Erasmus University Rotterdam
Erasmus School of Economics
Bachelor Thesis Econometrie en Operationele Research

Optimizing Asset Allocation With Dynamic Programming

Daniël Heemstra (623597)

Supervisor:	E. Kole
Second assessor:	E. Vladimirov
Date final version:	1st July 2024

The views stated in this thesis are those of the author and not necessarily those of the supervisor, second assessor, Erasmus School of Economics or Erasmus University Rotterdam.

Abstract

This research evaluates how dynamic programming methods can help optimize asset allocation among stocks, bonds, and bills, considering varying investors' preferences. Specifically, it investigates the effectiveness of these approaches compared to simpler alternative weight-choosing strategies using multiple performance measures. The study also aims to determine the optimal asset allocation for investors with different levels of risk aversion and investment horizons. By employing real financial data and dynamic programming techniques, the research identifies optimal allocation rules tailored to individual investor profiles. The findings demonstrate that dynamic programming methods yield superior portfolio performance and thus better allocation strategies compared to traditional methods, highlighting their potential for enhancing investment decision-making.

1 Introduction

Asset allocation is a critical component of investment strategy that aims at determining the optimal distribution of assets within a portfolio to balance risk and reward. According to Statman (2004), empirical evidence suggests that the typical investor does not diversify their portfolio sufficiently. Even if investors initially diversify their portfolio, this optimal asset allocation would eventually decay when it is not maintained and market conditions change. There are many more reasons to adjust portfolio weights intermediately instead of using a buy-and-hold strategy. First of all, returns are predictable using observable state variables and regime differences. Additionally, when external factors cause fluctuations in wealth. In all these situations, dynamic programming can be beneficial.

Traditionally, investors have relied on methods such as the Mean-Variance Optimization by Markowitz (1952) to guide them in optimizing their portfolios. However, these approaches often fall short of capturing the complexities of real-world markets, particularly when considering varying levels of risk aversion and investment horizons. Therefore these methods fail at predicting the optimal asset allocation out-of-sample. This was confirmed by DeMiguel et al. (2009), where the equally weighted portfolio yielded superior results compared to the more refined Mean-Variance portfolio and its extensions. Unlike traditional methods, dynamic programming can incorporate a wide range of variables and constraints, offering a more flexible and potentially more accurate framework for investment decision-making.

Dynamic Programming is an approach often used in mathematics and computer science to solve big complex problems by breaking them down into smaller sub-problems. Solving each sub-problem only once and storing the results avoids redundant computations, leading to more efficient solutions for a wide range of problems. Brandt et al. (2005) introduced this as a promising approach for optimizing asset allocation. Value and portfolio iteration were developed to remove the curse of dimensionality of the conventional DSS approach, making it more usable for complicated investment settings. Theoretical research shows great potential for these methods in asset allocation. Despite this potential, there is limited empirical research evaluating the effectiveness of these dynamic programming approaches in real-world asset allocation scenarios. This study aims to fill this gap by assessing how dynamic programming methods can optimize asset allocation among stocks, bonds, and bills, tailored to different investor profiles.

This research considers three potential investment types: stocks, bonds and bills. This is an extension on Van Binsbergen and Brandt (2007) that applies dynamic programming for choosing the weight an investor should put into stocks. The same methods and investment setting are employed, but the latest data will be utilized to connect this study to real-world applications. Furthermore, two additional state variables are used to enhance the simulation of asset returns. Campbell et al. (2003) provides us with the short-term interest rate and yield spread as additional state variables to the dividend-price ratio. The return simulations are a crucial part of the dynamic programming methods because they determine the optimal asset allocation based on the simulations.

Several performance measures are used to determine the empirical effectiveness of the dynamic programming methods, these are then compared to relevant benchmarks. This is important to ensure that the produced asset allocations are improved. The Sharpe and Sortino ratios are evaluated to ensure an optimal risk-reward ratio. Also, the turnover will be measured to give some insight into potential trading costs associated with the strategies. Lastly realized portfolio returns and return volatility are provided to identify differences between strategies in more detail. The resulting asset allocations are then discussed for the optimal dynamic programming approaches, given the different investor preferences.

The dynamic programming methods resulted in slightly superior results, which confirms their effectiveness and potential in investment decision-making. Especially when considering long investment horizons the dynamic programming approaches seem useful. This led to certain optimal asset allocations that are dependent on the investor's risk aversion and investment horizon. The findings have the potential to enhance investment decision-making processes and contribute to the broader field of financial optimization.

The paper is further structured as follows. In section 2, the data used in this research is provided. Following this, all methods, portfolios and performance measures are extensively explained in section 3. Next, the results are presented and discussed in section 4. Lastly, the main findings are summarized and the paper is concluded in section 5.

2 Data

This research considers three different assets: stocks, bonds and bills. The stock returns are based on the S&P 500 index returns because they give a good indication of the overall stock market returns. Next, the bond returns are based on the 10-year US bond returns. 10-year maturity since it fits the sample size the best. Lastly, the bill returns are established from the 90-day US treasury bill returns, these serve as the risk-free rate. Excess stock and bond returns are created by subtracting the risk-free rate.

To make predictions of future asset returns, state variables are needed to measure the direction of where the markets are going. The first state variable is the dividend price ratio, this is an often-used predictor for stock returns. It is created by first taking the difference of the last 12 monthly value-weighted returns on the S&P 500 index including and excluding dividends, which are then multiplied by their respective price indexes to make the sum of dividends over the last year. Finally, this sum is divided by the current price index. This tells us how the stocks are valued compared to their dividends. A high ratio, for example, would indicate that the stocks are undervalued which would lead to an increase in stock returns for the next periods. More on the dividend-price ratio can be found in Campbell and Shiller (1988). The next state variable considered is the short-term interest rate, for which the risk-free rate is used. The interest rate is a powerful predictor of stock returns and bond yields, which is confirmed in a recent paper by Rapach et al. (2016). When interest rates rise, it can make borrowing money for a company more expensive, which means they have less money to invest back in the company and less cash flow stability, leading to a decrease in stock returns. This inverse relation is also present for bonds. When interest rates rise, investors will no longer prefer the lower fixed interest rate paid by a bond, which eventually decreases bond prices. The third state variable employed is the yield spread, this depicts the difference between longer uncertain bond returns and shorter more certain bill returns. This research defines this the same as the excess bond returns. This state variable tells a lot about the confidence of investors in the economy. When investors are sceptical about the future economy they will put a greater portion of their wealth into the more certain short-term assets, this would lead to more demand for bills and thus a lower bill rate. Which further results in a higher yield spread. This variable has predictive power for all assets. Lastly, the lagged excess stock returns will be used as a state variable because stock returns are generally hard to predict.

All the data is obtained via Wharton Research Data Services (WRDS), which is a popular provider of financial and economic data for research applications. This connection is used to download the most commonly used data for stock and bond characteristics from the Center of Research in Security Prices (CRSP).

Quarterly data is considered because the paper's main goal is to study the optimal asset ratio between stocks and bonds given the investors' preferences and current state variables. Inside each quarter of a year, investors can determine what stocks or bonds specifically are best to hold to further increase returns, while taking into account the ratio between them. Furthermore, only the last 50 years or 200 quarters of data are considered. This is split into 150 quarters in-sample to estimate the VAR and 50 quarters out-of-sample to determine weights and test performance. Here a moving estimation window is imposed to keep the estimation of the VAR based on the last 150 quarters, but also dependent on the most recent information. Some characteristics of the used data are displayed in table 1.

Table 1: Data characteristics

Sample	Type	Mean	Standard dev.	Skewness	Kurtosis	Min	Max
Entire sample	Log excess stock returns	0.008	0.086	-0.967	4.664	-0.334	0.182
	Log excess bond returns	0.004	0.042	0.355	3.123	-0.104	0.117
	Gross bill returns	1.012	0.010	0.809	3.621	1.000	1.042
	Log dividend-price ratio	0.028	0.012	0.726	2.335	0.011	0.059
Out-of-sample	Log excess stock returns	0.023	0.079	-1.232	4.897	-0.231	0.182
	Log excess bond returns	0.002	0.038	0.242	3.528	-0.071	0.106
	Gross bill returns	1.003	0.004	1.639	4.797	1.000	1.014
	Log dividend-price ratio	0.019	0.003	-0.713	2.673	0.013	0.023

From the table, the risk type of the assets is easily noticed. The stock returns have the highest standard deviation, after this the bond returns and last the bill returns. But this is compensated by higher mean returns for the more risky assets. The stocks are negatively skewed, which means more stock returns are below the mean but this is compensated by some higher stock returns. The reverse is true for the bond and bill returns. Furthermore, all assets have relatively high kurtosis, which means that all asset returns are more likely to be near the mean compared to a normal distribution. The risk types also coincide with the return ranges, because riskier assets have a wider return range. The dividend-price ratio has a lower kurtosis and further a big difference between both samples. Its skewness changes from positive to negative between the samples, also the mean, standard deviation and range shrink when considering only the last 50 quarters. The assets also see a big difference between both samples. The stock returns are on average higher in the last 50 quarters, while the bond and bill returns are lower.

Table 2 shows the first-order autocorrelations between the variables.

Table 2: First order autocorrelations entire sample

	Log excess stock returns	Log excess bond returns	Gross bill returns	Log dividend-price ratio
Log excess stock returns	0.087	-0.087	-0.160	-0.094
Log excess bond returns	0.159	0.020	-0.152	-0.043
Gross bill returns	-0.119	0.014	0.936	0.167
Log dividend-price ratio	0.074	-0.007	0.158	0.972

Almost all variables have some first-order autocorrelation between them, this makes them suitable for a vector autoregressive model.

3 Methodology

In this section, first, the data-generating process is described. Next, the different employed dynamic programming methods, which are largely in line with Brandt et al. (2005), are extensively explained. Furthermore, the benchmark portfolios and performance measures are specified.

3.1 Data-generating process

The dynamic programming methods rely on simulated data, therefore an accurate data-generating process is necessary. This paper employs the following vector autoregressive (VAR) model to simulate asset returns.

$$\mathbf{y}_t = \mathbf{c} + \mathbf{\Phi}\mathbf{y}_{t-1} + \epsilon_t \tag{1}$$

where

$$\epsilon_t \sim N(\mathbf{0}, \mathbf{\Sigma}) \tag{2}$$

$\mathbf{y}_t := (R_t^s, R_t^b, r_t^f, Z_t)'$ is a 4×1 vector of endogenous variables at time t. Furthermore, $\mathbf{c}$ is a 4×1 vector of intercept terms and $\mathbf{\Phi}$ is a 4×4 matrix of coefficients. Lastly, ϵ_t is a 4×1 vector of error terms at time t.

Here R_t^s represents the quarterly log excess stock returns based on the S&P 500 index returns and the 90-day US treasury bill returns. Next R_t^b denotes the quarterly log excess bond returns, these are based on the 10-year US bond returns and again the 90-day US treasury bill returns. Third, r_t^f depicts the quarterly gross risk-free rate, which is established from the 90-day US treasury bill returns. Lastly Z_t is defined as the log dividend price ratio.

The results section 4.1 contains the estimates for the entire sample to grasp how the model will look approximately, and also to justify the validity of the simulation model.

3.2 Dynamic programming and problem setting

This paper considers a multi-period investing problem, given by the following general equations:

$$\max_{\{x_s\}_{s=t}^{T-1}} \mathbb{E}_t\left[u(W_T)\right] \tag{3}$$

$$\text{subject to } W_{s+1} = W_s(x_s' r_{s+1} + r^f) \text{ for } s = 1, \ldots, T-1 \tag{4}$$

for a given utility function u, initial wealth W_t and risk-free rate r^f. There are T portfolio decisions, that result in optimal weights $\{x_s\}_{s=t}^{T-1}$. n assets, where r_{s+1} vector of excess returns consists of the excess stock and bond returns in this case. Also a conditional expectation because of m state variables z_t, $\mathbb{E}_t[\cdot] = \mathbb{E}[\cdot \mid z_t]$.

Furthermore, we define the value function:

$$V(W_t, z_t) = \max_{\{x_s\}_{s=t}^{T-1}} \mathbb{E}[u(W_T) \mid z_t] \text{ s.t. equation 4} \tag{5}$$

This function depicts the value of the optimal solution, which depends on wealth W_t and state variables z_t at time t. Also, the intermediate solution to the problem can be found by using the next equation:

$$V(W_\tau, z_\tau) = \max_{\{x_s\}_{s=\tau}^{T-1}} \mathbb{E}[u(W_T) \mid z_\tau] \text{ for } t < \tau < T \tag{6}$$

The value function makes it possible to split the problem.

$$\begin{aligned} V\left(W_t, z_t\right) &= \max_{\{\mathbf{x}_s\}_{s=t}^{T-1}} \mathbb{E}_t\left[u\left(W_T\right)\right] \\ &= \max_{\mathbf{x}_t} \mathrm{E}_t\left[\max_{\{\mathbf{x}_s\}_{s=t+1}^{T-1}} \mathbb{E}_{t+1}\left[u\left(W_T\right)\right]\right] \\ &= \max_{\mathbf{x}_t} \mathrm{E}_t\left[V\left(W_{t+1}, \mathbf{z}_{t+1}\right)\right] \\ &= \max_{\mathbf{x}_t} \mathrm{E}_t\left[V\left(W_t\left(x_t' r_{t+1} + r^{\mathrm{f}}\right), z_{t+1}\right)\right] \end{aligned} \tag{7}$$

Where the law of iterated expectations is used. The last equality is called the Bellman equation, which allows us to split the problem into a set of recursive optimization problems.

Next, we need to define the power utility:

$$u\left(W_t\right) = \frac{W_t^{1-\gamma}}{1-\gamma} \tag{8}$$

Here γ represents the coefficient of relative risk aversion, which is the solution of $-u''(W_t)\,W_t/u'(W_t)$. The power utility function is homogeneous with degree $1-\gamma : u(cW_t) = c^{1-\gamma}u(W_t)$. Combining this with equation 4 yields:

$$u(W_{t+1}) = \frac{W_t^{1-\gamma}\left(x_t' r_{t+1} + r^{\mathrm{f}}\right)^{1-\gamma}}{1-\gamma} \tag{9}$$

From this, we can conclude that the optimal portfolio x_t^* in a single period optimization satisfies $\mathbb{E}_t\left[\left(x_t^{*\prime} r_{t+1} + r^{\mathrm{f}}\right)^{-\gamma} r_{t+1}\right] = \mathbf{0}$, independent of wealth since $W_t > 0$.

This allows us to rewrite the problem to the following form:

$$\begin{aligned}
&\max_{\{x_s\}_{==1}^{T-1}} \mathbb{E}_t\left[\frac{W_T^{1-\gamma}}{1-\gamma}\right] \\
&= \max_{\{x_s\}_{s=t}^{T-1}} \mathbb{E}_t\left[\frac{W_t^{1-\gamma}}{1-\gamma}\prod_{s=t}^{T-1}\left(x_s' r_{s+1} + r^f\right)^{1-\gamma}\right] \\
&= \max_{x_t} \mathbb{E}_t[\underbrace{\frac{W_t^{1-\gamma}\left(x_t' r_{t+1} + r^f\right)^{1-\gamma}}{1-\gamma}}_{\nu(W_{t+1})} \underbrace{\max_{\{x'\}_{==t+1}^{T-1}} \mathbb{E}_{t+1}\left[\prod_{s=t+1}^{T-1}\left(x_s' r_{s+1} + r^f\right)^{1-\gamma}\right]}_{\psi(z_{t+1})}]
\end{aligned} \tag{10}$$

We can set $W_t = 1$ because the optimization is independent of wealth W_t. Now we can recover the earlier denominated Bellman equation as:

$$\frac{1}{1-\gamma}\psi(z_t) = \max_{x_t} \mathbb{E}_t\left[u\left(x_t' r_{t+1} + r^f\right)\psi(z_{t+1})\right] \tag{11}$$

The term $1/(1-\gamma)$ is needed to get the first term in the product in equation 10 in the form $u\left(x_s' r_{s+1} + r^{\mathrm{f}}\right)$. $\psi(z_t)$ now can be interpreted as a scaled value function. This makes it possible for the optimization problem to simplify to a sequence of single-period optimizations that include the scaled value function, instead of a much more complex multi-period optimization.

DSS

Discretizing the State Space is seen as a benchmark in Van Binsbergen and Brandt (2007) and is the most popular dynamic programming approach. The entire procedure can be found in appendix A, this can be useful for readers who are unfamiliar with the above-described mechanism to understand the the next two methods better. It is however not used in this paper because of an important drawback of the DSS approach, which is the curse of dimensionality. This is caused by the fact that the number of grid points grows exponentially in the number of state variables k. If you consider n points per state variable, the number of combinations is n^k. This makes the optimization very computationally heavy when wanting to consider many grid points and multiple assets. Furthermore, interpolation becomes difficult when the number of state variables grows. In this research, the VAR model is too large, making it unreasonable and inefficient to use DSS.

Value and Portfolio Iteration

Brandt et al. (2005) proposed an alternative to DSS: value and portfolio iteration. The general idea is to generate a large set of paths for the returns and state variables, with the length of each path equal to the investment horizon T. Next, calculate for a given weight the utility at point s of path p, and use these to perform a regression on the utility of the different paths on its state variables and a constant. The fitted values of this regression give the expected utility of each path. Finally, use these fitted values to determine the optimal weight.

The main difference between both methods lies in the need for the value of future decisions at time $s+1$. It is possible to use the scaled value function or the optimal portfolio weights, that belong to path p, which resulted from the optimization at time $s+1$. Respectively these are called value iteration and portfolio iteration. When the number of paths goes to infinity both approaches are asymptotically the same. But simulating an infinite amount of paths is impossible, thus there remains a small difference between them.

An example can further explain this. For the decision at time $T-1$, calculate for a given weight x_{T-1} the expected utility as the fitted value of a regression of the utility $u\left(x_{T-1} r_{T,p} + r_f\right)$ on a constant and the state variables $z_{T-1,p}$. Select the optimal weight for each path, denote this as $x^*_{T-1,p}$, and use the expected utility to construct the scaled value function $\psi\left(z_{t-1,p}\right)$. Next for the decision at $T-2$, include the value of the decision at $T-1$. Value iteration includes the expected value of the decision at $T-1$, this leads to the following equation:

$$\max_{x_{T-2}} \mathbb{E}_{T-2}\left[\left(x_{T-2} r_{T-1} + r_f\right)^{1-\gamma} \hat{\mathbb{E}}_{T-1}\left[\frac{\left(x^*_{T-1} r_T + r_f\right)^{1-\gamma}}{1-\gamma}\right]\right] \tag{12}$$

But portfolio iteration includes the actual value for the path p of the decision at $T-1$, and the expected value comes from the equation:

$$\max_{x_{T-2}} \mathbb{E}_{T-2}\left[\frac{\left(x_{T-2} r_{T-1} + r_f\right)^{1-\gamma}\left(x^*_{T-1} r_T + r_f\right)^{1-\gamma}}{1-\gamma}\right] \tag{13}$$

There is thus a small difference between both methods when using a finite number of simulations because the moments are approximated.

Repeat these steps till time t, where the last iteration will produce the weights for the amount invested in stocks and bonds for the upcoming quarter. Then move forward one period inside the testing phases and repeat the same procedure. This would eventually generate asset allocations for the 50 out-of-sample investing moments, for both dynamic programming approaches.

3.3 Benchmark portfolios

Evaluating the relative performance of a portfolio calls for some benchmark portfolios. This research considers the following two strategies as benchmarks.

Myopic portfolio

This strategy only considers an investment horizon of one quarter. Therefore it does not use dynamic programming because it only has to solve the optimization problem for the upcoming quarter. Optimization is done using equation 9 with the same power utility function and weight restrictions. This makes it an excellent benchmark to test the dynamic programming part of the value and portfolio iteration.

Buy-and-hold portfolio

This strategy again employs the same power utility function and weight restrictions, but this time the utility will be calculated for the whole path at once. The weight that maximizes the total utility, which is the sum of all paths, will be used and held for the entire upcoming investment horizon. This portfolio will be calculated for each investor preference case. Again this strategy does not use dynamic programming but it does employ the same time horizons, which makes it suitable for determining the effectiveness of the dynamic programming part.

3.4 Performance measures

This paper evaluates different dynamic programming methods, which are then used for asset allocation strategies. Three performance measures will be used to assess their effectiveness, these are explained and motivated in this subsection.

Out-of-sample Sharpe ratio

One of the most used performance measures for portfolios is the out-of-sample Sharpe ratio, denoted as $S\hat{H}R_k$. This is defined as the ratio between the estimated mean excess return $\hat{\mu}_k$ and the estimated standard deviation of the returns $\hat{\sigma}_k$ of strategy k. This can be formulated as follows:

$$S\hat{H}R_k = \frac{\hat{\mu}_k}{\hat{\sigma}_k} \tag{14}$$

The Sharpe ratio serves as a gauge for the risk-return trade-off associated with a given allocation strategy, where a higher ratio denotes a better risk-adjusted return. To determine the statistical significance of the differences in Sharpe ratios across strategies, the Jobson-Korkie test (Jobson and Korkie (1981)), which assesses the equality of Sharpe ratios, is employed. The computation of this test's p-value facilitates the determination of whether the Sharpe ratios of two distinct strategies are statistically different. When it is stated that two Sharpe ratios significantly differ, it was tested using a Jobson-Korkie test with a significance level of 0.1 (*), 0.05 (**) and 0.01 (***).

Out-of-sample Sortino ratio

A downside of the Sharpe ratio is that it considers the total standard deviation of the returns. Thus, deviation caused by high positive returns is seen as bad risk, even though this is not necessarily true. The Sortino ratio only considers the negative returns standard deviations, this gives a better view of a portfolio's risk-adjusted performance since positive volatility is a benefit. In the equation below, $\hat{\sigma}_{d,k}$ represents the standard deviation of the negative portfolio returns.

$$S\hat{O}R_k = \frac{\hat{\mu}_k}{\hat{\sigma}_{d,k}} \tag{15}$$

Turnover

The turnover quantifies the frequency with which assets within a portfolio are bought and sold. High turnover suggests frequent trading activity, while low turnover implies a buy-and-hold strategy with less frequent trading. High turnover can lead to increased transaction costs, which could mean that a strategy becomes less desirable even when its shape ratio seems great. The turnover for the portfolio k is expressed in the following function.

$$\text{Turnover}_k = \frac{1}{T-1} \sum_{t=1}^{T-1} \sum_{j=1}^{N} \left(\left| \hat{\mathrm{w}}_{k,j,t+1} - \hat{\mathrm{w}}_{k,j,t^+} \right| \right) \tag{16}$$

$\hat{\mathrm{w}}_{k,j,t+1}$ is the portfolio weight before rebalancing at t + 1; and $\hat{\mathrm{w}}_{k,j,t^+}$ is the desired portfolio weight at time t + 1, after rebalancing. T depicts the investment period. Although this study does not account for transaction costs, turnover metrics are included to provide insights into the trading activity associated with each portfolio strategy.

4 Results

Everything was coded in RStudio, R version 4.4.0 by R Core Team (2023) and used third-party packages can be found in the footnote [1]. All code has been run on an Intel(R) Core(TM) i7-10870H CPU @ 2.20GHz 2.21 GHz laptop with 16,0 GB of RAM.

The dynamic programming weights were created using 1000 simulations to simulate the returns and state variables. Also, the whole procedure was repeated five times and the resulting portfolio weights are their average. There are two scenarios for the risk aversion coefficients considered; 5 and 20. Five represents a moderate risk-averse investor and twenty represents an extreme risk-averse investor. Also, four investment horizons are studied; 4, 8, 20 and 40 quarters. In total, this makes 8 distinct investor types, covering a wide range of investor preferences. The optimal portfolio weights are found by a grid search over the interval $[0, 1]$ in steps of 0.05.

4.1 VAR results entire sample

The resulting VAR model for the entire sample is displayed below. The p-values of their respective t-statistics are in parentheses after each coefficient.

$$\begin{pmatrix} R_t^s \\ R_t^b \\ r_t^f \\ Z_t \end{pmatrix} = \begin{pmatrix} 2.037(0.035) \\ 0.032(0.947) \\ 0.136(0.000) \\ -0.022(0.421) \end{pmatrix} + \begin{pmatrix} 0.058(0.412) & 0.404(0.005) & -2.045(0.035) & 1.378(0.072) \\ -0.046(0.197) & 0.030(0.673) & -0.027(0.956) & -0.017(0.963) \\ -0.000(0.924) & -0.043(0.000) & 0.863(0.000) & 0.074(0.006) \\ -0.003(0.206) & -0.009(0.029) & 0.023(0.409) & 0.964(0.000) \end{pmatrix} \begin{pmatrix} R_{t-1}^s \\ R_{t-1}^b \\ r_{t-1}^f \\ Z_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \\ \epsilon_{3,t} \\ \epsilon_{4,t} \end{pmatrix}$$

[1] The following third-party R packages were used: Pfaff (2008), Venables and Ripley (2002), Wickham (2016), Kassambara (2023), Ardia and Boudt (2015), Komsta and Novomestky (2022), Van Domelen (2018)

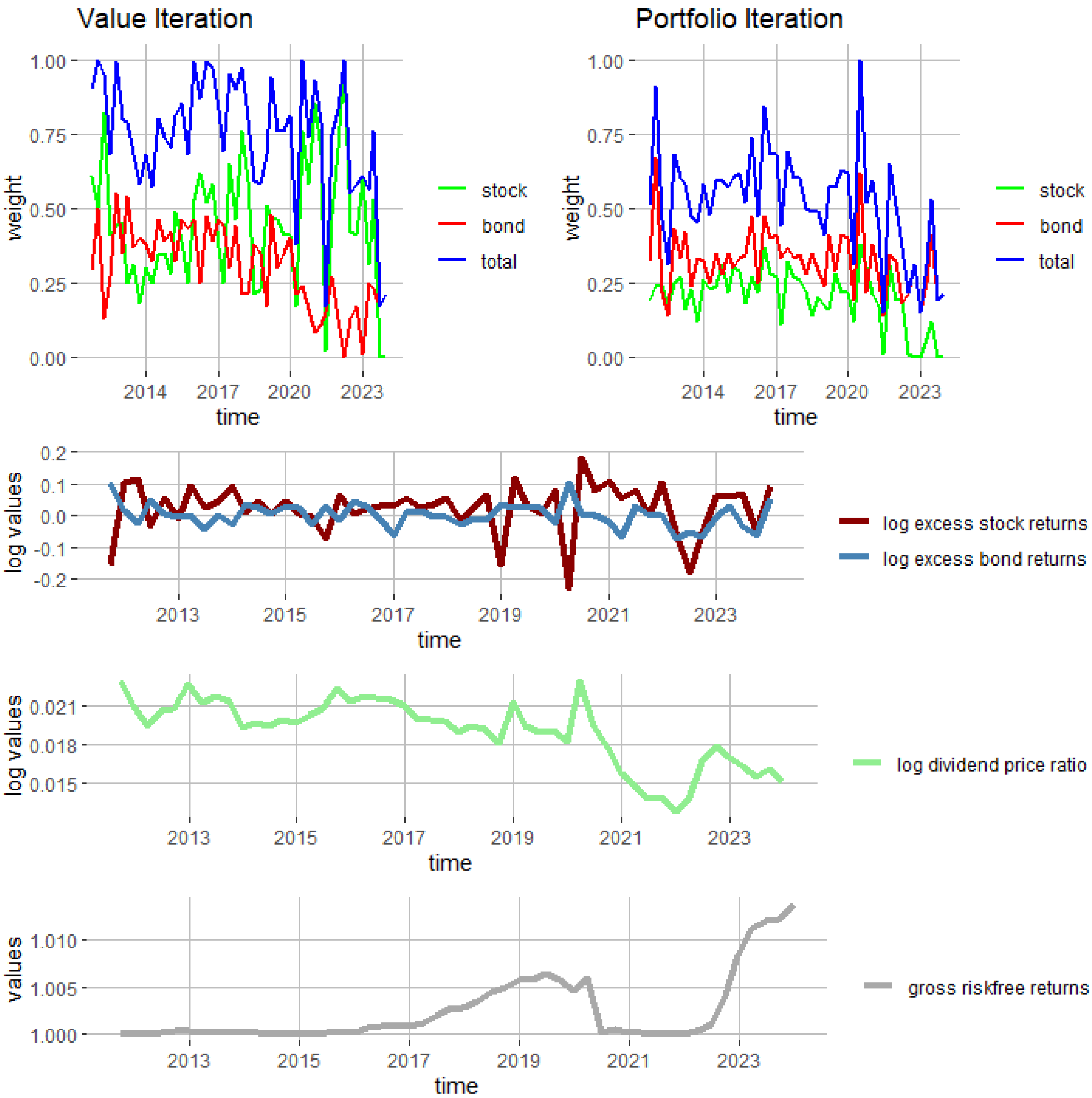
$$\begin{pmatrix}\epsilon_{1,t}\\ \epsilon_{2,t}\\ \epsilon_{3,t}\\ \epsilon_{4,t}\end{pmatrix} \sim N\left(\begin{pmatrix}-3.589e-18\\ -6.049e-19\\ -2.805e-19\\ 6.810e-21\end{pmatrix}, \begin{pmatrix}7.065e-03 & -9.329e-05 & -2.042e-05 & -1.814e-04\\ -9.329e-05 & 1.770e-03 & 1.188e-05 & -1.384e-05\\ -2.042e-05 & 1.188e-05 & 8.588e-06 & 8.058e-07\\ -1.814e-04 & -1.384e-05 & 8.058e-07 & 5.949e-06\end{pmatrix}\right)$$

The stock returns, bill returns and dividend-price ratio seem to be properly described by multiple significant variables. Only the bond returns lack a significant variable but their error term does not seem out of place compared to the others. The risk-free rate and the dividend price ratio are mostly explained by themself, compared to the stock and bond returns that depend more on other variables. Also, the earlier described economic relations between the variables are present. Therefore it can be concluded that the VAR model is suitable for this research purpose and will create accurate simulations.

4.2 Analyzing method dynamics

The upper two graphs in figure 1 show the results of the weights from the dynamic programming methods, with a risk aversion coefficient of 20 and a time horizon of 8 quarters or 2 years. This case is highlighted because it is the most sensitive to changes in the state variables. The other cases can be found under appendix B. The remaining graphs in figure 1 display the behaviour of the state variables during the out-of-sample testing phase, these are not results but it is easier this way to see the method dynamics.

Figure 1: Results for Time Horizon 8 and Risk Aversion 20



There appears to be a difference between the weights of both dynamic programming methods. The value iteration method allocates more wealth to stocks and bonds than the portfolio iteration. This was already expected because in Van Binsbergen and Brandt (2007) it is shown that for the same investor preferences, the weights in riskier assets were also substantially lower for the portfolio iteration. This is further demonstrated by the bigger part of the wealth in stocks than in bonds for the value iteration, while the portfolio iteration prefers to put more wealth in bonds.

Another difference between the two models can be found around the first quarter of 2020. There is a large decrease in the log excess stock returns, while the log excess bond returns increase slightly. The dividend price ratio also peaks at the same time. Both methods react to this by increasing the next period's total weight, but the value iteration does this by investing more in stocks, while the portfolio iteration increases the bond's weight. After this, the dividend price ratios decrease fast for the next couple of quarters. Value iteration keeps the weight invested in stocks high but portfolio iteration does the opposite. This again already came up in the paper Van Binsbergen and Brandt (2007), where they provided graphs of the weight invested in the risky asset relative to the state variable. All graphs showed higher weights in the risky asset when the state variable was in a poor state for the value iteration. Furthermore, the gross risk-free returns increase substantially towards the end of the testing phase. This would make it more plausible to invest more in this asset. Both methods show that they adapt to this by decreasing the total weight of stocks and bonds to put a bigger part of the wealth in the bills.

4.3 Performance measures

Table 3 contains all the resulting portfolio performance measures. There is a distinction between the investors' preferences because these can not be compared to each other. Only the Myopic portfolio can be compared to the other strategies within the same risk aversion portfolios. The realized return is also shown which is defined as the compounded excess portfolio returns. Also, all portfolio return volatilities are revealed in the last column of the table to address differences in Sharpe ratios better.

Table 3: Performance measures

Risk Aversion	Time Horizon	Strategy	Sharpe Ratio	Sortino Ratio	Turnover	Realized Return	Volatility
$\gamma = 5$	1	Myopic	0.400	0.651	1.581	0.926	0.034
	4	Value	0.375	0.659	1.566	0.949	0.038
		Portfolio	0.372	0.654	1.569	0.938	0.038
		Buy-hold	0.420	0.460	0.406	1.045	0.036
	8	Value	0.415	0.725	1.582	1.186	0.040
		Portfolio	0.419	0.753	1.574	1.192	0.039
		Buy-hold	0.396	0.460	0.205	0.985	0.037
	20	Value	0.403	0.662	1.647	1.481	0.048
		Portfolio	0.417	0.679	1.621	1.424	0.045
		Buy-hold	0.407	0.475	0.069	1.161	0.040
	40	Value	0.347	0.422	1.896	2.010	0.072
		Portfolio	0.402	0.639	1.643	1.462	0.048
		Buy-hold	0.360	0.380	0.035	0.915	0.038
$\gamma = 20$	1	Myopic	0.400	0.944	0.540	0.328	0.014
	4	Value	0.477	1.486	0.662	0.479	0.017
		Portfolio	0.444	1.138	0.625	0.401	0.016
		Buy-hold	0.374	0.439	0.142	0.801	0.033
	8	Value	0.395	0.624	1.204	1.111	0.040
		Portfolio	0.464	1.172	0.740	0.463	0.017
		Buy-hold	0.380	0.463	0.080	0.836	0.034
	20	Value	0.372	0.427	1.775	1.423	0.052
		Portfolio	0.453	0.956	0.849	0.486	0.018
		Buy-hold	0.412	0.505	0.045	1.009	0.036
	40	Value	0.177	0.256	1.646	0.434	0.047
		Portfolio	0.439	0.993	1.015	0.725	0.026
		Buy-hold	0.414	0.482	0.030	1.028	0.036

The tables make clear that for low values of the investment horizon T and risk aversion γ the value and portfolio iteration methods are almost equivalent. When these increase both methods begin to differ more, in these situations, the portfolio iteration seems to perform better.

First, the lower risk aversion scenarios are analyzed. The Sharpe ratios only slightly differ, but this difference is for no combination significant according to the Jobson-Korkie test. The Sortino ratios tell more, the buy-and-hold portfolios all have lower values than their respective dynamic programming portfolios. Thus the buy-and-hold portfolio returns have more volatile negative returns which is not desirable. This shows that the dynamic programming segment is effective in this setting, which is elongated by the higher realized returns for the value and portfolio iteration. Their return volatility is consistently higher but this is caused by more volatility from positive returns which can not be seen as bad risk. The turnovers of the dynamic programming are much higher than the buy-and-hold strategy because they are rebalanced more often. However, this is not seen as a problem, because the allocations are only updated quarterly and they are not extremely high. The improved portfolio performance significantly compensates for the increased trading activity. The myopic benchmark strategy does seem to produce similar results for most of the lower risk-averse scenarios, only its realized return is lower than the dynamic programming portfolios. This makes the dynamic programming methods slightly more favourable for this setting, where between both for higher investment horizons the portfolio iteration performs a little better.

Next, the extreme risk-averse investor scenarios. The Sharpe ratios seem to differ more, however they still do not differ significantly in any case. This is due to the low out-of-sample size of 50 quarters, which makes the Jobson-Korkie test not reject the null hypothesis of equal Sharpe ratios. Just as in the other risk aversion case, the portfolio iteration performs better when the investment horizon increases. The Sortino ratios indicate that the value iteration has more volatile negative returns, which is not preferable. Also, their turnover is higher in all cases. There is no clear winner when comparing their realized returns, but there is when looking at the portfolio volatilities. The portfolio iteration volatility is much lower for all scenarios. Only the myopic benchmark portfolio possesses lower volatility, but their other performance metrics do not justify choosing this portfolio over the others.

4.4 Mean weights

Table 4 contains some more insights into the weights that the dynamic programming methods produce for each asset given the investor's preference. Now these weights can be combined with the previous performance measures to determine the optimal asset allocation for an investor with a certain investment horizon and risk aversion. However, this optimal asset allocation is for this given setting and research and is not a guarantee of the optimal asset allocation in the future.

Table 4: Mean weights

Risk Aversion	Time Horizon	Method	Stocks	Bonds	Bills
$\gamma = 5$	4	Value	0.477	0.456	0.067
		Portfolio	0.472	0.461	0.067
	8	Value	0.519	0.420	0.061
		Portfolio	0.514	0.421	0.065
	20	Value	0.623	0.337	0.040
		Portfolio	0.584	0.372	0.044
	40	Value	0.909	0.077	0.014
		Portfolio	0.642	0.307	0.051
$\gamma = 20$	4	Value	0.195	0.272	0.533
		Portfolio	0.175	0.284	0.541
	8	Value	0.443	0.303	0.254
		Portfolio	0.199	0.321	0.480
	20	Value	0.636	0.356	0.008
		Portfolio	0.241	0.325	0.434
	40	Value	0.196	0.760	0.044
		Portfolio	0.320	0.332	0.348

First, a moderate-risk investor with a risk aversion coefficient of five. Both methods give approximately the same weights for the lower time horizons. When this given investor has a short investment horizon it is best to maintain a 50:40:10 split between stocks, bonds and bills respectively. If the investment horizon increases then the portfolio iteration method performs better, therefore these weights will be chosen. It can be seen that the relative weights between stocks and bonds change more in favour of the stocks, the optimal split would be approximately 65:30:5. This of course depends on the actual investment horizon an investor has, thus the optimal allocation is time-dependent.

Next, a low-risk investor with a risk aversion coefficient of twenty. Here it is immediately evident that a bigger portion is invested in the low-risk bills. Also, the optimal weights between the two dynamic programming methods differ more. When this given investor has a short investment horizon it seems best to maintain a 20:30:50 split between stocks, bonds and bills respectively. If the investment horizon increases then the relative weights between stocks and bills change more in favour of the stocks, and the optimal split would be approximately 30:30:40. These results are from the portfolio iteration which seems superior in these cases. The shift from wealth in bills to stocks can be explained by the reduced risk associated with stocks as the investment horizon lengthens. This is due to the stocks being able to recover from large negative returns by the increased likelihood of positive returns afterwards.

5 Conclusion

This study aimed to evaluate the effectiveness of dynamic programming methods in optimizing asset allocation among stocks, bonds, and bills, considering varying levels of investor risk tolerance and investment horizons. The results demonstrate that portfolios optimized using dynamic programming methods slightly outperformed those using simpler alternative weight-choosing strategies. The study also identified optimal asset allocation rules tailored to different investor profiles, emphasizing the importance of considering individual preferences in investment decisions.

The practical implications of these findings are significant for investors and portfolio managers. The superior performance of dynamic programming methods highlights their potential to enhance investment decision-making processes, providing more precise and effective asset allocation strategies compared to traditional approaches.

However, this study has some limitations. The assumptions made in the dynamic programming models may not account for all market variables and investor behaviours. This research only applied one testing phase, further studies could also investigate the long-term performance and robustness of these methods in different economic scenarios. Future research could also explore the application of dynamic programming methods to other asset classes and market conditions, as well as the integration of other more complex state variables.

References

Ardia, D. and Boudt, K. (2015). Testing equality of modified Sharpe ratios. *Finance Research Letters*, 13:97–104.

Brandt, M. W., Goyal, A., Santa-Clara, P., and Stroud, J. R. (2005). A simulation approach to dynamic portfolio choice with an application to learning about return predictability. *The Review of Financial Studies*, 18(3):831–873.

Campbell, J. Y., Chan, Y. L., and Viceira, L. M. (2003). A multivariate model of strategic asset allocation. *Journal of financial economics*, 67(1):41–80.

Campbell, J. Y. and Shiller, R. J. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *The review of financial studies*, 1(3):195–228.

DeMiguel, V., Garlappi, L., and Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The review of Financial studies*, 22(5):1915–1953.

Jobson, J. D. and Korkie, B. M. (1981). Performance hypothesis testing with the sharpe and treynor measures. *The Journal of Finance*, 36(4):889–908.

Kassambara, A. (2023). *ggpubr: 'ggplot2' Based Publication Ready Plots*. R package version 0.6.0.

Komsta, L. and Novomestky, F. (2022). *moments: Moments, Cumulants, Skewness, Kurtosis and Related Tests*. R package version 0.14.1.

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

Pfaff, B. (2008). VAR, SVAR and SVEC models: Implementation within r package vars. *Journal of statistical software*, 27:1–32.

R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rapach, D. E., Ringgenberg, M. C., and Zhou, G. (2016). Short interest and aggregate stock returns. *Journal of Financial Economics*, 121(1):46–65.

Statman, M. (2004). The diversification puzzle. *Financial Analysts Journal*, 60(4):44–53.

Van Binsbergen, J. H. and Brandt, M. W. (2007). Solving dynamic portfolio choice problems by recursing on optimized portfolio weights or on the value function? *Computational Economics*, 29:355–367.

Van Domelen, D. R. (2018). *stocks: Stock Market Analysis*. R package version 1.1.4.

Venables, W. N. and Ripley, B. D. (2002). *Modern Applied Statistics with S*. Springer, New York, fourth edition. ISBN 0-387-95457-0.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

A DSS

It follows the following procedure to solve the optimization problem.

1. Define a grid over the state variables.
2. Start with the last optimization at $T-1$. For each grid point simulate the returns at T and solve the optimization problem.
3. Calculate the scaled value function for each grid point.
4. Move to the decision at $T-2$ and simulate for each grid point the returns and state variables at $T-1$.
5. Determine the scaled value function $\psi(\tilde{z}_{T-1})$ for simulated state variables $\tilde{z}_{T-1}$ by interpolation of the grid values of step 3.
6. Again solve the optimization problem for each grid point and calculate its scaled value function.
7. Repeat steps 4 to 6 for decisions $T-3$ till $t+1$.

Now make a final decision for time t by performing a final simulation for the current known value of the state variable. Determine the final optimal weight x_t by using the scaled value function of $t+1$ and optimizing one last time.

B Weights

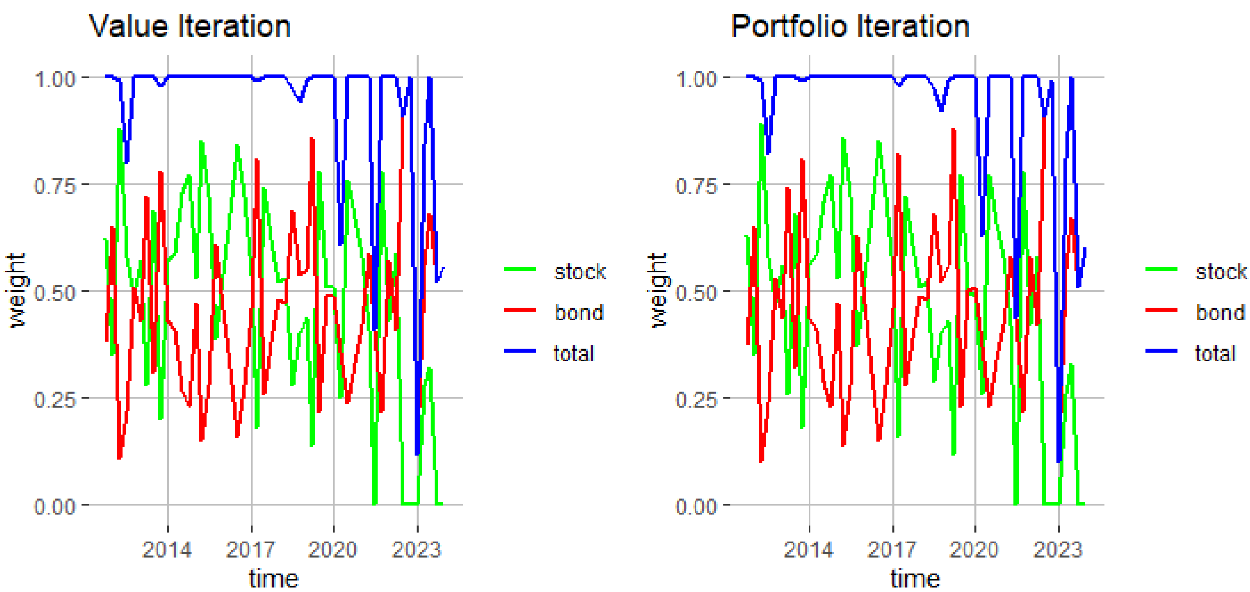


Figure 2: Results for Time Horizon 4 and Risk Aversion 5

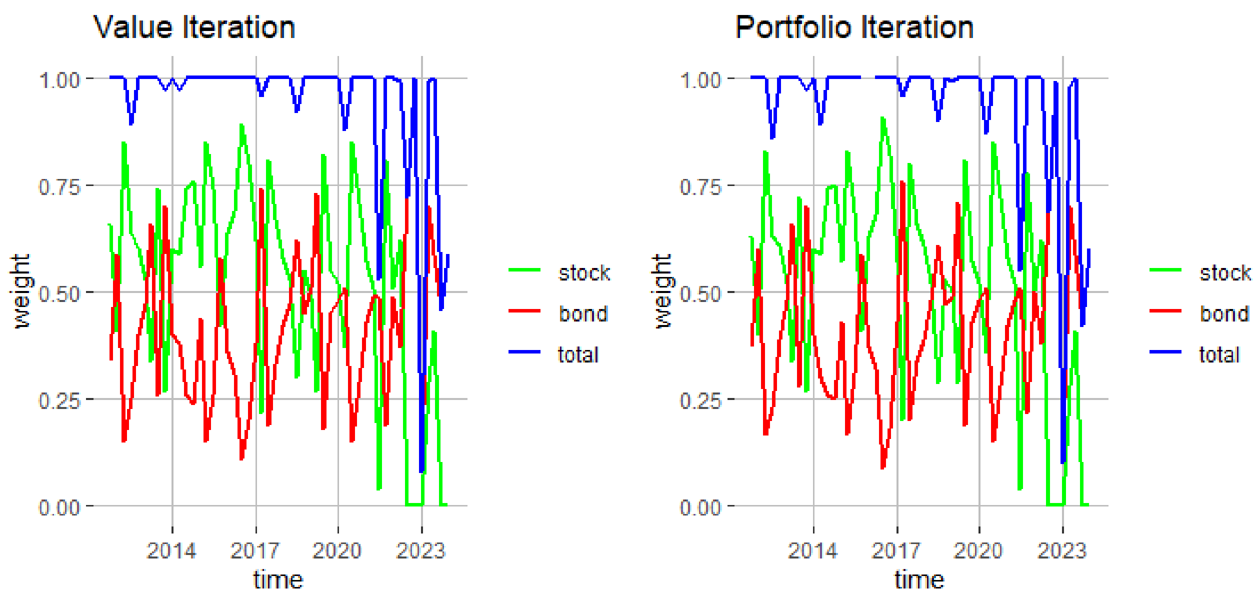


Figure 3: Results for Time Horizon 8 and Risk Aversion 5

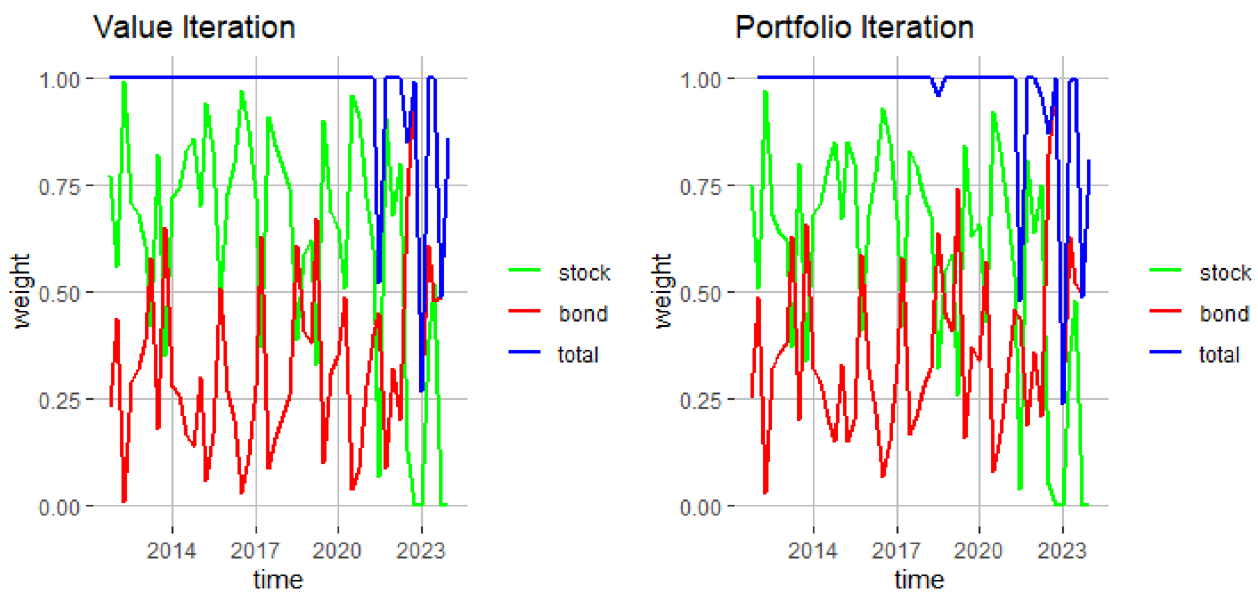


Figure 4: Results for Time Horizon 20 and Risk Aversion 5

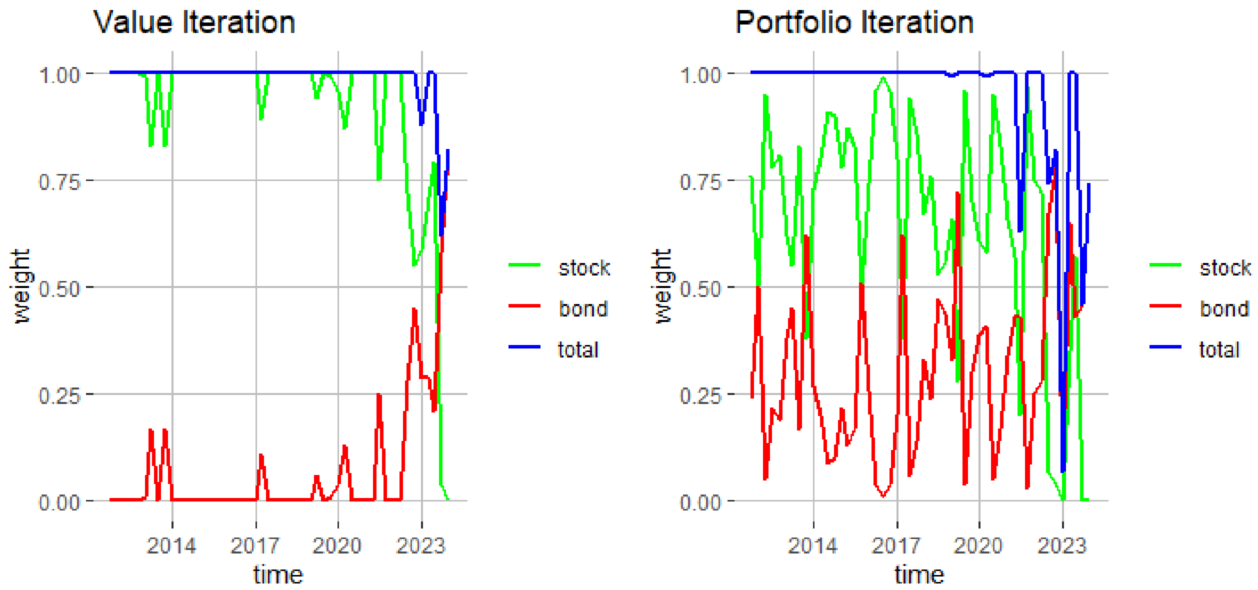


Figure 5: Results for Time Horizon 40 and Risk Aversion 5

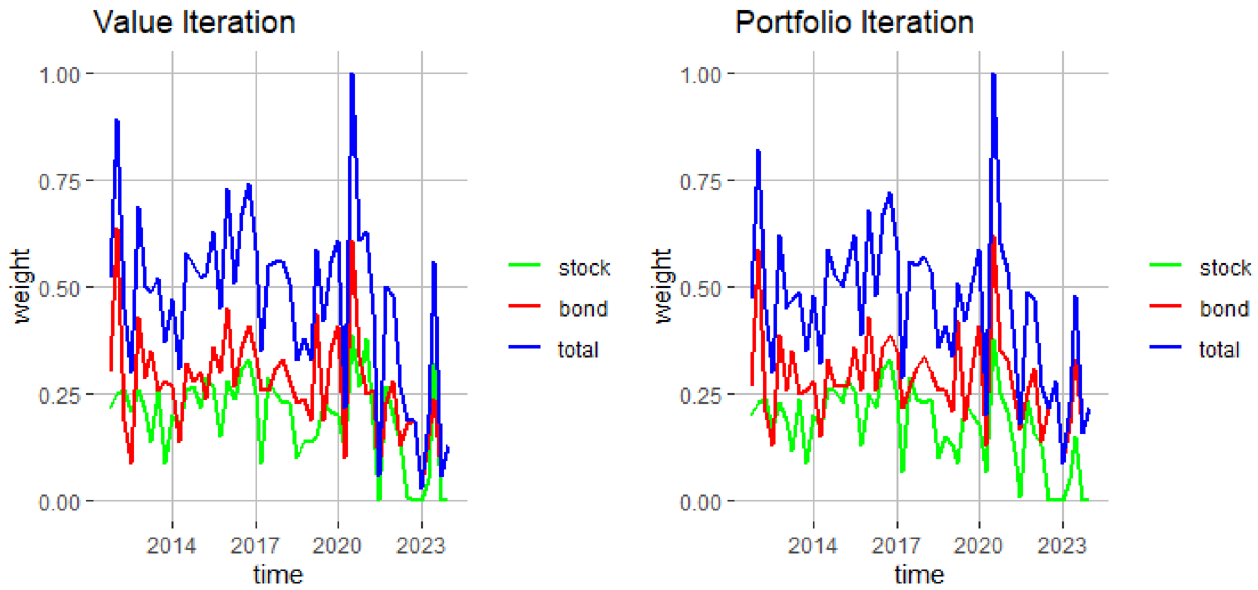


Figure 6: Results for Time Horizon 4 and Risk Aversion 20

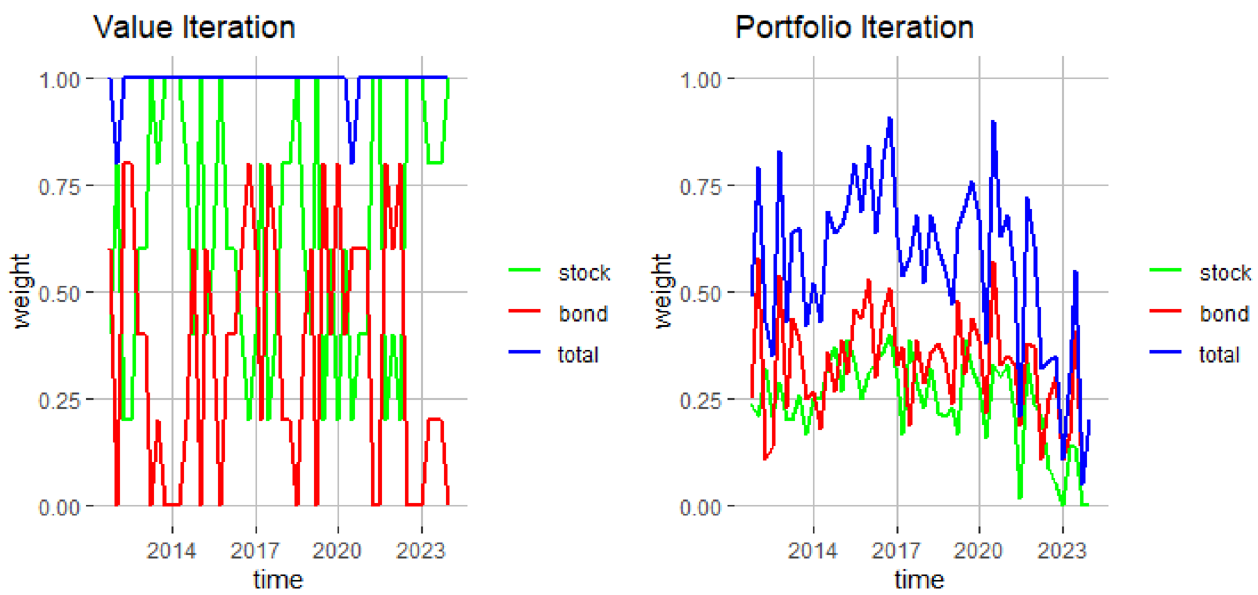


Figure 7: Results for Time Horizon 20 and Risk Aversion 20

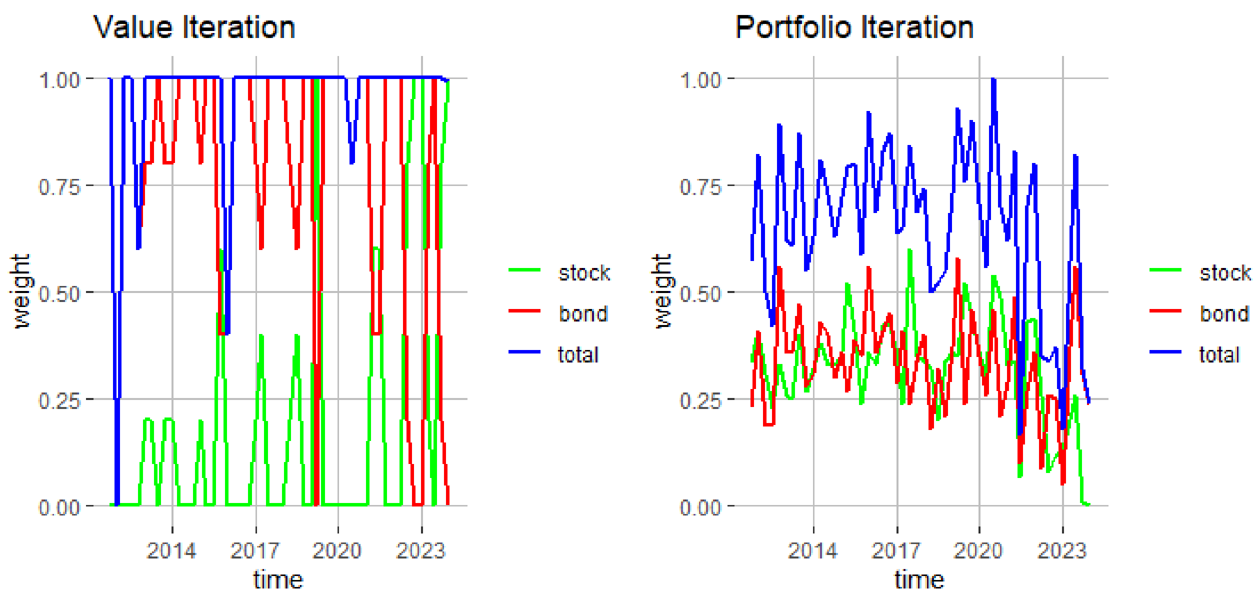


Figure 8: Results for Time Horizon 40 and Risk Aversion 20