Pontifícia Universidade Católica
do Rio de Janeiro

André Frederico Maciel Gutierrez

Goal-Based Investments: A Dynamic Stochastic Programming Approach

Dissertação de Mestrado

Dissertation presented to the Programa de Pós–graduação em Engenharia de Produção of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Engenharia de Produção.

Advisor: Prof. Davi Michel Valladão

Rio de Janeiro
fevereiro 2024

Pontifícia Universidade Católica
do Rio de Janeiro

André Frederico Maciel Gutierrez

Goal-Based Investments: A Dynamic Stochastic Programming Approach

Dissertation presented to the Programa de Pós–graduação em Engenharia de Produção of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Engenharia de Produção. Approved by the undersigned Examination Committee:

Prof. Davi Michel Valladão
Advisor
Departamento de Engenharia Industrial – PUC-Rio

Prof. Bruno Fanzeres Dos Santos
Departamento de Engenharia Industrial – PUC-Rio

Prof. Bernardo Kulnig Pagnoncelli
SKEMA – Business School

Rio de Janeiro, fevereiro the 27th, 2024

André Frederico Maciel Gutierrez

Majored in Industrial Engineering by the Pontifical Catholic University of Rio de Janeiro – PUC-Rio

Bibliographic data

Gutierrez, André F. M.

Goal-Based Investments: A Dynamic Stochastic Programming Approach / André Frederico Maciel Gutierrez; advisor: Davi Michel Valladão. – Rio de janeiro: PUC-Rio , Departamento de Engenharia Industrial, 2024.

v., 38 f: il. color. ; 30 cm

Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Industrial.

Inclui bibliografia

1. Industrial Engineering – Teses. 2. Production Engineering – Teses. 3. Previdência;. 4. Otimização Linear;. 5. Cadeia de Markov Escondidas;. 6. Simulação;. 7. Modelo Financeiro Orientado a Objetivo de Longo Prazo;. I. Valladão, Davi M.. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Industrial. III. Título.

CDD: 620.11

Acknowledgments

Obrigado aos meus pais pelo incentivo e pelo exemplo iniqualável, o trabalho pertence a vocês.
Obrigado a todos meus familiares pelo apoio e aos meus amigos por tornarem tudo mais leve.
Obrigado a todos meus professores, em especial meu orientador, Davi Valladão, pela incansável paciência.
This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001

Abstract

Gutierrez, André F. M.; Valladão, Davi M. (Advisor). **Goal-Based Investments: A Dynamic Stochastic Programming Approach**. Rio de Janeiro, 2024. 38p. Dissertação de Mestrado – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

The aim of this study is to develop an investment policy that minimizes the total contribution required to achieve a long-term financial objective. To achieve this goal, we developed a multi-stage optimization problem that integrates a Hidden Markov Model to capture the stochastic dynamics of asset returns. Unlike conventional portfolio optimization models which are based on unrealistic assumptions, our approach is based on the goal-oriented investment framework which provides a more practical and effective solution. In addition, by using the Hidden Markov Model in our optimization process, we obtain a more accurate estimate of the dynamics of asset returns, which translates into better investment decision-making. By using our model, the contribution required to achieve a desired financial goal is minimized through an investment policy that considers current levels of wealth and prevailing economic conditions.

Keywords

Social Security; Linear Optimization; Hidden Markov Model; Simulation; Goal Oriented Investment

Resumo

Gutierrez, André F. M.; Valladão, Davi M.. **Política de Investimento Orientada a Objetivo de Longo Prazo**. Rio de Janeiro, 2024. 38p. Dissertação de Mestrado – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

O objetivo deste estudo é desenvolver uma política de investimento que minimize a contribuição total necessária para atingir um objetivo financeiro a longo prazo. Para atingir este objetivo, desenvolvemos um problema de otimização multi-estágios que integra um modelo de Markov oculto para captar a dinâmica estocástica dos retornos dos ativos. Ao contrário dos modelos convencionais de otimização de carteiras, que se baseiam em pressupostos irrealistas, a nossa abordagem baseia-se no quadro de investimentos orientado a objetivos, que proporciona uma solução mais prática e eficaz. Além disso, ao utilizar o modelo de Markov oculto no nosso processo de otimização, obtemos uma estimativa mais precisa da dinâmica dos retornos dos ativos, o que se traduz numa melhor tomada de decisões de investimento. Ao utilizar o nosso modelo, a contribuição necessária para atingir um objetivo financeiro desejado é minimizada através de uma política de investimento que tem em conta o estado atual da riqueza e as condições economicas prevalecentes.

Palavras-chave

Previdência; Otimização Linear; Cadeia de Markov Escondidas; Simulação; Modelo Financeiro Orientado a Objetivo de Longo Prazo;

Table of Contents

1	Introduction	**11**
1.1	Relevant Literature	13
2	Proposed Model	**16**
2.1	Optimization Model	16
2.2	Uncertainty Characterization	20
3	Solution Methodology	**23**
3.1	Solution Methodology	23
4	Empirical study	**27**
4.1	Case Study Set Up	27
4.2	Case Study Results	29
4.3	Sensitivity Analysis	31
4.4	Benchmarking against Traditional Approaches	34
5	Conclusion	**36**
Bibliography		**37**

List of Figures

Figure 1.1	One Stage Iteration	12
Figure 1.2	HMM States In the Historical $S\&P500$	13
Figure 2.1	Objective Functions Set that Prevents Underachievement	17
Figure 2.2	Objective Function that Prevents Overachievement	18
Figure 2.3	Objective Function with Wrong Calibration	18
Figure 2.4	Stochastic Iteration Between Stages	19
Figure 2.5	Stochastic Process Between States	21
Figure 2.6	Distribution Conditional to States	22
Figure 3.1	Future Cost Function Estimation	24
Figure 4.1	Costs Function per Stage	29
Figure 4.2	Allocation Policy - State = 1	30
Figure 4.3	Allocation Policy - State = 2	30
Figure 4.4	Allocation Policy - State = 3	31
Figure 4.5	Allocation Policy - State = 4	31
Figure 4.6	Contribution Sensitivity to λ_1	32
Figure 4.7	Wealth Trajectory Sensitivity to λ_1	33
Figure 4.8	Less Conservative Model	34
Figure 4.9	Comparative Funds Trajectories	35
Figure 4.10	Comparative Contributions	35

List of Tables

Table 1.1 Model Outputs Across Various Scenarios 11

Table 4.1 Markov Text P-Values 28
Table 4.2 HMM Transition Matrix 28
Table 4.3 Parameters of Log-Normal Distribution by State 28

List of Notations

W_t – Wealth at stage t.

$W_{t+1}(s)$ – Wealth calculated at stage t considering returns from scenario s.

$\bar{W}_{t+1}(d)$ – Point of Initial wealth at stage t+1 defined as a new solution.

$Q_t^j(W_t)$ – Cost at stage t and state j as a function of initial wealth.

c_t – Total Contribution at stage t.

c_t^+ – Positive contribution at stage t.

c_t^- – Negative contribution (withdrawal) at stage t.

p_s – Probability of a given scenario.

$p_{j,k}$ – Probability of going from state j to k.

$\theta_{s,k}$ – Auxiliary variable.

q – Free risk rate.

M – Contribution upper limit.

$\pi^k_{t+1,\bar{W}_{t+1}(d)}$ – Dual from next problem (t+1) at a particular point, $\bar{W}_{t+1}(d)$, in state k.

G – Financial objective, goal.

x_t – The amount of wealth in each asset at the stage t. (an array)

d – Tax, duty to be paid.

1
Introduction

In the realm of financial planning, a highly promising investment strategy emerges in the form of goal-based investing. This innovative approach harnesses the tools and techniques of financial engineering to deliver personalized assistance on a broader scale, tailored to the unique needs and aspirations of individual investors. At its essence, this strategy revolves around the pursuit of any specific financial objectives, ranging from retirement savings to home purchases and funding a child's education. The investment plan is meticulously crafted with these goals in mind, considering essential factors such as time horizon, risk tolerance, and requisite savings.

With our model, we have a clear mission: to craft an investment policy that provides actionable guidance, flexibly adjusting to the ever-changing economic landscape and our evolving wealth status as we pursue our financial goals. For example, in Table 1.1 we display the model output for three cases, considering a 10-year time horizon: one for an individual who has reached 30% of their financial goal, another for someone at 50%, and a third for an individual at 80%. For each case, we reveal the allocation in the $S\&P500$ according to different possible economic states identified by our model, one representing a crisis situation ("Bear Market"), one representing a stable period ("Neutral Market"), and another representing a prosperous moment ("Bull Market").

Wealth	Bear Market	Neutral Market	Bull Market
30 % of Goal	0%	100%	100%
50 % of Goal	0%	86%	100%
80 % of Goal	0%	54%	100%

Table 1.1: Model Outputs Across Various Scenarios

The present study offers a novel approach to personalized investment policy that is tailored to individual objectives and financial resources. Specifically, for all $t = 1, ..., T-1$ we formulate an optimization problem that utilizes a multi-stage stochastic optimization framework

$$Q_t(W_t) = \min_{c_t, \mathbf{x_t} \in X} c_t + \mathbb{E}[Q_{t+1}(W_{t+1}(\mathbf{x_t}, \mathbf{r_{t+1}}))],$$

where for the last stage ($t = T$) we have $Q_T(W_T) = \|W_T - G\|$. The model minimizes the contribution at time t, c_t, required to achieve a pre-determined financial goal G, while co-optimizing investment decisions $\mathbf{x_t}$ and computing the one step ahead wealth $W_{t+1}(\mathbf{x_t}, \mathbf{r_{t+1}})$ based on market returns simulations. In Figure 1.1 we show how those variables interact at any stage t.

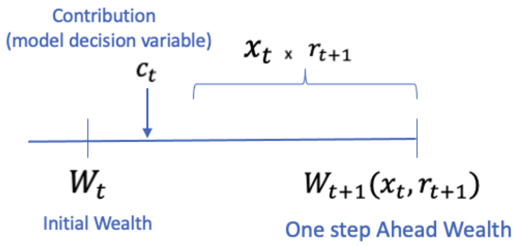

Figure 1.1: One Stage Iteration

Unlike conventional investment models that focus on portfolio variance as a measure of risk, the goal-based approach emphasizes risk as the probability of not achieving the desired financial outcome. Specifically, in our optimization model, we always want to reach the target but we search to minimize the contribution needed. The proposed optimization model thus offers a novel and effective tool for a personalized investment policy that aligns with individual goals and resources.

This study presents a pioneering approach to investment modeling that exploits machine learning techniques to enrich the accuracy of asset return dynamics. To be precise, we incorporate Hidden Markov Models (HMMs) to capture the empirical stylized facts of asset returns, including excess kurtosis and clustered volatility, as well as different economic regimes such as bull and bear markets. Through this method, the study develops a more nuanced and precise understanding of asset return dynamics, enabling more accurate and effective investment modeling.

The use of HMMs offers a sophisticated and flexible framework that can adjust to changing market conditions and capture complex patterns in asset return behavior. As an example, in Figure 1.2 we have a historical series of the $S\&P500$ where the HMM defines four different economic states, according to market returns and volatility. By blending different distributions according to the economic state, this framework can incorporate different regimes and account for the inherent uncertainty in asset returns. Overall, this study represents a significant advancement in goal -based investment modeling, offering a powerful tool with simple and direct outputs for individuals seeking to achieve their financial objectives.

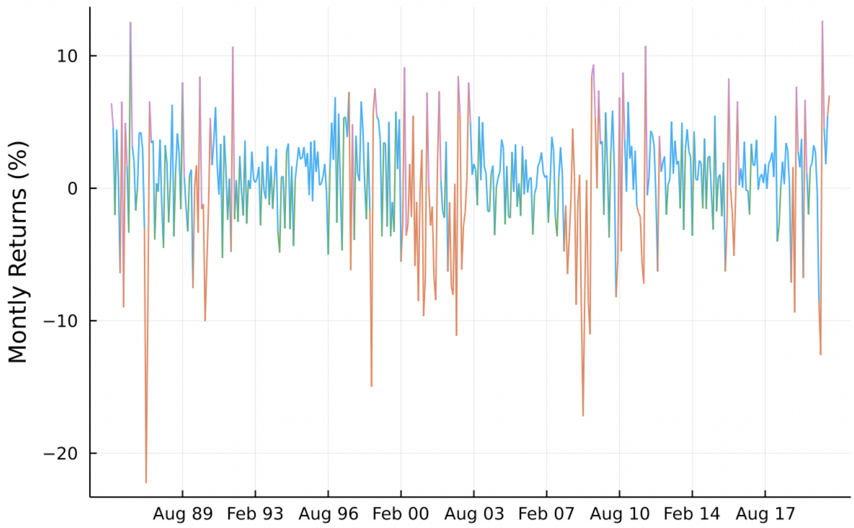

Figure 1.2: HMM States In the Historical $S\&P500$

1.1 Relevant Literature

The renowned modern portfolio theory approaches risk as portfolio variance or standard deviation, and looks to minimize it. The reward is defined as the expected return, which is maximized accordingly. This approach is based on the assumptions that investors are risk averse, and that there exists a linkage between risk and returns. These concepts form the bedrock of the mean-variance framework, as originally introduced by Markowitz [1], and have since been the subject of considerable research.

Despite the substantial body of literature in this area, which we will explore in further detail, it cannot be directly applied to the realm of goal-based investing. In general, Mean-Variance (MV) framework application in long-term investments is flawed thanks to one precarious assumption: it does not consider a time horizon. Consequently, it does not incorporate inflation risk in the treatment of financial wealth and it is harder to define an end goal. Recent research has attempted to address this limitation by incorporating an optimal investment strategy with protection against inflation. For instance, [2] considers stochastic inflation and a broader model where all assets may entail some degree of risk.

Various researchers have addressed the portfolio problem from different perspectives while operating within the same framework. As an example, a multi-period mean-variance model that accounts for stochastic income and mortality risks is presented by [3]. In another study, [4] considers both security returns and salaries as uncertain variables. Furthermore, recent literature has attempted to enhance the robustness of the MV framework by incorporating ambiguity in the decision problem, such as manager ambiguity [5] or by integrating ambiguity aversion and stochastic volatility and income into the model [6].

Another significant criticism of the MV framework is its optimal precommitment solution being inconsistent over time. To address this issue, some researchers have aligned this criterion with the game-theoretic framework, as in [7], [8], and [9], but have found suboptimal solutions. It is also essential to note that the classical MV framework assumes a constant level of risk and return in the financial market, which is not consistent with empirical evidence. As [10] have pointed out, real returns in the next ten years are highest when stocks have the lowest valuations relative to earnings or dividends and vice versa. This observation, known as mean-reversion, has been incorporated into some papers, such as [11] and [12].

Despite numerous attempts to enhance the Mean-Variance (MV) framework and improve its applicability over time, it is adopted mainly by institutional investors [13] and thus still overlooks the constraints faced by individual investors who have specific financial requirements and liabilities to consider. The goal-based framework was introduced to address these shortcomings, with a primary focus on applying portfolio theory to individual investors and their specific needs. In contrast to the MV framework, goal-based investors define risk as the possibility of not achieving their desired goals, allowing risk to be more precisely defined using mathematical models, as can be seen in [14].

Numerous studies have delved into the goal-based framework. For instance, [15] sets a target rate of return tailored to investors' goals and contributions, dynamically adjusting it across various scenarios. Additionally, [16] and [17] aim to fulfill multiple financial objectives by hierarchizing them. Lastly, [13] introduces considerations of both present and future cash flows alongside prioritized financial goals.

A goal-based investing plan serves as a valuable tool for tackling concrete challenges in investment management, offering flexibility to set any desired objective. One particularly significant application lies in pension scheme management, which has encountered formidable obstacles globally. Historically, many pension schemes adhered to a defined benefit (DB) structure, guaranteeing predetermined payments to employees based on earnings-related formulas. However, sustaining these schemes has proven arduous for numerous emerging and developed nations, prompting calls for reform [**?**].

Currently, a defined contribution approach appears promising for addressing these challenges. Despite hurdles such as the imperative for enhanced market design and regulatory infrastructure [19], goal-based investing plans emerge as a practical solution by establishing retirement savings targets. This framework facilitates the development of cost-effective, personalized investment strategies which are crucial for pension scheme sustainability.

2
Proposed Model

2.1
Optimization Model

Our investment problem revolves around minimizing future contributions while addressing numerous uncertainties stemming from factors like market returns and future wealth levels across multiple stages. As we will discuss further in Section 3.1, solving multi-stage problems presents a considerable challenge, and there is no straightforward approach.

Broadly speaking, our methodology entails approximating the objective function in the final stage through a first-order Taylor approximation. Subsequently, we utilize this approximation to construct the future cost function for the preceding stage. This iterative process continues until we reach the first stage, enabling us to optimize the overall problem in a simplified fashion.

Beginning with the last stage (T) we have a particular optimization problem given by Equation 2-1, with the following variables:

1. c_T^+: non-negative contribution at stage T, a decision variable.
2. c_T^-: non-negative withdrawal at stage T, a decision variable.
3. G: Goal to be achieved in the last stage T, given as an input to the model.
4. W_T: initial wealth at stage T, given as an input to the model at this particular stage.

$$\begin{aligned} Q_T(W_T) = \quad & \min_{c_T^+, c_T^- \geq 0} \quad \lambda_1 \; c_T^+ - \lambda_2 \; (1-d) \; c_T^- \\ & \text{s.t.} \qquad G = W_T + c_T^+ - c_T^- \end{aligned} \tag{2-1}$$

The optimization problem defined is a straightforward linear program that can be efficiently solved using standard optimization techniques. Our constraint ensures that the goal G is met by contributing the necessary amount in the final stage. Additionally, in this last stage, we do not need to estimate returns or consider the state of the economy.

As revealed by Equation 2-1, our objective is to minimize the non-negative contributions, subject to a penalty factor λ_1, minus the non-negative withdrawals, subject to λ_2, while attaining the financial goal. Additionally, we account for a tax rate of d, which applies when cashing out from our portfolio.

The choice of penalty values holds paramount significance. In most cases, we want a model geared toward averting underachievement, preventing a massive contribution in the last stage, even if it means the possibility of contributing excessively along the way. Accordingly, we want to strike a balance between the penalties λ_1, and λ_2 to ensure that the costs incurred from non-negative contributions outweigh the potential benefits of withdrawals. This is done by fixing $\lambda_1 = 1$ and constrain λ_2 within the interval $]-1/(1-d), 1/(1-d)[$. In Figure 2.1 we represent the range of possible objective functions considering such a set.

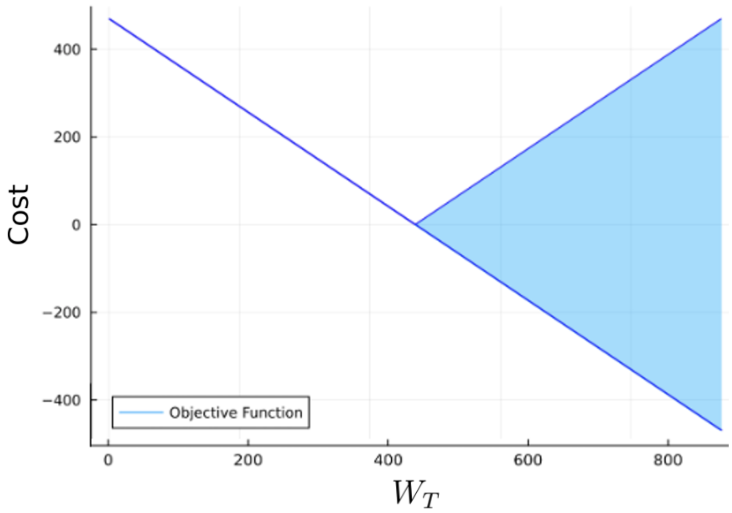

Figure 2.1: Objective Functions Set that Prevents Underachievement

Conversely, when we keep $\lambda_1 = 1$ while setting λ_2 lower than $-1/(1-d)$, we aim to circumvent overachievement at the cost of potentially contributing a great deal in the last stage. A possible function in this set is depicted in Figure 2.2, which assigns a higher cost to withdrawals compared to contributions.

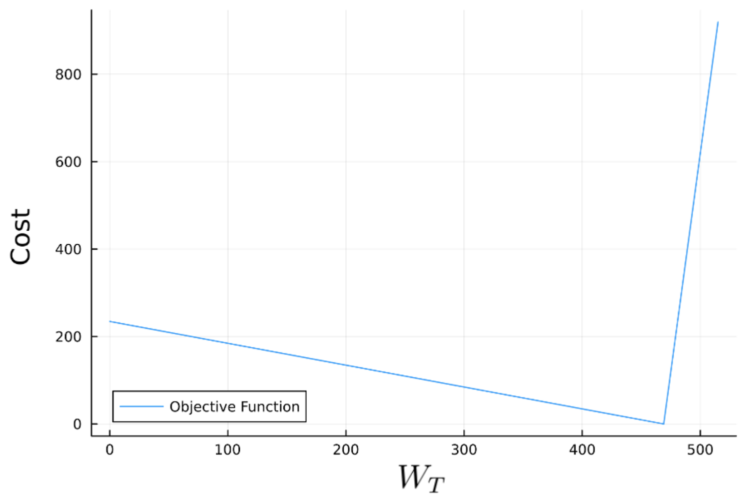

Figure 2.2: Objective Function that Prevents Overachievement

Lastly, if λ_2 surpasses $1/(1-d)$ with λ_1 remaining positive and below one, it signifies an incorrect penalty choice, as our objective function would prioritize overachieving the target to an extreme extent. In other words, the model would maximize the wealth generation at any cost trying to withdraw the most in the last stage. This case is visually represented by Figure 2.3.

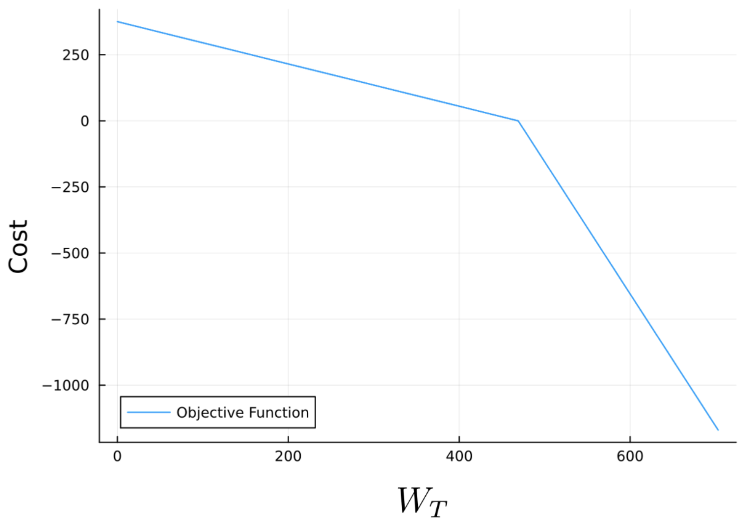

Figure 2.3: Objective Function with Wrong Calibration

After estimating stage T, we have the future cost function of stage T-1. Since our objective is to minimize the expected cost required for the entire investment journey, we take a backward approach: the optimization process begins in the intermediate stages.

We simulate returns for each stage using a Hidden Markov Model, which generates diverse samples based on different states. These states are modeled to correspond with different possible economic conditions and offer a significant advantage by introducing a time dependency to our problem. The HMM methodology will be explained in detail in Section 2.2.

Consequently, for the intermediate stages, we estimate a cost function denoted as $Q_t^j(W_t)$. This function is dependent on the initial wealth in the stage, represented by W_t, as well as the Hidden Markov Model (HMM) state (j) and the stage (t). As exemplified in Figure 2.4, given a function $Q_t^j(W_t)$, we simulate a sample of returns (S_k) for each possible state k of the n possible Markov states.

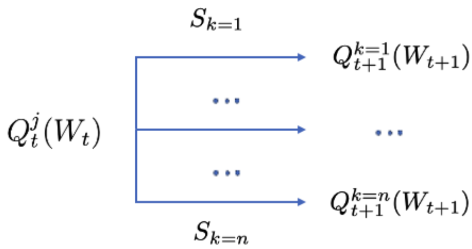

Figure 2.4: Stochastic Iteration Between Stages

As we can see at Formulation 2-2, we minimize the objective function by determining optimal values for the variables c_t^+, c_t^-, and for the wealth allocated to each asset, given by the array x_t. The problem constraints ensure the balance of the contributions and withdrawals with the initial wealth W_t and the amount invested in each financial asset within the set I. Furthermore, they also limit the value of c_t to a maximum value M. However, solving such a problem is not an easy task since calculating the expected value of future cost functions is not straightforward. As explained in Section 3.1, we have a particular methodology do so.

$$
\begin{aligned}
Q_t^j(W_t) = \min_{c_t^+, c_t^- \geq 0, x_t} \quad & c_t + \sum_k E[Q_{t+1}^k(\sum_I x_{i,t}(1 + r_{i,t+1}(s_k))) \mid K_{t+1} = k] * p_{j,k})] \\
\text{s.t} & \sum_I x_{i,t} - (c_t^+ - c_t^-) = W_t \\
& c_t = c_t^+ - (1-d)\ c_t^- \\
& c_t \leq M
\end{aligned}
\tag{2-2}
$$

2.2 Uncertainty Characterization

Given the nature of the uncertainty modeled in our problem, it is paramount to consider the "stylized facts" associated with asset returns. These are fundamental traits that have been consistently observed across various independent studies in the field, and the Hidden Markov Model has demonstrated promising outcomes in modeling them. As stated in [24], the prevailing observations of financial asset returns include:

1. Daily asset returns generally do not exhibit significant autocorrelation, but longer time intervals such as weeks or months may show some degree of autocorrelation.

2. The unconditional distribution of asset returns often displays a heavy tail. This implies that extreme events like market crashes or booms occur more frequently than predicted by a normal distribution. To deal with this, we use a normal log-return variable in our modeling.

3. High-volatility events tend to cluster over time, and volatility shows positive autocorrelation over a certain period. Therefore, financial returns lack independence. The HMMs effectively introduce time dependency into our model and can handle the complex volatility patterns in financial data by utilizing a mixture of distributions.

The Hidden Markov Model is a well-established probabilistic framework that can be effectively utilized for generating observations from a specific distribution, contingent upon the underlying state of an unobserved process. The process is typically represented by a Markov chain, with a discrete, first-order Markov chain being the focus of our discussion. In this regard, the likelihood of a state at any given time step depends solely on the immediately preceding state within the chain. Considering q_t and the state S that we are at a moment t, the stochastic process is represented in Equation 2-3.

$$Pr(q_{t+1} = S_j | q_t = S_i, ..., q_1 = S_u) = Pr(q_{t+1} = S_j | q_t = S_i) \tag{2-3}$$

Given that the probability of a particular state within a Markov chain is solely contingent upon the preceding state, we can represent the state transitions using a state transition probability $a_{i,j}$. This parameter denotes the probability of transitioning from a state i to another state j, as outlined in Equation 2-4.

$$a_{i,j} = Pr(q_t = S_j | q_{t-1} = S_i) \tag{2-4}$$

As an illustrative example, suppose we consider an economy with three distinct states, denoted as S_1 representing a bull market, S_2 denoting a bear market, and S_3 signifying a neutral market. Considering the state transition matrix $\mathbf{A}$, we can gain insight into the underlying Markov process between these states, as elucidated in Figure 2.5.

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}$$

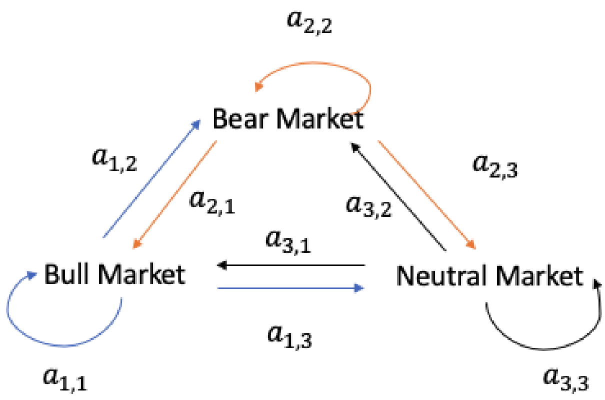

Figure 2.5: Stochastic Process Between States

As we are operating within the context of a Hidden Markov Model, it is important to note that two distinct stochastic processes are at play. The first process pertains to the underlying, unobserved states, as discussed above, and the second process pertains to the observations made given a specific state. We can represent this bivariate stochastic process as (S_t, X_t), where X_t represents the observation generated at time t given a state S_t. In the context of our concrete example, S_t would denote the state of the economy, and X_t would signify the generated asset's returns. This process is succinctly described in Equation 2-5 and exemplified in Figure 2.6.

$$b_t(j) = Pr(X_t | X_{t-1}, q_t = S_j) = Pr(X_t | q_t = S_j) \tag{2-5}$$

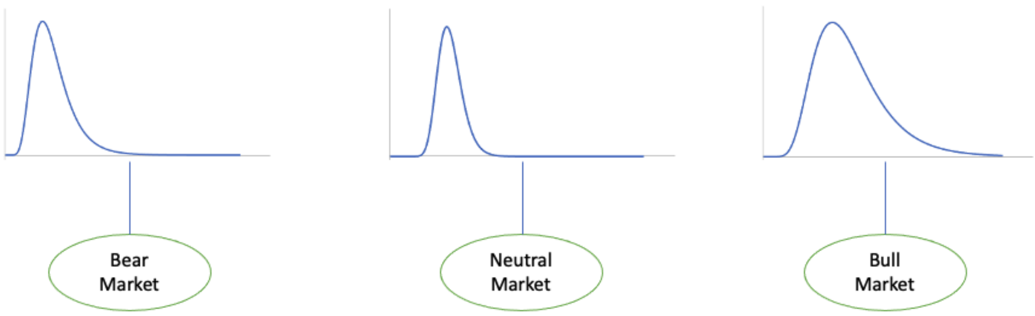

Figure 2.6: Distribution Conditional to States

To summarize, in a Hidden Markov Model (HMM), the transition probabilities between states are described by a transition matrix denoted by **A**. This matrix is determined by the probabilities of transitioning from one state i to another state j at each time step t, denoted by $a_{i,j}$. The conditional probability distribution of the observations for each state j is given by $b_t(j) = Pr(X_t|q_t = S_j)$, which represents the likelihood of observing an output X_t given that the system is in state j at time t. Additionally, a complete HMM will be defined if the initial distribution $\pi_i = Pr(q_1 = S_i)$ is specified, which provides the first state of the sequence.

In this work, we will use a log-normal distribution with varying scale and mean according to the state that we are in. However, determining the appropriate number of states to use in the HMM is a critical decision. As our study does not primarily focus on developing a methodology for determining the number of states, we employ a well-established approach known as the Markov Test.

Given a market return at time t, denoted as r_t, and a specified coverage rate denoted as p, such that $Pr_{t-1}(r_t \leq -VaR(p)) = p$, we establish a sequence of violations using a binary variable, I_t, defined as follows: $I_t = 1$ if $r_t \leq -VaR_t(p)$ and $I_t = 0$ otherwise. Consequently, we introduce a variable $\pi_{l,j}$ with $l = I_{t-1}$ and $j = I_t$, representing the probability associated with a particular sequence of violations. As an example, $\pi_{1,1}$ would represent the probability of two consecutive violations in the sequence.

The Markov test then considers the null hypothesis $\pi_{0,1} = \pi_{1,1} = p$, meaning that the probability of a violation equals the coverage rate and that they are independent of each other. If the null hypothesis is rejected, then the HMM is not well fitted.

3 Solution Methodology

3.1 Solution Methodology

A commonly used modeling approach to represent multi-stage programming is the discrete event tree, which captures the decision variables at each stage, the different possible realizations of uncertainty, and the resulting scenarios. As discussed in [21], the tree reveals the optimal decision to be made at each stage for all scenarios. Thus it provides a comprehensive framework for analyzing the decision-making process considering uncertainty and the development of effective strategies. In other words, in our case, it contains a developed investment policy.

As is well-known, when tackling large multistage problems, discretizing the decision variable into a set of values, and solving the following-stage problem for each of these values is not a viable approach. As pointed out in [22], this method leads to an exponential increase in the number of combinations, even for a small number of variables. This is easy to grasp when we think about the tree structure being developed.

Furthermore, as the number of stages increases, approximating the future cost function becomes increasingly challenging and necessitates advanced techniques. This challenge is commonly referred to as the "curse of dimensionality", a hurdle often surmounted by making approximations and assuming the independence of the stochastic process for each stage, as proposed by [22].

However, our multi-stage problem does not assume stage-wise independence. Instead, it assumes Markov Dependence. In such a case, the probability of a scenario occurring adjusts according to the state given by the HMM, and, since we are working with a first-order Markov Process, the probability of those states occurring remains constant. Therefore, all the scenarios are accessible and depend only on the previous state. In any case, we still need to approximate the future cost function.

To make this estimate, we can represent $Q_t^j(Wt)$ through a first-order Taylor approximation since the last stage problem, like the intermediate one, is convex. Therefore, we need solutions to the problem at specific points (i.e. initial wealth, W_t) and their respective duals. The estimation precision improves progressively as we increase the number of points used. In theory, if we did not have computational limits, we would use an infinite number of points, achieving an exact representation of the function.

For an accurate approximation of the future cost function, selecting the appropriate set of points to construct the cuts is crucial. The most relevant ones are the initial wealth (W_t) that gives different decision variables solutions. Since our problem is convex, we can use the strong duality theorem to perform a sensitivity analysis, providing precisely those values. Considering all those points for constructing the problem cuts would yield an exact representation of the future cost function, once they contain all the possible new solutions to the problem. However, we only utilize a subset of these values due to potential computational constraints.

As an example in Figure 3.1, we illustrate a first-order Taylor approximation to the cost function for a particular stage, represented in red. The estimation is made by selecting the maximum among several linear functions.

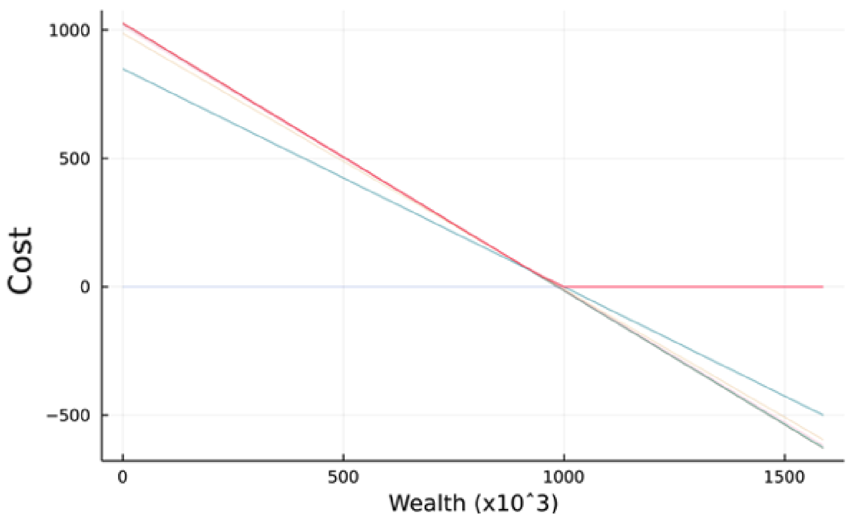

Figure 3.1: Future Cost Function Estimation

Hence, we can compute our future cost function as the expected value of a variable θ which is dependent on scenarios denoted as s drawn from a sample set S_k according to the Markov Model state k. This state k follows from a prior state j with a probability denoted as $p_{j,k}$. This is calculated by $\sum_K \left(\sum_{S_k} \frac{p_s \theta_s}{1+q} \right) \times p_{j,k}$ where we incorporate the term $1 + q$ to represent the adjustment of the future cost to its present value.

The θ_s value is used in the first-order Taylor approximation, so we ensure it remains above the cutoffs that represent the linear functions corresponding to the cost function for each possible HMM state (k). In those restriction we consider a grid (D_k), that provides a solution ($Q^k_{t+1}(\bar{W}_{t+1}(d))$) and its associated dual variable, denoted as $\pi^k_{t+1,\bar{W}_{t+1}(d)}$, from the next stage problem at the selected point $\bar{W}_{t+1}(d)$. It is important to note that this grid is constructed with the sensibility analysis made in the following stage and it gives the cutoffs used in the estimation. Finally, we also compute the future wealth given by $W_{t+1}(s)$ according to our allocation and the return $r_{t+1}(s)$ simulated. Given our methodology, we arrive at the following formulation for intermediate states described at the Problem 3-1.

$$
\begin{aligned}
Q^j_t(W_t) = \quad & \min_{c^+_t, c^-_t, x_t, \theta_s} \quad c_t + \textstyle\sum_K \left(\sum_{S_k} \frac{p_s \theta_s}{1+q} \right) \times p_{j,k} \\
s.t \quad & \\
& W_{t+1}(s) = \sum_I x_{i,t}(1 + r_{i,t+1}(s)) \qquad \forall s \in S_k, \forall k \in K \\
& \theta_s \geq Q^k_{t+1}(\bar{W}_{t+1}(d)) + \pi^k_{t+1,\bar{W}_{t+1}(d)}(W_{t+1}(s) - \bar{W}_{t+1}(d)), \\
& \qquad\qquad \forall s \in S_k, \forall d \in D_k, \forall k \in K \\
& \theta_s \geq 0, \qquad \forall s \in S_k, \forall k \in K \\
& \textstyle\sum_{i \in I} x_{i,t} - (c^+_t - c^-_t) = W_t \\
& c_t = c^+_t - (1-d)\ c^-_t \\
& c_t \leq M \\
& c^+_t, c^-_t \geq 0
\end{aligned}
\tag{3-1}
$$

In conclusion, our approach can be encapsulated within Algorithm 1. This algorithm has the desired financial goal, the maximum allowable contribution per stage, the number of stages remaining until retirement, and the fitted Hidden Markov Model (HMM) as inputs. By following our previously described backward approach, which involves both our intermediary and final stage models, we construct stage-specific cost functions based on the initial wealth available at each respective stage. By employing these cost functions, we determine the investment policy, specifying the contributions and allocations required at a given moment.

Algorithm 1 Financial Planning Algorithm

Input: Goal, Contribution Upper limit, Number of stages, Fitted HMM
Output: Investment Policy

```
for t ← T to 1 do
    if t == T then
        Run Last Stage problem for each selected point, start with W_T = 0.
        Store cuts from stage T (solutions and respective duals at those points).
        Store model's optimized variables for each selected point.
    else
        for j ← 1 to K do
            Simulate returns from HMM distribution in state j.
            Run intermediate problem for each selected point, starting with W_t = 0 (use all cuts from stage t + 1 and the returns simulated).
            Store cuts from stage t with HMM j
            Store model's optimized variables for each selected point.
        end for
    end if
end for
return The decision variables optimized for each state, stage, and selected point.
```

4
Empirical study

4.1
Case Study Set Up

This case study aims to demonstrate the model's capabilities by employing it to craft a pension plan. We want to formulate a policy that will enable individuals making monthly contributions to secure sufficient funds for retirement. The allocation considers two investment options: the $S\&P500$ and a monthly LIBOR rate, a safety-focused interest rate measure.

To achieve this, our HMM is estimated based on the historical performance of the asset's real returns since 1986. In this way, we account for inflation, ensuring planned safeguards against rising prices.

The model's parameters are calibrated using a real-world scenario. We factor in a tax rate (d) of 15% on withdrawals and opt for a conservative approach, prioritizing a higher likelihood of withdrawing funds over contributing more in later stages. As our preference is to surpass the target rather than fall short of it, we assign values of $\lambda_1 = 10000$ and $\lambda_2 = 1/(1-d)$.

Our individual, aged 25, plans to retire at 65, entailing 480 monthly contributions and allocations. We limit these contributions at \$3000 and set the goal as the amount needed to ensure a consistent monthly payout of 70% of their pre-retirement income (assumed to be \$15000) until they reach the average life expectancy in Brazil.

Determining the goal involves calculating the present value (PV) at the retirement date of the required pension payments, as shown in Equation 4-1. Here, PMT represents the payment made in each period, q denotes the discount rate (calculated as the historical LIBOR average), and n signifies the number of stages.

$$PV = PMT * [(1 - 1/(1+q)^n)/q] \tag{4-1}$$

For estimating the number of states in the HMM, we utilize the Markov test, discussed in Section 2.2. The model is fitted on 80% of the dataset, generating a probabilistic prediction with a coverage ratio of 10%. Subsequently, we validate this prediction using the remaining 20% of historical returns. As a result, we obtain the P-value from the Markov test for three HMMs considering different states as follows:

Number of states	Three	Four	Five
P-value	1.70 %	7.96 %	23.4 %

Table 4.1: Markov Text P-Values

We use a P-value below 0.05 to determine statistical significance, indicating rejection of the null hypothesis. As a result, we cannot conclude that the Hidden Markov Model with three states fits well. However, the four-state model successfully passes the test and emerges as the most parsimonious option. This is the model with fewer parameters, which suggests that it is less susceptible to overfitting, and it passes the test, justifying its selection.

Considering the fitted HMM, Table 4.2 provides the transition matrix among the economy states. Furthermore, Table 4.3 presents the expected value and variance resulting from the fitting of each log-normal distribution for each economic state. This reveals a pattern of two states characterized by negative expected returns, indicative of moments of crisis, and two states exhibiting positive expected returns.

-	State 1	State 2	State 3	State 4
State 1	73.4 %	0 %	0 %	26.6 %
State 2	3.6 %	9 %	82.7 %	4.6 %
State 3	8 %	22 %	70%	0 %
State 4	25.3 %	8.3 %	61 %	5.1 %

Table 4.2: HMM Transition Matrix

-	State 1	State 2	State 3	State 4
Expected Return	-0.0273	-0.030	0.0183	0.0670
Variance	0.0192	0.0202	0.0124	0.0584

Table 4.3: Parameters of Log-Normal Distribution by State

4.2 Case Study Results

The exponential compounding effects of returns maximize the significance of early contributions to wealth accumulation. This phenomenon is elucidated by the model's results, depicted in Figure 4.1. It becomes apparent that the optimal point—i.e., the minimum in the cost function—can be reached with considerably less wealth in the early stages. Consequently, the model provides valuable guidance, encouraging users to commence saving at the earliest opportunity. This advice is particularly pertinent given the model's precise recommendations regarding the contribution amount. While many individuals recognize the importance of saving, determining the appropriate amount often remains unclear.

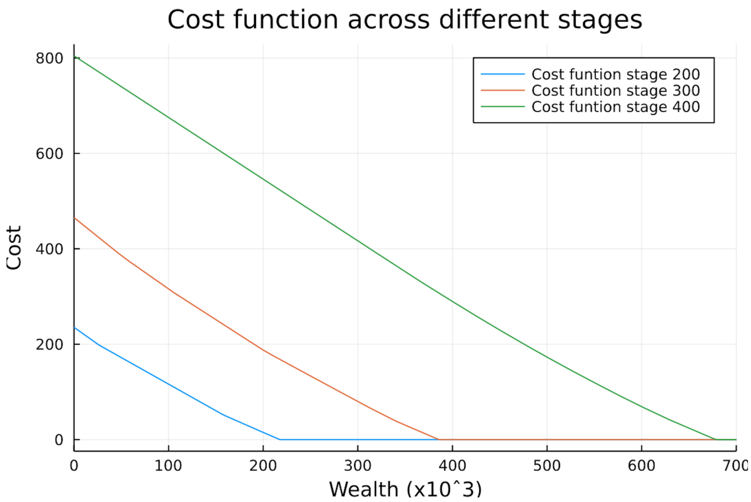

Figure 4.1: Costs Function per Stage

Furthermore, the model provides a clear allocation policy according to the present state that we are in. The heatmaps below illustrate recommended allocations to the $S\&P500$ for each of the four economic states at every stage of the financial plan. Considering that the red line represents the goal, we can see how the portfolio changes as we get close to our target and as we go to different market conditions:

– State 1: According to the HMM transition matrix, we have a high probability of staying in the same state, with negative returns for $S\&P500$. Thanks to that, in this case, the model will always allocate 100% to fixed income.

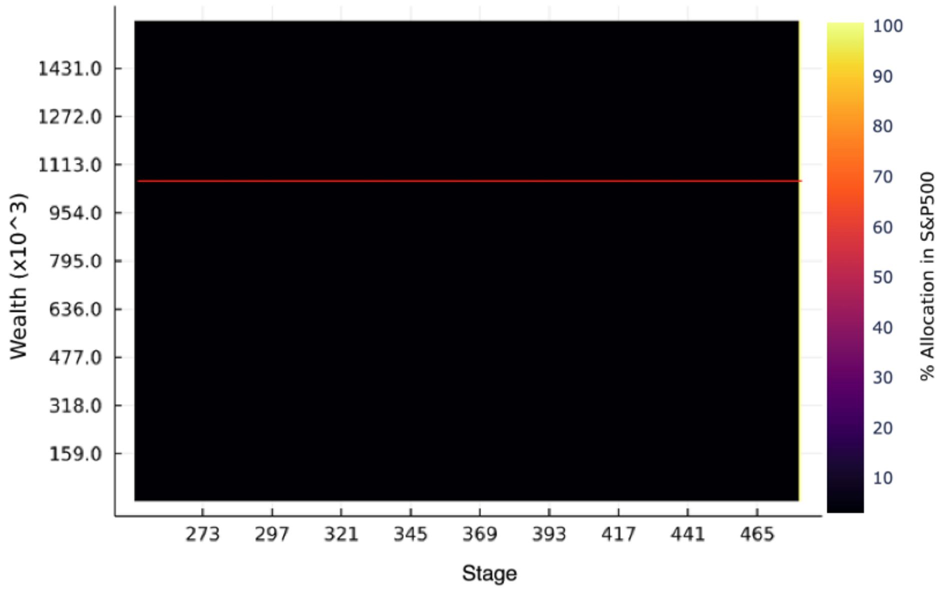

Figure 4.2: Allocation Policy - State = 1

– State 2: In this state, we have a probability of 82% of going to the third state, which has positive returns for the $S\&P500$. Consequently, the model allocates 100% to the $S\&P500$.

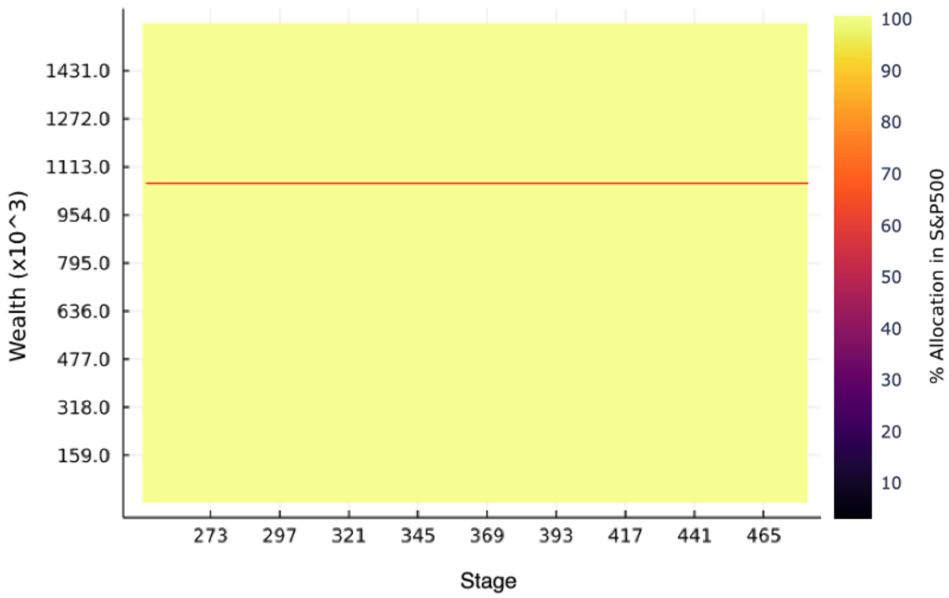

Figure 4.3: Allocation Policy - State = 2

– State 3: Now we have a probability of 30% of going to first or second state, which implies a negative return for stocks, and a probability of 70% of staying in the same state, with positive returns for the $S\&P500$. We tend to allocate less to the fixed income, but this varies according to our level of wealth.

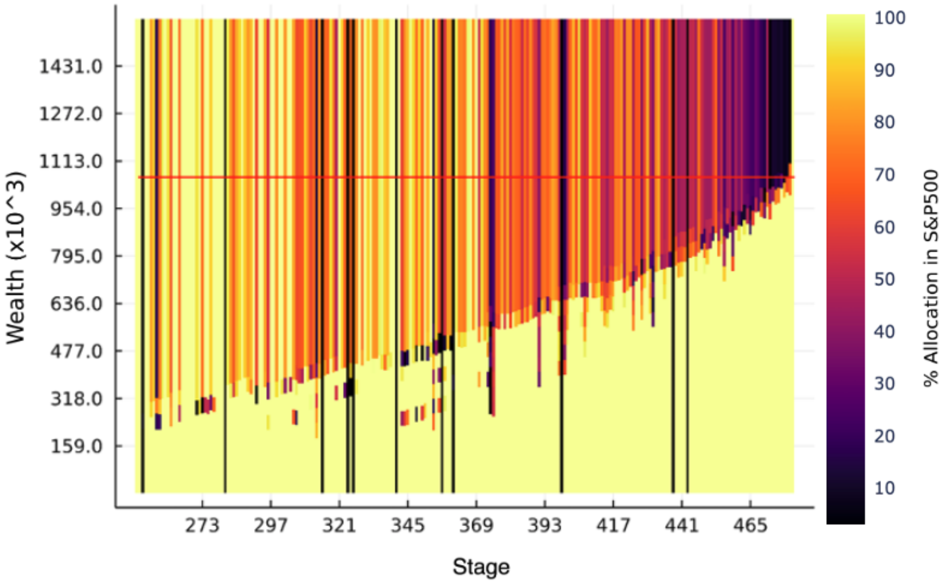

Figure 4.4: Allocation Policy - State = 3

– State 4: This state is similar to the third one although it has a probability of 66% of going to states with positive returns for stocks, a little less than the previous case. We can see that the model tends to allocate a little more to fixed income when compared to the policy in the third state.

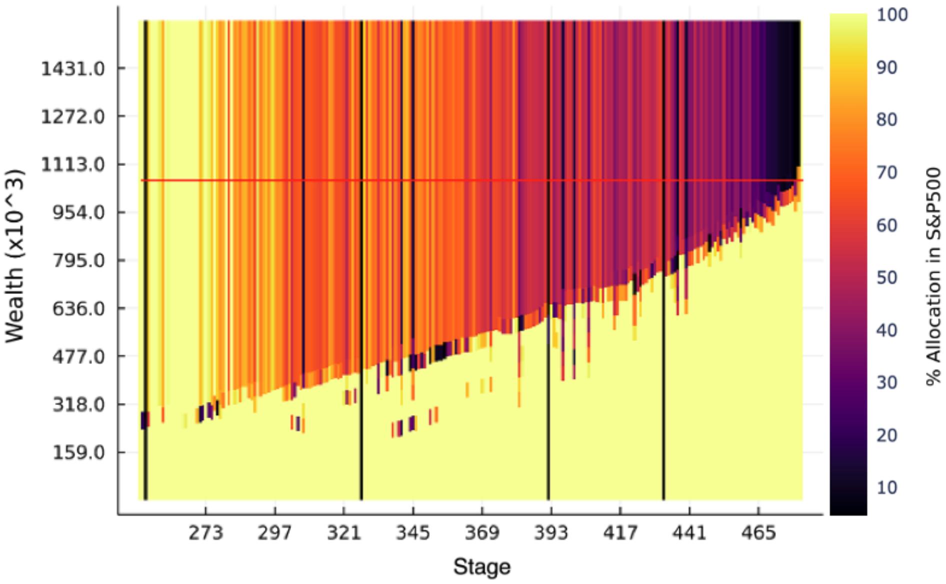

Figure 4.5: Allocation Policy - State = 4

4.3 Sensitivity Analysis

As previously explained, in all intermediary stages we have an upper limit to contributions, and in the final stage we contribute what is needed to reach our goal. Therefore, it is necessary to control our risk tolerance of contributing too much in the last stage, in the case of falling short of the goal or contributing in excess in the intermediary stages, in the case of exceeding the goal in the last stage. In our model, λ_1 and λ_2 are the parameters responsible for such control. A sensitivity analysis to see how these factors interact follows.

An analysis is made by conducting a 10-year simulation, generating 500 market return trajectories based on our fitted Hidden Markov Model. By varying the λ_1 parameter, while keeping λ_2 constant, four different models construct an investment policy that is tested in each scenario.

In Figure 4.6, we depict the first quartile, median, and third quartile of contributions made at each stage across these 500 trajectories. In other words, we expose scenarios where the model had to contribute a small amount, an average amount, and a big amount. While the median of those trajectories exhibits minimal differences, a notable observation emerges in the first and third quartiles.

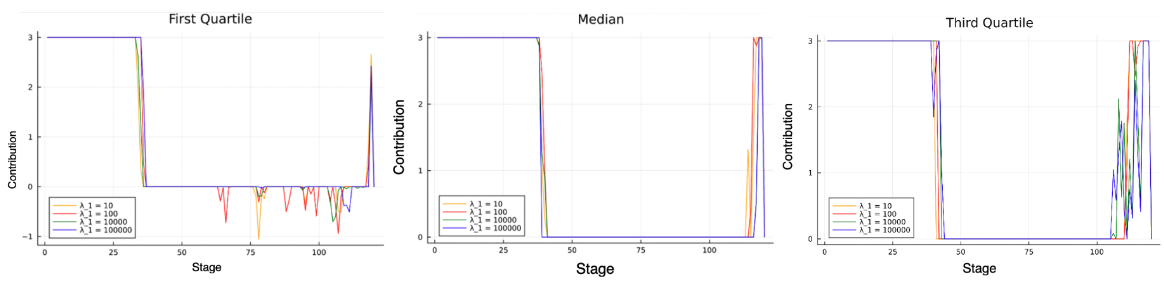

Figure 4.6: Contribution Sensitivity to λ_1

The first quartile corresponds to scenarios with favorable returns and as we increase λ_1, the model withdraws less in intermediate stages. Conversely, in the third quartile, where we encounter scenarios requiring more contributions due to unfavorable outcomes, models with higher λ_1 values initiate contributions faster than the others as they approach the last stage. As expected, in both cases, it becomes evident that with higher λ_1, the model is more cautious, prioritizing a lower probability of falling short of the target in the final stage.

In Figure 4.7, we present the mean, the first quartile, and the third quartile of the wealth in those trajectories from only two models, one with $\lambda_1 = 10$ and the other with $\lambda_1 = 100000$. As we can see, the model with the highest λ_1 has a smaller interval between the quartiles, which means the model has a more defensive approach along the way.

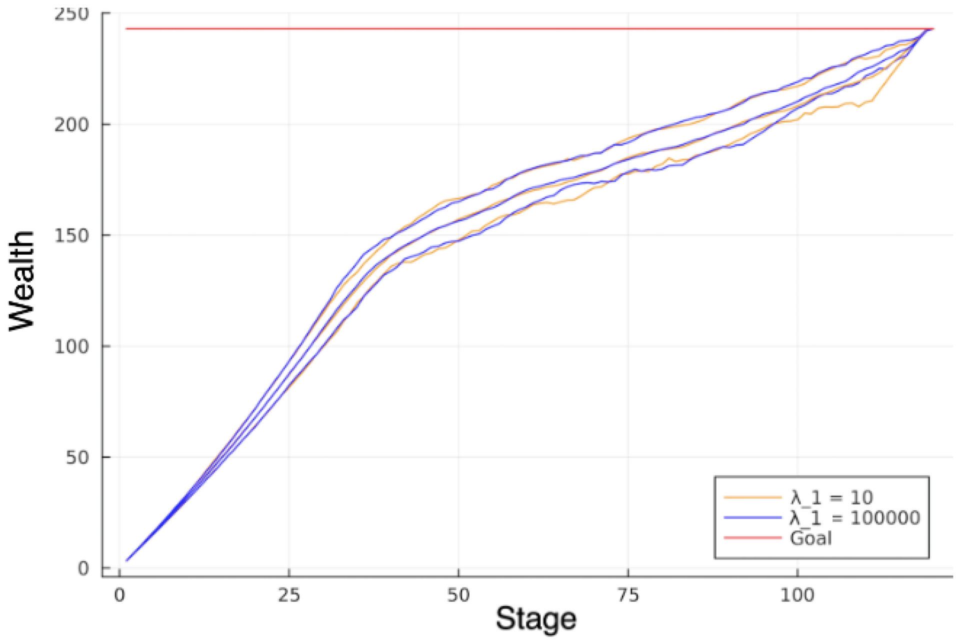

Figure 4.7: Wealth Trajectory Sensitivity to λ_1

Finally, in Figure 4.8, a new analysis is presented concerning a model tailored to mitigate the risk of surpassing our financial goal. In this scenario, the parameters are configured as $\lambda_1 = 1$ and $\lambda_2 = -10/0.85$. The left plot of the figure illustrates the wealth trajectory, while the right plot depicts the contributions made at each stage.

As depicted in the graph on the left side, there are instances in the intermediate stages where our wealth appears to decrease. However, upon closer examination of the contribution trajectory on the right side, it becomes evident that these declines are attributable to planned withdrawals. The model adopts a less conservative approach, which results in a reduced need for contributions along the way.

Nevertheless, this approach may necessitate a higher level of contributions in the final stage, which may exceed the user's available resources. In our specific case, we exceeded the prescribed limit for intermediate-stage contributions at the last moment to successfully achieve our financial goal.

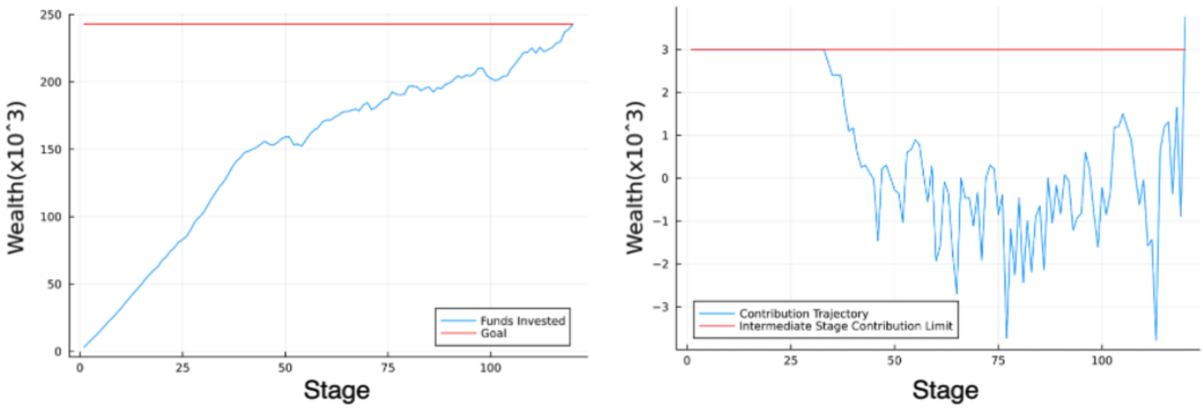

Figure 4.8: Less Conservative Model

4.4 Benchmarking against Traditional Approaches

A crucial inquiry is whether the policy generated by the model outperforms traditional approaches in achieving long-term goals. This section compares our model's allocation policy, termed the HMM Policy, with another common long-term investment strategy, the Fixed Policy, which allocates 60% to stocks and 40% to fixed income.

To conduct this comparison, simulations of return trajectories by the fitted HMM are utilized to evaluate the performance of our model, which optimizes both contributions and allocations, against an alternative model that exclusively optimizes contributions and employs the Fixed Policy for allocation. Both models have parameters set to $\lambda_1 = 10000$ and $\lambda_2 = 1/0.85$, aimed at mitigating the risk of underperformance in the final stage.

In Figure 4.9, several simulated trajectories of the funds for both the HMM Policy and the Fixed Policy models are observed. Furthermore, Figure 4.10 displays the average contributions made in the simulations for each stage.

Notably, the Fixed Policy model surpasses the financial goal much faster than our model at the cost of a higher overall contribution. The model lacks adaptability in allocation, it cannot respond to market conditions. Therefore, it enforces a margin of safety due to its consideration of potential negative stock returns. It becomes evident that the Fixed Policy model requires substantially more contributions: while total contributions across the years average around $\$155,499$, the HMM Policy model averages a significantly lower value of $\$77477$.

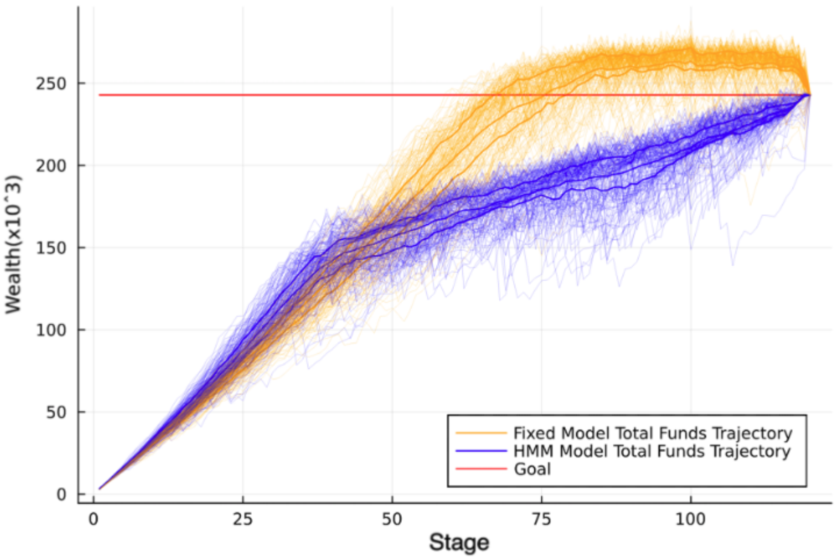

Figure 4.9: Comparative Funds Trajectories

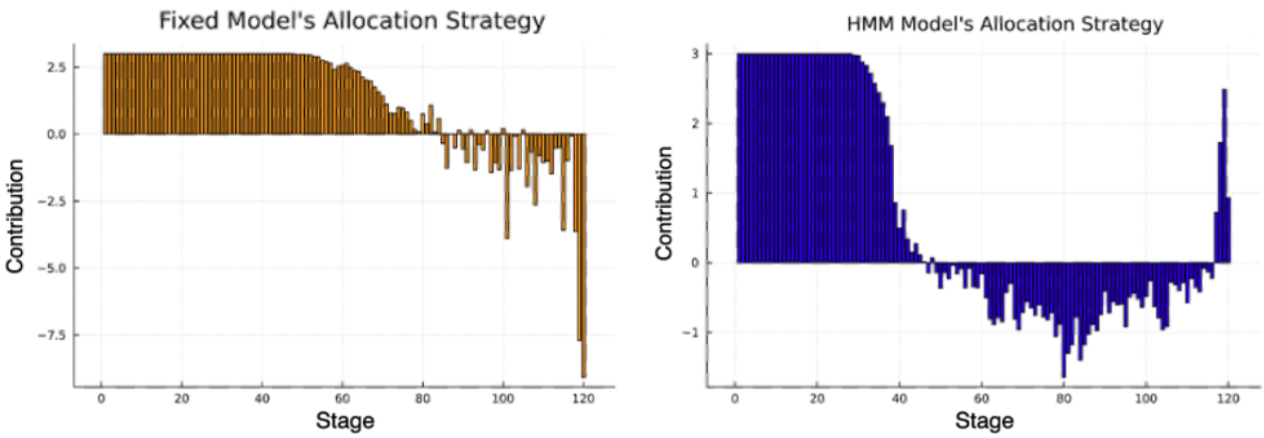

Figure 4.10: Comparative Contributions

5
Conclusion

Our approach offers a practical and effective solution for investors, addressing the limitations of traditional portfolio optimization methods while incorporating the state of the economy for informed investment strategies. However, some areas require further improvement.

The calibration of the Hidden Markov Model presented challenges. Despite our study presenting a straightforward method for fitting it, the HMM has several local minimums, making estimation harder. Exploring more efficient techniques would undoubtedly enhance its performance.

Additionally, the long-term nature of our investment approach is constrained by the limited availability of data points, hampering the model's calibration. Therefore, it is crucial to employ data generation techniques to enhance our out-of-sample tests and improve the reliability of our results.

With ongoing research and advancements, our approach has the potential to become a valuable tool for investors seeking to optimize their portfolios and achieve long-term financial goals.

Bibliography

[1] MARKOWITZ, H.. **Portfolio selection.** The Journal of Finance, 7(1):77–91, 1952.

[2] YAO, H.; LAI, Y.; MA, Q. ; JIAN, M.. **Asset allocation for a dc pension fund with stochastic income and mortality risk: A multi-period mean–variance framework.** Insurance: Mathematics and Economics, 54:84–92, 2014.

[3] YAO, H.; YANG, Z. ; CHEN, P.. **Markowitz's mean–variance defined contribution pension fund management under inflation: A continuous-time model.** Insurance: Mathematics and Economics, 53(3):851–863, 2013.

[4] YALI LIU, MEIYING YANG, J. Z.; BA, M.. **Portfolio selection of the defined contribution pension fund with uncertain return and salary: A multi-period mean-variance model.** Journal of Intelligent Fuzzy Systems, 2018.

[5] LIYUAN WANG, Z. C.; YANG, P.. **Robust equilibrium control-measure policy for a dc pension plan with state dependent risk aversion under mean-variance criterion.** Journal of Industrial and Management Optimization, 2021.

[6] HAO CHANG, JIAAO LI, H. Z.. **Robust optimal stategies of dc pension plans with stochastic volatility and stochastic income under mean-variance criteria.** Journal of Industrial and Management Optimization, 2022.

[7] LI, Y.; WANG, S.; ZENG, Y. ; QIAO, H.. **Equilibrium investment strategy for a dc plan with partial information and mean–variance criterion.** IEEE Systems Journal, 11(3):1492–1504, 2017.

[8] WU, H.; ZHANG, L. ; CHEN, H.. **Nash equilibrium strategies for a defined contribution pension management.** Insurance: Mathematics and Economics, 62:202–214, 2015.

[9] WU, H.; ZENG, Y.. **Equilibrium investment strategy for defined-contribution pension schemes with generalized mean–variance criterion and mortality risk**. Insurance: Mathematics and Economics, 64:396–408, 2015.

[10] VICEIRA, L.. **Life-Cycle Funds**. University of Chicago Press, 2008.

[11] WANG, P.; LI, Z.-F. ; SUN, J.. **Robust portfolio choice for a dc pension plan with inflation risk and mean-reverting risk premium under ambiguity**. Optimization, 70:1–34, 11 2019.

[12] GUAN, G.; LIANG, Z.. **Mean–variance efficiency of dc pension plan under stochastic interest rate and mean-reverting returns**. Insurance: Mathematics and Economics, 61:99–109, 2015.

[13] WOO CHANG KIM, DO-GYUN KWON, Y. L. J. H. K.; LIN, C.. **Personalized goal-based investing via multi-stage stochastic goal programming**. Quantitative Finance, 20(3):515–526, 2020.

[14] PARKER, F. J.. **Goal-based portfolio optimization**. The Journal of Wealth Management, 2016.

[15] CONSIGLI, G.. **Chapter 17 - asset-liability management for individual investors**. p. 751–827, 2008.

[16] GORDON B. FOWLER, V. D. V.. **Holistic asset allocation for private individuals**. The Journal of Wealth Management, 2006.

[17] DEMPSTER, M.; KLOPPERS, D.; MEDOVA, E.; OSMOLOVSKIY, I. ; USTINOV, P.. **Lifecycle goal achievement or portfolio volatility reduction?** The Journal of Portfolio Management, 42:99–117, 01 2016.

[19] HUITRON, M. E. G.; RODRIGUEZ-MONTEMAYOR, E.. **Pension goals and institutional arrangements: Reforms dc 2.0 for latin america**. INSEAD Working Paper No. 2017/54/EPS, August 2017 2007.

[21] DAVI M. VALLADAO, ÁLVARO VEIGA, G. V.. **A multistage linear stochastic programming model for optimal corporate debt management**. European Journal of Operational Research, 237(1):303–311, 2014.

[22] M.V.F. PEREIRA, L. P.. **Multi-stage stochastic optimization applied to energy planning**. Mathematical Programming, 1989.

[24] CONT, R.. **Empirical properties of asset returns: stylized facts and statistical issues**. Quantitative Finance, 1(2):223–236, 2001.