. The model satisfies the efficient market hypothesis and for most asset classes assume some type of Ornstein-Uhlenbeck process that is a mean reverting random walk process. See Smith and Speed (1998) [35] for a discussion on the use of deflators in the ESG model.

Figure 2.1: Diagram of the model framework to apply the optimization methods



A portfolio replication model was used to calibrate both the size and interest rate on the deposit portfolio. This is based on deposit data from January 2000 to December 2014. This model is used to project both the size of the deposit portfolio (X_t^2) and the interest rate (x_t^2) at time t per the ESG scenarios. The portfolio replication approach follows the methodology set out by Paraschiv (2011) [32] where the deposit portfolio behavior is represented as a portfolio of risk free assets at various duration.

A regression model was used to calibrate the relationship between the historic credit loss CL_t from January 2007 to December 2014 to prevailing interest rates. This model is used to project the CL_t underlying the asset portfolio for each ESG scenario. The methodology is similar to Havrylchyk (2010) [19] who developed a regression type model to empirically test the impact on the credit loss due to a change in a set of macro-economic variables on the South African banking sector.

A two step projection process is used to project the cost of wholesale funding (x_t^3, x_t^4, x_t^5 and x_t^6). The first is the credit spread paid by the bank over and above the risk free rate, and the second is a liquidity premium. The Leland and Toft (1996) [25] model is used to calculate the credit risk component. The portion of the observed spread not explained by the credit spread is termed the liquidity spread. A poison stochastic jump process was calibrated using historic liquidity spreads from January 2007 to December 2014. This model is used to introduce the large sudden jumps observed in the cost of wholesale funding and thus liquidity risk as part of the funding. The methodology per Bates (1996) [2] is used for the poison stochastic jump process. The poison stochastic jump process calculates the liquidity risk premium and the Leland ad Toft model the credit spread to calculate the cost of funding underlying each of the ESG scenarios. 20 unique paths are produced for each of the 600 ESG simulations across the 60 month projection period.

Per figure 2.1 the SLP and RRL optimization is applied to the 600 scenarios times 20 unique liquidity risk paths. The results in 12000 outcomes projected for 60 months from December 2014 to December 2019. The optimization methodologies are used to determine the optimal mix of wholesale funding given the uncertainty presented via the 12000 scenarios.

Chapter 3

Stochastic Linear Programming

3.1 Eventtree

The computing resources required to solve certain algorithms operating in higher dimensions grow exponentially causing intractable problems (curse of dimensionality). Methods to approximate the continuous nature will attempt to cover only the realizations of the random process that are truly needed to obtain the near-optimal decision. In the case of the stochastic linear optimization problem this is achieved by breaking down the problem to a finite approximation. The event tree is a tool to express the continuous distribution with a simple discrete approximation via a set of nodes and branches see Dupacova *et al.* (2000) [13]. It is important to recognize that the event tree is an approximation of the process only.

There are a number of methods available to construct an event tree. The approach discussed in Gulpnar *et al.* (2004) [17] was used in this paper to calibrate the event tree. This procedure is based on a simulated and randomized clustering approach. The event tree consist of decision nodes and branches originating from the same base. The structure of the event tree supporting this paper is two event branches originating at each node. The sub set of branches created under this structure is independent. Thus moving down from node 1 and up from node 2 will not end in the same position.

The projection horizon supporting this paper is 60 months. This results in $1.152 * 10^{15}$ unique nodes at $t = 60$. This dimension exceed the number of scenarios to calibrate the event tree. To overcome this challenge we partition the 60 month time period into 12 decision time intervals.

3.2 Methodology

The Stochastic Linear Program ("SLP") is used to optimize the NII function per equation 2.1. The optimization decision is focused on the duration mix of funding issued to fill the monthly funding gap G_t^k (see equation 2.4) at time t for scenario k. The subscript notation for the remainder of this section is t for time period and k for the scenario.

The objective is to minimize the funding cost to the bank. The cost impact of the new funding is a function of the current interest rates and the size of the funding gap, where the previous funding decisions drive the size of the funding gap. Choosing mostly long term funding will lock in historic interest rates and reduce the exposure of jumps in funding costs as the funding gap will be smaller. However longer term funding is generally more expensive.

$\overline{F_t^k}$ is the decision vector representing the funding mix $< F_t^{3,k}, F_t^{4,k}, F_t^{5,k}, F_t^{6,k} >$ to fill the gap G_t^k such that $G_t^k = F_t^{3,k} + F_t^{4,k} + F_t^{5,k} + F_t^{6,k}$. The setup needs to be expanded to explicitly allow decisions made in time $t-1$ to influence the optimal decision in time t. To achieve this add $F_t^{7,k}$ to vector $\overline{F_t}$ and to the NII function, where F_t^7 is the sum of all the wholesale funding not maturing in month t. Thus F_t^7 is known based on previous funding decisions. $F_t^{7,k}$ introduce the path dependency of previous decisions. Note $F_t^{3,k} \neq X_t^{3,k}$ as $F_t^{3,k}$ is only the portion of the funding gap filled by the 6 month instruments, where $X_t^{3,k}$ will also include 6 month instruments issued over the last 5 months. The interest rate paid on an instrument relates to the rate as at issue date, thus the rate $x_t^{3,k}$ will only apply to $F_t^{3,k}$. The NII function for the SLP is as follows:

$$NII = X_t^{1,k} * x_t^{1,k} - X_t^{2,k} * x_t^{2,k} - F_t^{3,k} * x_t^{3,k} - F_t^{4,k} * x_t^{4,k} - F_t^{5,k} * x_t^{5,k} - F_t^{6,k} * x_t^{6,k} - F_t^{7,k} * x_t^{7,k}. \tag{3.1}$$

Let the vector $x_t^k :< x_t^{1,k}, x_t^{2,k}, x_t^{3,k}, x_t^{4,k}, x_t^{5,k}, x_t^{6,k}, x_t^{7,k} >$ represent the interest rate earned or paid on the various instruments under scenario k.

Let d_t^k be the outcome at time t for scenario k, where d_t^k represent the change in the deposit funding from month $t-1$ to month t. Thus $d_t^k = X_{t-1}^{2,k} - X_t^{2,k}$. If the level of the deposit portfolios reduce then $d_t^k > 0$ and thus the size of the wholesale funding will increase.

Per Chapter 2 $Xm_t^{i,k}$ is the level of the wholesale funding $i = 3, 4, 5, 6$ to mature in month t, for scenario k. A 6 month instrument issued in month $t - 6$ will mature in month t, thus $Xm_t^{i,k} = F_{t-Mi}^{i,k}$, where Mi is the term of the instrument i. Based on the above definition the gap G_t defined in equation 2.4 summarize as follows:

$$G_t^k = \sum_{i=3}^{6} Xm_t^{i,k} + d_t^k \tag{3.2}$$

Per the model setup the bank needs to fill the funding gap G_t by the funding choice such that:

$$G_t^k = F_t^{3,k} + F_t^{4,k} + F_t^{5,k} + F_t^{6,k} \tag{3.3}$$

From the path dependency discussion above $F_t^{7,k}$ is defined as follows:

$$F_t^{7,k} = \sum_{i=3}^{7} F_{t-1}^{i,k} - \sum_{i=3}^{6} Xm_t^{i,k} \tag{3.4}$$

Let $x_t^{7,k}$ be the interest rate paid on the remaining wholesale liabilities prior to funding the gap in month t. This interest rate is a function of the previous funding decisions and corresponding interest rates that applied, thus is fully computable using information from the previous known outcomes at $t = 1, 2...t - 1$.

$$x_t^{7,k} = \frac{\sum_{i=3}^{6}[F_{t-1}^{i,k} x_{t-1}^{i,k}] - [\sum_{i=3}^{6} Xm_t^{i,k} x_{t-Mi}^{i,k}]}{F_t^{7,k}} \tag{3.5}$$

Define $F_t^{1,k} = X_t^{1,k}$ to be the size of the asset portfolio and $F_t^{2,k} = X_t^{2,k}$ to be the size of the deposit portfolio. This notation is used to support the linear model formulation in F rather than X. The only change in the size of $F_t^{2,k}$ is due to the change in the deposit portfolio, where $F_t^{1,k}$ is constant over time. Thus the following equality holds $F_t^{2,k} = F_{t-1}^{2,k} + d_t^k$.

Formulating the linear model

The NII is formulated in F per equation 3.3, this is formulated in terms of the SLP optimization methodology as:

$$Max(x_t)^T F_t. \tag{3.6}$$

Equation 3.6 is the same as minimizing the cost of funding $\sum_{i=3}^{7} -x_t^i F_t^i$. The expanded form of the linear program can be written as per the L-shape method:

Maximize $(x_t)^T F_t + E_\xi[(x_{t+1})^T F_{t+1} + E_\xi[(x_{t+2})^T F_{t+2}] + ...]$. Where the realization of the random event in stage $t+1$, $t+2$, .. is $\xi \in \Omega$. Applying the master and sub problem per the L-shape the problem simplify to Maximize $(x_t)^T F_t + \theta_t$, where θ_t is iteratively expanded.

The constraints applicable to this linear problem are:

$$F_t^{1,k} = F_{t-1}^{1,k} = X_1 \tag{3.7}$$

$$F_t^{2,k} = F_{t-1}^{2,k} - d_t^k \tag{3.8}$$

$$F_t^{3,k} + F_t^{4,k} + F_t^{5,k} + F_t^{6,k} = \sum_{i=3}^{6} Xm_t^{i,k} + d_t^k \tag{3.9}$$

$$F_t^{7,k} = F_{t-1}^{3,k} + F_{t-1}^{4,k} + F_{t-1}^{5,k} + F_{t-1}^{6,k} + F_{t-1}^{7,k} - \sum_{i=3}^{6} Xm_t^{i,k} \tag{3.10}$$

$$\tag{3.11}$$

The constraints can be written in the form of equation $Wx_t^k = h_t^k - T_t^k x_{t-1}^{a(k)}$. The multi period nested L-Shape algorithm was used to determine the optimal strategy, if feasible.

3.3 Results

Table 3.1 show three trading strategies where F_3 represent the 6 month instruments, F_4 the 12 month instruments, F_5 the 18 month instruments and F_6 the 24 month instruments. The % represents the portion of the funding gap to be filled by the various instruments. Trading strategy 1 is more weighted towards longer dated instruments (mainly 24 month instruments) where strategy 3 focus on short dated instruments. Trading strategy 2 is a mix of the above, however still more weighted towards the longer dated funding.

The SLP optimization methodology is used to select the optimal trading strategy for the bank. The SLP optimization is designed to maximize return only. Other performance metric such as the Sharp Ratio (average return divided by the standard deviation), Value at Risk and Conditional Value at Risk is not considered as part of the SLP optimization. Equation 3.6 can be extended to target other performance metric however a more complex optimization methodology will apply due to the

Table 3.1: Funding strategies

Trading strategy	F_3	F_4	F_5	F_6
Strategy 1	0	0	12.5%	87.5%
Strategy 2	0	12.5%	25%	62.5%
Strategy 3	87.5%	12.5%	0%	0%

non-linearity of the optimization criteria.

The SLP optimization method selected trading strategy 1 as optimal in terms of maximizing the return. The performance of strategy 2 and 3 is shown for comparison purposes only. Short dated debt was cheaper compared to longer dated debt per the model setup. Funding the bank with short dated debt exposes the bank to funding at a very high cost during periods to distress. The SLP optimization methodology selected a longer funding approach to cushion the bank from these liquidity events.

Strategy 1 maximizes the average return over a 60 month projection period and across the 12000 scenarios. The preference to fund the bank with longer dated instruments mitigate the liquidity risk introduced by continuously rolling funding at shorter durations. Table 3.2 show the return distribution for each of the strategies split into 4 buckets for simplicity. Strategy 1 has the biggest portion in the high return bucket, this is the driving force of the superior returns for Strategy 1. This coincide with periods of higher interest rates where the return on assets reprice faster than the cost of funding due to the longer funding duration, confirming the importance of funding at longer durations.

Table 3.2: Strategy 1 has a higher portion in the high return category

Return category	Strategy 1	Strategy 2	Strategy 3
Loss	8.1%	7.1%	7.3%
Low return	23.4%	24.4%	24.3%
Medium return	57.9%	66.4%	65.8%
High return	10.6%	2.2%	2.6%

8% of the outcomes under Strategy 1 results in a loss compared to 7% for strategy 2 and 3. The 95% VAR and CVAR is based on the return of assets instead of the nominal loss. This return should be multiplied with the size of the asset portfolio to obtain an absolute level. This confirms the slightly worst 95% VAR and CVAR for Strategy 1 as shown in table 3.3. The positive skewness in the results distribution results in a higher standard deviation of the return under Strategy 1 impacting the Sharp ratio per table 3.3. A summary of the performance of the three trading strategies across a number of performance metric are shown in the table 3.3.

Table 3.3: Comparison of the performance metric across the three strategies

Trading strategy	Average return	Sharp Ratio	95% VAR	CVAR
Strategy 1	3.1%	5.65	-0.2%	-0.64%
Strategy 2	3.0%	6.56	-0.2%	-0.61%
Strategy 3	3.0%	6.77	-0.1%	-0.52%

The optimal solution is a function of both the scenarios considered and the assumptions on the sub-components such as the credit loss, deposit portfolio behavior and cost of wholesale funding. The impact of choosing a different starting date for the projection and lower liquidity risk in the cost of funding was tested. This resulted in a shorter optimal funding compared to Strategy 1 above.

The power of the above methodology is to isolate specific impacts to facilitate the bank to determine the optimal wholesale funding mix given specific outcomes. We investigated the impact of reducing the liquidity risk via the liquidity premium projection using a poison jump process with less jumps. The optimal strategy approaches the short strategy from table 3.1 as the frequency of the jumps is reduced. This is intuitive as the bank will seek shorter dated instruments which are cheaper if liquidity risk diminishes. This confirms the importance of this tool to assist the bank with scenario planning. A further research topic from this paper is determining the optimal funding strategy under various scenarios and assumptions, isolating the key drivers of specific funding strategies.

Chapter 4

Recurrent Reinforcement Learning

4.1 Methodology

The optimization methodology per section 2 considered 4 durations for wholesale funding. For the purpose of the RRL methodology we simplify this to two durations, namely a 6 and 12 month instrument only. The same projection period, ESG scenarios and sub models to project the NII was used as per the SLP method. As per the SLP optimization the trading decision is made every 6 months. This setup simplify the complexity of the trading decision, the return function and the algebra required to support the RRL optimization methodology. The methodology can be extended to more instruments and monthly trading rules with an increase in the complexity of the solutions; this will also require more data to train the trading function.

The funding gap each month was defined as G_t. Let $\bar{F}_t =< F_t^3, F_t^4 >$ represent the decision vector at time t, where F_t^3 represent the portion of the gap G_t to be filled by issuing 6 month instruments.

The policy is a function with explicit weights to be trained during the reinforcement learning process. For the purposes of this paper the policy function is a trading function shown below:

$$F_t^3 = \tanh(\exp(\theta * (x_t^4 - x_{t-1}^4 - 0.005))) \tag{4.1}$$

where θ is the parameter to be solved and controls the speed of change in the trading rule. See Moody and Saffel (2001) [28] for a discussion on the choice of this trading signal. The choice of the trading function seems fairly arbitrary, however the properties of this function have intuitive appeal. The month on month change in the 12 month interest rate is the main driver of credit losses on the asset portfolio, which in turn drives the probability and the size of the liquidity jumps in the liquidity premium calibration. Due to this relationship we expect the trading strategy to move to a longer duration to protect the bank from liquidity risk that increase during an interest raising cycle. The *tanh* function ensures that F_t^3 is bounded between $[0, 1]$, where the *exp* function allows for a fairly steep change in the trading strategy as Δx_t^4 changes. The θ parameter controls the speed of this change. Per this setup $F_t^4 = 1 - F_t^3$

The NII equation (2.1) present the initial setup of the net interest rate margin, or return function supporting the RRL system. This equation simplify for the RRL application as only 2 types of wholesale funding instruments are used in the RRL method compared to the 4 types in the SLP method:

$$R_t^* = x_t^1 * X_t^1 - x_t^2 * X_t^2 - x_t^3 * X_t^3 - x_t^4 * X_t^4 \tag{4.2}$$

Per this construction optimizing R_t^* is the same as minimizing $R_t = x_t^3 * X_t^3 + x_t^4 * X_t^4$. The return in month t is a function of the previous funding decision X_{t-1}^4 and the current funding decision X_t^4 and X_t^3. This is because X_{t-1}^3 matures by t where X_{t-1}^4 only mature by $t + 1$. Based on this R_t follows as:

$$R_t = F_{t-1}^3 * [x_t^3 * F_t^3 + x_t^4 * (1 - F_t^3)] + x_{t-1}^4 * (1 - F_{t-1}^3) \tag{4.3}$$

The Sharpe ratio is used as the optimization function for the purposes of the RRL optimization. The Sharpe ratio is a well known performance function used in portfolio management as this use both average returns and the standard deviation of these returns. The Sharpe ratio as time t is defined below.

$$S_t = \frac{Average(R_t)}{Std(R_t)}.$$

$$S_t = \frac{A_t}{K_t(B_t - A_t^2)^{0.5}}. \tag{4.4}$$

Where $A_t = 1/t \sum R_t$, $B_t = 1/t \sum R_t^2$ and $K_t = (\frac{t}{t-1})^{0.5}$.

The differential Sharpe ratio is key if an on-line learning algorithm is required. This paper use the differential Shape ratio as the reward signal for the RRL problem. For the differential Sharpe ratio A_t and B_t are defined below.

$$A_t = A_{t-1} + \eta(R_t - A_{t-1}).$$
$$B_t = B_{t-1} + \eta(R_t^2 - B_{t-1}). \quad (4.5)$$

Where η is the adaption rate.

The recurrent reinforcement leaning algorithm aims to maximize S_t using an online learning approach via the differential Sharpe ratio. This is done by adjusting the policy function via the θ from F_t^3 with each time step across all simulations. The weight is updated using the gradient method as discussed in detail in Williams (1992) [40].

$$\Delta\theta = \alpha\frac{dS_t}{\theta} \quad (4.6)$$

where α is the learning rate of the RRL process. The equation for $\Delta\theta$ can be broken down into $\frac{dS_T}{d\theta} = \frac{dS_T}{dR_T} * \frac{dR_T}{d\theta}$. Consider the components in two steps.

First consider $\frac{dS_T}{dR_T}$

As S_t is a function of both B_t and A_t the derivative above can be written as $\frac{dS_T}{dR_T} = \frac{dS_T}{dA_T} * \frac{dA_T}{dR_T} + \frac{dS_T}{dB_T} * \frac{dB_T}{dR_T}$. Using equation 4.5 to define B_t and A_t the derivation follows from algebra.

$$\frac{dS_T}{dR_T} = \eta * \frac{B_{T-1} - A_{T-1} * R_T}{(B_{T-1} - A_{T-1}^2)^{3/2}}. \quad (4.7)$$

Next consider $\frac{dR_T}{d\theta}$

The real-time recurrent learning (”RTRL”) set out in Rumelhart *et al.* (1985) [33] is used for the derivation of the recursive learning algorithm. As per Moody and Saffel (2001) [28] the RRL algorithm is given as $\sum_{t=1}^{T}[\frac{dR_t}{dF_t^3} * \frac{dF_t^3}{d\theta} + \frac{dR_t}{dF_{t-1}^3} * \frac{dF_{t-1}^3}{d\theta}]$. The second term in this equation is required as the return function R_t is a function of the incremental decision, thus both F_{t-1}^3 and F_t^3 directly affect the calculation of the R_t.

Note that the quantity $\frac{dF_t^3}{d\theta}$ is a total derivatives that depend upon the entire sequence of previous trades from time t=0 to t.

The derivation of the first elements is relative straight forward from equation 4.3, $\frac{dR_t}{dF_t^3} = F_{t-1}^3 * x_t^3 - F_{t-1}^3 * x_t^4$ and $\frac{dR_t}{dF_{t-1}^3} = F_t^3 * x_t^3 + (1 - F_t^3) * x_t^4 - x_{t-1}^4$. The derivation of the second element is obtained using the recurrent learning algorithm RTRL.

$$\frac{dF_t^3}{d\theta} = \frac{\partial F_t^3}{\partial \theta} + \frac{dF_{t-1}^3}{d\theta}. \tag{4.8}$$

Where $\frac{dF_0^i}{d\theta} = 0$ and thus the above equation is solved recursively.

The derivative of $\frac{\partial F_t^3}{\partial \theta}$ is shown below:

$$\frac{\partial F_t^3}{\partial \theta} = sech^2(\exp(\theta * (x_t^2 - x_{t-1}^2 - 0.005))) * \exp(\theta * (x_t^2 - x_{t-1}^2 - 0.005)) * (x_t^2 - x_{t-1}^2 - 0.005). \tag{4.9}$$

Figure 4.1 set out the real-time recurrent learning framework. The optimization framework is initiated with a predefined θ per the trading rule per equation 4.1 in step 0. This trading rule is applied across the 12000 unique scenarios to calculate the return at time $t = 1$. The recurrent learning algorithm per equation 4.6 is applied to update θ to obtain the new trading rule updated with the information up to time $t = 1$ (Step 2 per figure 4.1). The new trading rule is applied across the 12000 unique scenarios from time $t = 0$ to obtain the return at time $t = 2$. The recurrent learning algorithm per equation 4.6 is applied to update θ to obtain the new trading rule updated with the information up to time $t = 2$. This process repeats till time $t = 60$. Important to note that the new trading rule will be applied from time $t = 0$ for every step.

4.2 Results

Figure 4.2 show the trading function, tagged with the "optimal" data label, calibrated per the RRL methodology. Per this trading rule the bank would issue 70% short dated and 30% long dated instruments when there is no change in Δx_t^4. The bank would increase the portion short dated instruments if Δx_t^4 is negative, while increasing the long dated instruments if Δx_t^4 is positive.

Similar to the SLP methodology we tested the impact on the trading rules if we reduce the impact of liquidity risk via the probability and size of the jump parameters in the cost of wholesale funding. This trading rule is shown as "Sensitivity 1" in figure 4.2. The reduced impact of liquidity risk will results in the bank continuing

Figure 4.1: Steps in the RRL optimization methodology

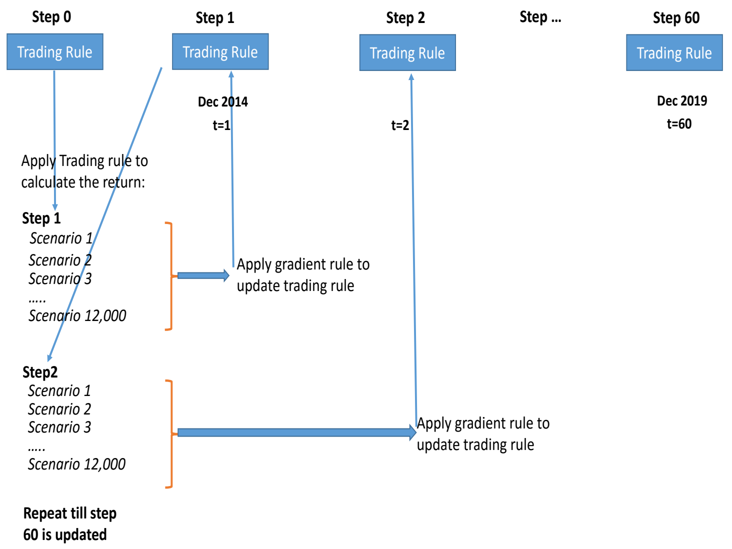
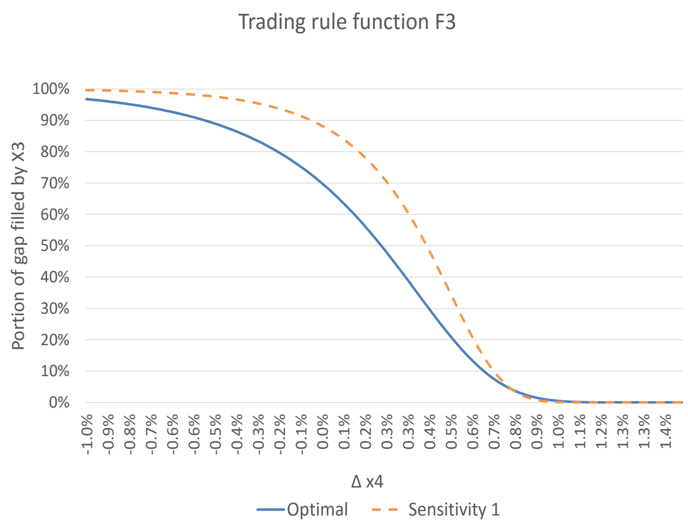


Figure 4.2: Portion of funding gap filled with short dated debt as credit losses change



to issue short dated instruments as credit losses change.

Chapter 5

Results

The SLP optimization aims to define the trading strategy to follow over the entire projection period. The trading strategy is chosen to target the optimal return. The SLP optimization method selected strategy 1 as optimal in terms of maximizing the return. Strategy 1 utilize mainly longer dated instruments to fund the bank. This strategy was selected to minimize the liquidity risk. This confirmed that the introduction of liquidity risk via jumps in the cost of funding of the bank requires the bank to switch funding to longer term instruments.

The RRL method dynamically adjust the trading strategy over the projection period. The credit and liquidity premium paid by banks to issue debt increase as credit losses increase in the underlying bank portfolios. The RRL methodology attempts to capture this dynamic by calibrating the trading rule based on changes in interest rates that drives credit losses. This allows the bank to maintain cheaper funding via short dated instruments when credit losses are low, switching to longer dated instruments to protect against liquidity risk as credit losses start to deteriorate. The RRL methodology provides a higher average return compared to the SLP method.

The trading rule supporting the RRL method was based on a change in interest rates. The calibration of the trading rule resulted in funding with shorter duration instruments when the month-on-month change in interest rates are very small. This switch to longer dated instruments when the interest rates start to increase. The switch is fairly aggressive once beyond a certain point.

Table 5.1 compares the return distribution for the SLP and RRL methodologies, split into 4 buckets for simplicity. The RRL method has a higher portion in the high return bucket with a similar portion in the loss making bucket. Strategy 1 from the SLP method provides superior returns compared to other static funding strategies when liquidity risk are high due to the longer dated funding. The RRL also benefit from this as the trading rules drive longer dated funding as liquidity risk builds up, while focusing on shorted dated instruments during benign periods.

Table 5.1: The RRL method has a higher portion in the high return category

Return category	SLP:Strategy 1	RRL
Loss	8.1%	8.3%
Low return	23.4%	18.7%
Medium return	57.9%	31.8%
High return	10.6%	41.2%

Table 5.2 compares the average return, Sharp ratio ,95% value at risk and CVAR measure for two methods.

Table 5.2: Metric to compare performance of the two methods

Trading strategy	Average return	Sharp Ratio	95% VAR	CVAR
RRL	3.32%	4.33	-0.4%	-0.9%
SLP: Strategy 1	3.07%	5.65	-0.2%	-0.6%

The average NII improved significantly when using the RRL method with the dynamic trading rule. Most notable is the shift in the NII distribution towards higher profits. The positive skewness of the RRL method results in a higher standard deviation and thus lower Sharp ratio. Although the loss distribution has a fatter tail indicating a higher level of large losses than under the SLP optimization (supported by the higher 95% VAR and CVAR).

The scenarios and assumptions supporting the optimization does impact the optimal strategy under both the RRL and SLP methodologies. Choosing a different starting position for the projection and a higher liquidity risk assumptions did results in a different SLP optimal strategy and a dynamic trading rule more weighted towards short dated funding due to the lower liquidity risk. A further research topic from this paper is the determining the optimal funding strategy under various scenarios and assumptions, isolating the key drivers of specific funding strategies.

Bibliography

[1] B. Aouni, C. Colapinto, and D. La Torre, *Financial portfolio management through the goal programming model: Current state-of-the-art*, European Journal of Operational Research **234** (2014), 536–545.

[2] D. S. Bates, *Jumps and stochastic volatility: exchange rate process implicit in deutsche mark options*, The review of Financial studies **9** (1996), 69–107.

[3] J.F. Benders, *Partitioning procedures for solving mixed-variables programming problems*, Numerische Mathematik **4** (1962), 238–252.

[4] F. Bertoluzzo and M. Corazza, *Reinforcement learning for automated financial trading: Basics and application*, Smart Innovation, Systems and Technology **26** (2014), 197–213.

[5] J.R. Birge and D. S. Louveaux, *Introduction to stochastic programming*, Springer (1997).

[6] L. Busoniu, R. Babuska, B. De Schutter, and D. ErnstHull, *Reinforcement learning and dynamic programming using function approximators*, Taylor and Francis (2009).

[7] D.R. Carino, T. Kent, D.H. Myers, C. Stacy, M. Sylvanus, A.L. Turner, K. Watanabe, and W.T. Ziemba, *The russell-yesuda kasai model: An asset liability model for a japanese insurance company using multistage stochastic programming*, Interfaces **24** (1994), 29–49.

[8] D. Chambers and A. Charnes, *Inter-temporal analysis and optimization of bank portfolios*, Management Science (1961), 393–410.

[9] G.B. Dantzig, *Linear programming and extensions*, Princeton University Press (1963).

[10] G.B. Dantzig and P. Wolfe, *The decomposition principle for linear programs*, Operations Research **8** (1960), 101–111.

[11] M.A.H. Dempster, *Introduction to stochastic programming*, Stochastic Progamming (1980), 3–59.

[12] M.A.H. Dempster and V. Leemans, *An automated fx trading system using adaptive reinforcement learning*, Expert System with Applications **30** (2006), 543–552.

[13] J. Dupacova, G. Consigli, and S.T. Wallace, *Scenarios for multistage stochastic programs*, Annals of Operational Research **100** (2000), 25–53.

[14] E. Edirisinghe and E.I. Patterson, *Multi-period stochastic portfolio optimization: Block-separable decomposition*, Annals op Opertaional Research **152** (2007), 367–394.

[15] D.E. Goldberg, *Genetic algorithms in search, optimization and machine learning*, Addison-Wesley (1989).

[16] D. Gorse, *Application of stochastic recurrent reinforcement learning to index trading*, European symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (2011).

[17] N. Gulpnar, B. Rustern, and R. Settergren, *Simulating and optimizing approaches to scenarios tree generation*, Journal of Economic Dynamics and Control **28** (2004), 1291–1315.

[18] S. Guven and E. Persentili, *A linear programming model for bank balance sheet management*, International Journal of Management Science **25** (1997), 449–459.

[19] O. Havrylchyk, *A macroeconomic credit risk model for stress testing the south african banking sector*, South African Reserve Bank **Working Paper 3** (2010), 10.

[20] P. Hilli, M. Koivu, T. Pennanen, and A. Ranna, *A stochastic programming model for asset liability management of a finnish pension company*, Annals of Operational Research **152** (2007), 115–139.

[21] P. Kall, *Stochastic linear programming*, Springer-Verlag (1976).

[22] K. Kosmidou and C. Zopounidis, *Combining goal programming model with simulation analysis for bank asset liability management*, Springer Optimization and Application **18** (2008), 281–300.

[23] R. Kouwenberg and Zenios S.A., *Stochastic programming models for asset liability management*, Handbooks in Finance - North Holland (2001).

[24] M.I. Kusy and W.T. Ziemba, *A bank asset and liability managenet model*, Operations Research **34** (1986), 356–376.

[25] H. Leland and K. Toft, *Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads*, Journal of Finance **51** (1996), 987–1019.

[26] D. Maringer and T. Ramtohul, *Regime-switching recurrent reinforcement learning for investment decision making*, Computational Management Science **9** (2012), 89–107.

[27] S. Marsland, *Machine learning an algorithmic perspective*, Chapman and Hall (2009).

[28] J. Moody and M. Saffel, *Learning to trade via direct reinforcement*, IEEE Transactions on neural networks **12** (2001), 875–889.

[29] J. Moody, L. Wu, Y. Liao, and M. Saffel, *Performance function and reinforcement learning for trading systems and portfolios*, Journal of forecasting **17** (1998), 441 –470.

[30] J. Murphy, Cambridge University Engineering Department Technical Report.

[31] R. Neuneier, *Advances in neural information processing systems*, MIT Press (1996), 925–958.

[32] F. Paraschiv, *Modeling client rate and volumes of non-maturing savings accounts*, University of St. Gallen, School of Management **3924** (2011).

[33] D.E. Rumelhart, G.E. Hinton, and R.J. Williams, *Learning internal representations by error propagation*, Parallel distributed processing: exploration in the micro structure of cognition **1** (1985), 318–362.

[34] T.J. Sheldon and A.D. Smith, *Market consistent valuation of life assurance business*, British Actuarial Journal **10** (2004), 543–626.

[35] A. Smith and S. Speed, *Gauge transformation in stochastic investment modeling*, Proceedings of the 8th International AFIR Colloquium (1998).

[36] R.S. Sutton, *Introduction: The challenge of reinforcement learning*, Springer US (1992).

[37] R. Van Slyke and R.J. Wets, *L-shaped linear program with application to optimal control and stochastic programming*, SIAM Journal on Applied Mathematics **17** (1969), 2638–663.

[38] P.J. Werbos, *Back propagation through time: What it does and how to do it*, IEEE Transactions on neural networks **78** (1990), 1550–1560.

[39] R.J. Wets, A.H.G. Nehauser, K. Rinnooy, and M.J. Todd, *Stochastic programming*, Optimization (Handbooks in Operations Research and Management Science) **1** (1990).

[40] R.J. Williams, *Simple statistical gradient-following algorithms for connectionist reinforcement learning*, Machine Learning **8** (1992), 229–256.

[41] C. Ying-jie and Q. Cheng-iin, *A stochastic programming model of commercial bank asset and liability management*, Journal of Shanghai University **Natural Science** (2000).

[42] S.A. Zenios and W.T. Ziemba, *Handbook of asset and liability management*, North Holland **2** (2007).