UNIVERSITA' DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI "M.FANNO"

DIPARTIMENTO DI MATEMATICA "TULLIO LEVI-CIVITA"

CORSO DI LAUREA MAGISTRALE IN ECONOMICS AND FINANCE

TESI DI LAUREA

THE BLACK-LITTERMAN MODEL IN CONTINUOUS-TIME: ANALYSIS OF THE EFFECT OF BIASED EXPERT FORECASTS ON ASSET ALLOCATIONS

RELATORE:

CH.MO PROF. CLAUDIO FONTANA

LAUREANDO: LUCA BOSCATO

MATRICOLA N. 2019070

ANNO ACCADEMICO 2022 – 2023

Dichiaro di aver preso visione del “Regolamento antiplagio” approvato dal Consiglio del Dipartimento di Scienze Economiche e Aziendali e, consapevole delle conseguenze derivanti da dichiarazioni mendaci, dichiaro che il presente lavoro non è già stato sottoposto, in tutto o in parte, per il conseguimento di un titolo accademico in altre Università italiane o straniere. Dichiaro inoltre che tutte le fonti utilizzate per la realizzazione del presente lavoro, inclusi i materiali digitali, sono state correttamente citate nel corpo del testo e nella sezione ‘Riferimenti bibliografici’.

I hereby declare that I have read and understood the “Anti-plagiarism rules and regulations” approved by the Council of the Department of Economics and Management and I am aware of the consequences of making false statements. I declare that this piece of work has not been previously submitted – either fully or partially – for fulfilling the requirements of an academic degree, whether in Italy or abroad. Furthermore, I declare that the references used for this work – including the digital materials – have been appropriately cited and acknowledged in the text and in the section ‘References’.

Firma (signature)

TABLE OF CONTENTS

Introduction

In 1952 Markowitz presented his work on portfolio theory and established a clever quantification of the two primary goals of investing: maximizing return while minimizing risk. The main takeaway from the mean-variance formulation of the portfolio optimization problem was that a well-selected collection of assets could add up to a portfolio that would maximize expected return and minimize volatility better than any individual asset could.

Although Markowitz's mean-variance paradigm provides a foundation for modern portfolio theory, actual attempts to apply it to portfolio optimization have proven to be more challenging than initially anticipated. Mostly due to the challenge in accurately estimating the expected return and the correlation between different assets. The implication is that the mean-variance approach generates highly concentrated portfolios with a small number of assets and great sensitivity to input changes (Michaud, 1989).

The aforementioned worries suggested a somewhat demanding challenge even for expert investment managers, including the Goldman Sachs fixed income research department. Fischer Black was the first to propose using the international CAPM equilibrium as a reference point, which sparked discussions on how to choose wise investments during the optimization process. Except for using the global CAPM equilibrium as a starting point, the model, now known as the Black-Litterman Asset Allocation Model, mathematically blends qualitative and quantitative research in an optimization model. Without imposing a full set of expected returns, the investor has the freedom to state subjective opinions in absolute and relative terms apart from the reference point (Black and Litterman, 1991). Combining the equilibrium portfolio with the subjective view vector can result in a portfolio that takes into account both the implied market expectations and the investor's personal views. The model has received widespread praise in practice and is still used today by fund managers due to its intuitive portfolio composition and less extreme weights.

Recently, Davis and Lleo (2013) widened the scope of investigation of the Black-Litterman model from a static framework to continuous time, transforming the original mixed estimation problem into a filtering problem. Similarly to Black and Litterman, they used expert opinions and financial market data to formulate beliefs about future asset performance. However, they adopted a dynamic approach, whereas Black and Litterman formulated a static, single-period model. Davis and Lleo also showed how to incorporate views in a continuous time asset allocation using standard filtering techniques.

Despite this implementation and the success in the practice of the Black-Litterman model over the one proposed by Markowitz, the inclusion of expert opinions does not leave the model without some concerns. More specifically, the model ignores the effect of behavioral biases on expert views. Biases may have a significant impact on portfolio managers, analysts, or individual investors, and then affect asset allocation and portfolio construction. Since 1980 the field of behavioral finance has highlighted the importance of understanding the psychological, emotional, and social factors that influence financial decisions and has led to the development of investment strategies that aim to account for these biases.

The objective of this thesis is to analyze the impact of biased expert opinions on asset allocations by comparing different portfolios optimized with the Black-Litterman model in continuous time. Our analysis shows that debiased forecasts improve portfolio efficiency while biased forecasts produce the opposite effect. Therefore, identifying behavioral biases in expert forecasts and addressing them is essential in controlling risks in asset allocation.

The thesis is organized as follows. The first chapter introduces the literature around the Black Litterman model and its implementations in continuous time. Then, an overview of the main behavioral biases is given and some methods to debiased them are provided. In the last section of the first chapter is presented the portfolio optimization problem and some metrics to measure the performance of a portfolio. The second chapter is focused on deeply presenting the mathematics behind the Black Litterman model in continuous time following the one proposed by Davis and Lleo (2013). In this chapter, the main components of the model are introduced such as the parameters of the model, how to debias the expert views, the Kalman filter, and the stochastic problem. The third chapter is devoted to the practical implementation of the model which is then compared with other portfolios to analyze the impact of biased expert views on asset allocations.

Chapter 1
The Black-Litterman Model

In this chapter, we aim at introducing an overview of the literature regarding the Black-Litterman model and its extension in continuous time with a particular focus on the BLCT model presented by Davis and Lleo (2013). The major components of this model will be analyzed in detail during all the sections of the chapter providing a deeper understanding of the literature that led to the formulation and study of the model.

We first provide a brief introduction to the Black-Litterman model and the literature around the model. Second, we present the evolution of the Black Litterman model in continuous time comparing the models with expert opinions arriving in discrete time and those with expert opinions arriving in continuous time. Third, we analyze the main behavioral biases affecting expert opinion and provide a solution to address them in a risk-sensitive asset allocation problem. We then describe the portfolio optimization strategies in continuous-time underlying the risk-sensitive strategy utilized by Davis and Lleo to construct their model. Finally, we provide an overview of the portfolio performance measures that we utilize in Chapter 3 to analyze our model implementation.

1.1. Black-Litterman Asset Allocation Model

The Black-Litterman Global Asset Allocation model was developed in 1990 by Fischer Black and Robert Litterman at Goldman Sachs to structure international bond portfolios in a manner consistent with the portfolio manager's unique view of markets. The model was expanded in Black and Litterman (1991, 1992), and further developed in He and Litterman (1999), and Litterman (2003).

The Black-Litterman asset allocation model is a refined portfolio construction method that addresses the issues of unintuitive, overly concentrated portfolios, sensitivity to inputs, and maximizing estimation error, which are common criticisms of traditional mean-variance optimization as a portfolio optimization strategy that aims to maximize returns for a given level of risk. The Black-Litterman model uses a Bayesian method to merge the market equilibrium vector of expected returns (the prior distribution) with an investor's personal views on the expected returns of one or more assets (view distribution) to create a new, mixed estimate vector of expected returns. The new vector of return (the posterior distribution) results in portfolios that are more reasonable in terms of portfolio weights and are more intuitive.

More in detail, the Black-Litterman model blends two sources of information to form an expected return formula. The first source is derived quantitatively, it represents the expected returns that follow from the Capital Asset Pricing Model (CAPM) and are assumed to hold when the market is in equilibrium. These CAPM returns provide a foundation for the process and are used to temper potentially extreme views from the second source of information. The second source of information is the investor's personal views, which are based on information not available to the market and may differ from the equilibrium expectations. The views of the investor are used to tilt the equilibrium views, they provide information to invest less or more in a certain asset, then would follow from the equilibrium views. Combining these two sources of information results in a new vector of expected returns that can then be used in the portfolio optimization process.

Mathematically, Black and Litterman define view portfolios, specify expected returns and degrees of confidence in the view portfolios, and apply the following Black-Litterman formula:

$$\mu^* = [(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1}[(\tau\Sigma)^{-1}\Pi + P'\Omega^{-1}Q]$$

The expected excess return vector, μ^*, is obtained from the information in k views:

$$P\mu = Q + \varepsilon$$

and in a prior reflecting equilibrium:

$$\mu = \Pi + \varepsilon^e$$

In these formulas, Q is a k-vector expressing the expected excess returns on the k view portfolios. Π is the n-vector of equilibrium risk premiums. P is a $k \times n$ matrix specifying k view portfolios in terms of their weights on the n assets. Ω is the covariance matrix of the random variables to account for the uncertainty in the investor's views. Finally, the scaling factor τ is used to adjust the covariance matrix of returns in order to determine the covariance matrix of the zero-mean distribution for ε^e.

Figure 1 below shows the process leading to the new combined return vector.

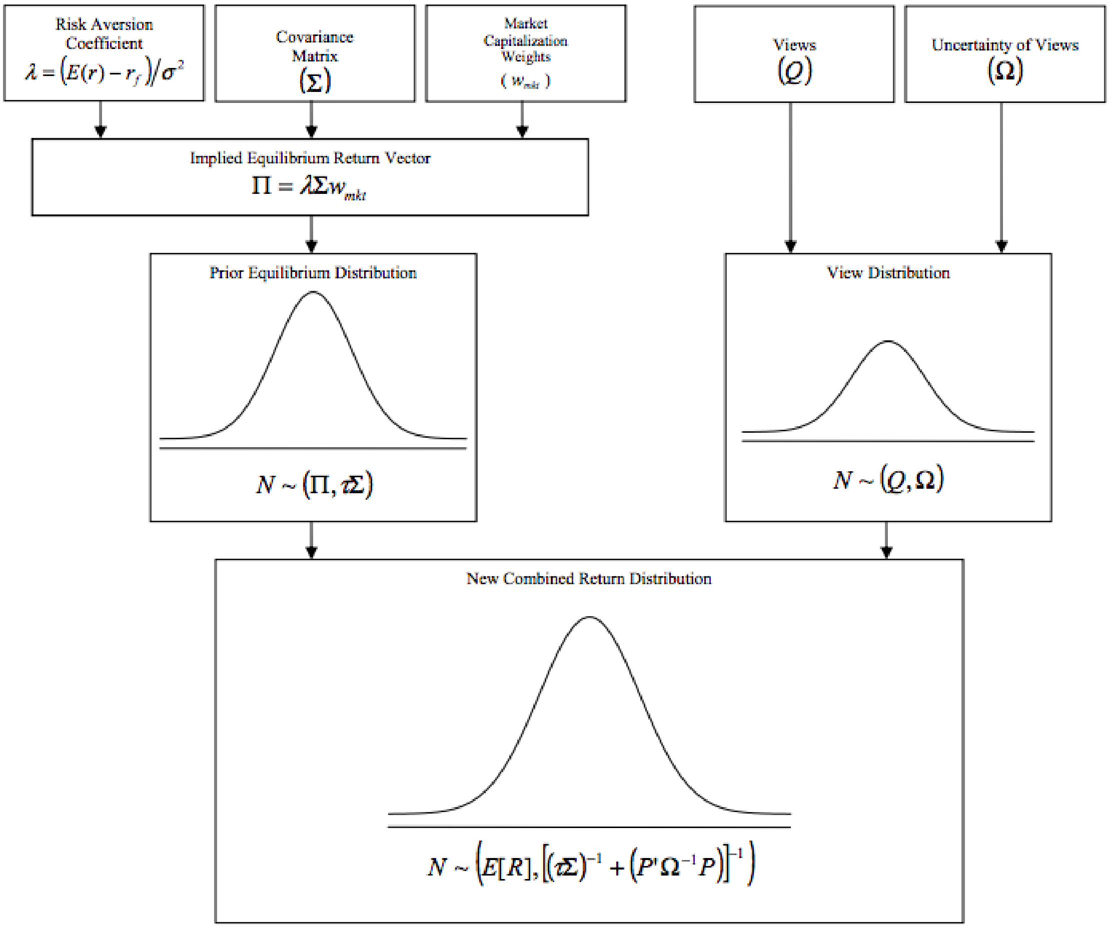


Figure 1 – How to derive the New Combined Return Vector(E[R])
Source: Idzorek (2002)

The papers published by Black and Litternan (1991,1992) provide a good overview of the features of the model and some information on the derivation, however, do not show all the formulas or a full derivation. Similarly, He and Litterman (1999) provide more detail on the workings of the model but do not present a complete set of formulas.

Other authors such as Idzorek (2005), Bertsimas et al. (2012) and Mankert (2012) provide a detailed description of the BL model and suggest some implementation.

Idzorek (2005) provides step-by-step instructions to implement the model and introduces a new method for controlling the tilts and the final portfolio weights caused by views. More specifically, he introduces a technique for specifying Ω (variance of the view) such that the impact of the shrinkage was controlled by the user-specified confidence level based on an intuitive 0% to 100% confidence level. His technique can be applied to the canonical Black-Litterman model because it is sensitive to the value of τ specified by the investor, even if his paper used the Alternative Reference model.

Bertsimas et al. (2012) provide a new perspective on the Black-Litterman model. They offer a rigorous inverse optimization interpretation of the BL model. In particular, they introduce a method to evaluate the consistency of the investor's views with the prior estimate by comparing the weights of the view portfolio to the eigenvalues of the prior covariance matrix. They also provide computational methods for estimating the new "BL"-type estimators and the corresponding portfolios. The first is a mean-variance inverse optimization (MV-IO) portfolio; the second is a robust mean-variance inverse optimization (RMV-IO) portfolio. Using numerical simulation and historical backtesting, they show that both methods often demonstrate a better risk-reward trade-off than their Black-Litterman (BL) model counterparts and are more robust to incorrect investor views.

Mankert and Sailer (2012) provide a detailed walkthrough of the model and discuss the impact of overconfidence on the original Black-Litterman model. In the model, the user indicates levels of confidence associated with each asset view in the form of confidence intervals. The authors of the model point out that people tend to be overconfident in financial decision-making, particularly when expressing confidence intervals, which can be a challenge for the implementation of the model.

Recently, many authors widened the scope of investigation from a static framework to continuous time, transforming the original mixed estimation problem into a filtering problem.

1.2. Continuous-Time Models

The recent literature on expert opinions in continuous-time portfolio selection can be divided into two sections. In one, experts express views on the current drift of financial securities at discrete, possibly random, times (Frey et al., 2012, 2014; Gabih et al., 2014; Sass et al., 2017). In the other, experts provide predictions on the evolution of factors over the investment horizon (Davis & Lleo, 2013, 2016, 2020). Both sections are closely connected as shown in the work proposed by Sass et al. (2021).

1.2.1. Expert Views at Discrete Time

The first papers addressing Black-Litterman in continuous time are Frey et al. (2012) (2014) who study expert opinions in the context of a dynamic portfolio optimization problem in continuous time. They widen for the first time the Black-Litterman model's scope of investigation from a static framework to continuous time, transforming the original mixed estimation problem into a filtering problem. More specifically, they consider the case in which the drift is modeled as a continuous time Markov chain with time-discrete expert opinions, and they transform the original problem into an optimization problem under full information by using stochastic filtering. They show that the value function of the control problem is a viscosity solution of the dynamic programming equation.

Sass et al. (2017) with a generalization of the results derived by Gabih et al. (2014) study the optimal trading strategies in a financial market with multidimensional stock returns where the drift is modeled as an Ornstein-Uhlenbeck process with time-discrete expert opinions. In their paper, the optimal trading strategy of investors maximizing the expected logarithmic utility of terminal wealth depends on the filter which is the conditional expectation of the drift given the available information. They rely on a basic theory about the matrix Riccati equation since the conditional covariance matrices of the filter follow ordinary differential equations of the Riccati type. They first consider the asymptotic behavior of the covariance matrices for an increasing number of expert opinions on a finite time horizon. Then, they state conditions for the convergence of the covariance matrices on an infinite time horizon with regularly arriving expert opinions. Finally, they derive the optimal trading strategy of an investor and find that the value function is a function of the conditional covariances matrix.

All these papers can be seen as a continuous-time version of the static Black-Litterman approach which combines an estimate of the asset returns with expert opinions on the performance of the assets. These papers share the commonality of analyzing the scenario in which experts' opinions are expressed at a discrete time.

On the other side, Davis and Lleo (2013) propose the implementation of the Black-Litterman model in a continuous time setting with expert opinions arriving continuously over the investment horizon. This model will be explained more in detail in the next section and chapter. In particular, the literature regarding the key elements of the model will be presented in this chapter, while the mathematics behind the model will be the foundation of Chapter 2.

Most recently, the work of Saas et al. (2021) shows the strict connection between the cases for experts' opinions arriving at discrete time and in continuous-time. They study a financial market where returns depend on an unobservable Gaussian drift process and expert opinions arrive at discrete times. More specifically, they consider a model where information dates are deterministic and equidistant and another model where the information dates arrive randomly as the jump times of a Poisson process. In both cases, they derive limit theorems stating that the information obtained from observing the discrete-time expert opinions is asymptotically the same as that from observing a certain diffusion process. The latter can be interpreted as the case of expert opinions arriving continuously over time. For estimating the hidden drift, they use a filter and study in detail the asymptotic behavior of the filter as the frequency of the arrival of expert opinions tends to infinitely. However, they find that as the frequency of expert opinions increases, the variance of expert opinions becomes larger. Their paper provides an important contribution to the literature because shows how the Black-Litterman model in continuous time (BLCT) proposed by David and Lleo (2013) can be obtained as a limit of the models with discrete-time experts.

1.2.2. Expert views in Continuous-Time

Davis and Lleo (2013) propose for the first time the implementation of the Black-Litterman model in a continuous-time setting with expert opinion arriving continuously. More specifically, Davis and Lleo consider the case in which the drift is modeled as an Ornstein-Uhlenbeck process with continuous time expert opinions. They formulate a dynamic approach where financial market data and expert opinions are merged to estimate the unobservable factors driving asset returns.

According to their model, experts views on the evolution of the risk premia are expressed at initial time only, and each forecast contains three pieces of information: the risk factor, a central view expressing the trajectory of the factor, and a confidence interval around the forecasts. To perform the aggregation and blend data with views they apply linear filtering. In particular, a Kalman filter is used to estimate the unobservable risk factors. Finally, they solve a risk-sensitive stochastic control problem, which uses the filter estimate, to optimize the portfolio. The key to their model is that "the filtering problem and the stochastic control problem are effectively separable. This insight allows incorporating analyst views and non-investable assets as observations in the filter even though they were not present in the portfolio optimization" (Davis and Lleo, 2013). The model has four key components: the financial market, the views, the linear filter and the stochastic control problem. The mathematics behind the model will be presented more in detail in the next chapter.

Davis and Lleo proposed several extensions to their model. Davis and Lleo (2020) take into consideration that not all experts' opinions are equally accurate since behavioral biases affect expert forecasts. Most of the literature related to the Black-Litterman model, such as He and Litterman (1999), Bertsimas et al. (2012) and Idzorek and Kowara (2013), ignore the effect of behavioral biases on analyst views. Mankert and Seiler (2012) represent an exception because discuss the impact of overconfidence on the original Black-Litterman model. However, they do not provide any practical solution to address this bias. To address the effect of behavioral biases on the formulation of experts' forecasts, Davis and Lleo (2020) build a new model based on their previous BLCT model where behavioral biases are handled through a process called debiasing. Their findings are consistent with the literature on behavioral finance and show that carefully formulated debiased forecasts improve portfolio efficiency while biased forecasts produce the opposite effect. More specifically, they find that "biases have a significant impact on portfolios, explaining nearly 70% of excess risk-taking in their implementation". Therefore, identifying behavioral biases in expert forecasts and addressing them is essential in controlling risks in asset allocation.

Since behavioral biases strongly impact expert opinions, the next section presents a detailed description of the most important behavioral biases, which we also take into consideration in our model implementation, and a solution to address them in a risk-sensitive asset allocation problem.

1.3. Behavioral Biases and Debiasing

Over the past 40 years, behavioral finance has shown that human beings are subject to behavioral biases and psychological pitfalls (Kahneman and Tversky, 2000). However, there are different opinions on the exact number and classification of those biases. Hirschleifer (2001), in his study of the impact of investor psychology on asset pricing, classifies 22 different psychological biases into four categories: self-deception, heuristic simplification, emotion/affect, and social. Shefrin (2005, 2008, 2010) identifies 12 main psychological pitfalls.

However, all authors agree on the major biases such as overconfidence, excessive optimism, conservatism, confirmation bias, and groupthink. Davis and Lleo (2020) address these five main psychological biases and apply general modeling principles based on Shefrin (2002, 2016) to counteract them. Removing the effect of behavioral biases from the forecasts is crucial to obtaining well-specified forecasts and is a process known as debiasing. The same biases are taken into consideration in our model implementation in Chapter 3.

The first bias is overconfidence. Kahneman and Riepe (1998) argue that overconfidence causes people to overestimate their knowledge, undervalue risks, and overestimate their ability to control events. Many researchers find evidence for the presence of overconfidence bias in different financial decisions. Doukas and Petmezas (2007) examine whether superior abnormal returns can be generated in the case of acquisitions by overconfident managers. They also investigate if self-attribution could be a source of managerial overconfidence. Allen and Evans (2005) determine the extent of trader overconfidence using experimental bidding data. Odean (1998) argues that overconfident investors engage in trading more readily as compared to rational investors. Overconfidence is not limited to secondary market traders but also affects primary market investors. Hsu and Shiu (2010) found that frequent investors in discriminatory auctions in the Taiwan stock market tend to underperform infrequent investors. This suggests that overconfidence may cause analysts to overestimate the precision of their views, leading to overly narrow confidence intervals. The natural solution proposed by Davis and Lleo (2020) is to widen the confidence interval.

The second bias is excessive optimism. According to Malmendier and Tate (2005), confidence and optimism differ. They suggest that confidence is related to skill-related outcomes while optimism is connected with exogenous outcomes. Thus, optimism is about expecting a favorable outcome irrespective of the actual effort or skills devoted by the individual (or group) to bring about the outcome. According to Ramnath et al. (2008), overoptimism is the tendency to overvalue the possibility of desired outcomes and undervalue the occurrence of unfavorable events. Thus, excessively optimistic analysts will overestimate the probability of scenarios that they perceive as positive. To address this bias, Davis and Lleo (2020) propose to widen the confidence interval to reflect the possibility that the actual realization may differ significantly from the prediction.

The third bias is conservatism. It has been observed that the degree to which an individual adjusts her beliefs is commonly less than would be expected in a Bayesian model of belief revision. This phenomenon in which a person underreacts to new information is known as conservatism bias. It is a specific type of belief perseverance error that is generally held to be an extension of Tversky and Kahneman’s theory on anchoring and adjusting. Conservatism is a phenomenon in which a person underreacts to new information. According to David and Lleo (2020), this bias affects the point estimate given by analysts, as well as the confidence interval and it is a serious concern, especially for multiperiod and continuous time models. However, their model does not require analysts to update their views; views are formulated at the initial stage when the model is parameterized, and analysts are not asked to update them later. Thus, once formulated at the initial stage, expert forecasts are processed by the Kalman filter in a Bayesian manner, reducing gradually the effect of conservatism.

The fourth bias is confirmation bias. This bias refers to the trend of acquiring or assessing new information in a way it is consistent with the person’s pre-existing beliefs. Thus, individuals don’t take divergent information into account (Schwind et al. 2012). Such a phenomenon can be described as the capacity people have to convince themselves about everything they want to believe in. The confirmation bias may overemphasize decision-makers’ beliefs and make them underestimate important information that leads to evidence opposite to their positions, thus, impairing the decision (Pompian, 2012). Davis and Lleo (2020) address this bias through the Kalman filter which weighs forecasts based on their accuracy.

The last bias is groupthink. Groupthink occurs when members of a group reach a consensus of opinion without considering alternative solutions. This usually happens because they don't want to upset the "status quo" of a situation because consensus has otherwise been reached. As a result, people can reach faulty conclusions as a group, and investors are no exception. The term "groupthink" was originally invented by Yale University social psychologist Irving Janis in 1972. According to Janis, “groups of intelligent people sometimes make poor decisions because groups prevent contrary information from being given the proper level of due diligence”. This bias was a well-studied phenomenon in behavioral finance. Groupthink can reduce anxiety for investors in the short term, but in the long- term it creates herding behavior that can lead to bubbles. This bias frequently leads to overreaction to new information creating significant market inefficiencies such as post-earnings announcement drift. To address this bias, Davis and Lleo (2020) consider dissenting analysts whose views differ markedly from the majority, to ensure heterogeneity of the expert pool. They also add a stress test scenario as an additional view to broaden the range of forecasts.

Behavioral biases have a strong impact on individuals, and debiasing becomes an essential procedure in the model to unbiased expert opinion and to obtain more precise results.

After expert opinions have been debiased, the next step in the process is to blend those opinions with market data to estimate the current factor level. Davis and Lleo (2013) apply linear filtering to perform the aggregation and blend data with views. More specifically, they utilize the Kalman-Bucy linear filter which will be described more in detail in the next chapter.

Once the factors estimates have been estimated, all the inputs for the stochastic portfolio optimizer are available. In discrete time models, stochastic programming algorithms are used. In continuous time, the methods of choice come from the stochastic control theory.

1.4. Portfolio Optimization

The problem of optimal investment-consumption in continuous-time was presented for the first time by Robert Merton (1969, 1971) who provides a more advanced approach to portfolio selection that accommodates the risk aversion of the investor and is not subject to the static nature of the Markowitz mean-variance approach. In his papers, Merton extends the model previously presented by Samuelson (1969) who considers a discrete-time consumption-investment model with the objective of maximizing the overall expected utility of consumption. Using a dynamic stochastic programming approach, he succeeds in obtaining the optimal decision for the consumption-investment model. Merton (1969) extends the model of Samuelson (1969) to a continuous-time framework and uses stochastic optimal control methodology to obtain the optimal portfolio strategy. Under certain assumptions for the preference structure and asset price dynamics, Merton obtains a closed-form solution to the optimal asset allocation problem, which devised investing a constant proportion in a risky asset. This constant proportion is also known as the Merton ratio.

All theories of continuous-time optimal investment developed since then, including duality (Karatzas et al., 1987; Karatzas and Shreve, 1998) and risk-sensitive investment management (Bielecki and Pliska, 1999; Kuroda and Nagai, 2002; Davis and Lleo, 2014) are direct descendants of these early papers.

Duality was introduced by Cox and Huang (1989) and Karatzas et al. (1987). Under complete market assumptions, they show how the portfolio choice problem could be decomposed into two subproblems. The first subproblem solved for the optimal terminal wealth and it could be formulated as a static optimization problem given the complete market assumption. The second subproblem was then solved to find the trading strategy that would replicate the optimal terminal wealth. This new approach helped to expand the class of dynamic problems that could be solved. Dual methods were then used by a number of authors (Xu, 1990; Shreve and Xu, 1992a, 1992b; Cvitanic and Karatzas, 1992; Karatzas et al., 1991; He and Pearson, 1991a, 1991b) to extend the martingale approach to problems where markets are incomplete, and agents face portfolio constraints. Duality methods have since been very popular for tackling other classes of portfolio optimization problems. For example, models where trading impacts security prices and problems with transaction costs. Applying some of these dual methods, Haugh, Kogan and Wang (2006) show how suboptimal dynamic portfolio strategies could be evaluated by computing lower and upper bounds on the expected utility of the true optimal dynamic trading strategy. In general, a better suboptimal solution is shown by a narrower gap between the lower and upper bounds, explaining how far the sub-optimal strategy is from optimality. These techniques apply directly to multidimensional diffusion processes with incomplete markets and portfolio constraints such as no-short selling or no borrowing constraint

On the other side, risk-sensitive control is a generalization of classical stochastic control in which the degree of risk aversion or risk tolerance of the optimizing agent is explicitly parameterized in the objective criterion and directly influences the outcome of the optimization. In risk-sensitive control, the objective of the decision maker is to select a control policy *h(t)* to maximize the criterion:

$$J(x,t,h;\theta) := -\frac{1}{\theta}\ln \boldsymbol{E}\left[e^{-\theta F(t,x,h)}\right]$$

where x is the state variable, t is the time, F is a given reward function, and the risk sensitivity $\theta \in]-1,0[\cup]0,\infty)$ represents the decision maker's degree of risk aversion. A Taylor expansion of the previous expression around $\theta = 0$ evidence the vital role played by the risk sensitivity parameter:

$$J(x,t,h;\theta) = \boldsymbol{E}[F(x,t,h)] - \frac{\theta}{2}\boldsymbol{Var}\,[F(x,t,h)] + O(\theta^2)$$

This criterion amounts to maximizing $\boldsymbol{E}\,[F(x,t,h)]$ subject to a penalty for variance. Hence risk-sensitive control explicitly models the risk-aversion of the decision maker as an integral part of the control framework differentiating from traditional stochastic control that imports the risk aversion in the problem through an externally defined utility function.

Bielecki and Pliska (1999) are the first to apply continuous time risk-sensitive control as a practical tool that could be used to solve real-world portfolio selection problems. They propose the logarithm of the investor's wealth as a reward function, that is $F(x,t,h) = ln\,V(t)$ where h is the investment strategy determining the portfolio process $V(t)$. This results in the risk-sensitive asset management criterion:

$$J(x,t,h;\theta) := -\frac{1}{\theta}\ln\boldsymbol{E}\,[e^{-\theta lnV(t)}] = -\frac{1}{\theta}\ln\boldsymbol{E}\,[V(t)^{-\theta}]$$

Thus, the investor's objective is to maximize the risk-sensitive (log) return of his/her portfolio or alternatively to maximize a function of the power utility (HARA) of terminal wealth. The contribution of Bielecki and Pliska to the field is immense: they studied the economic properties of the risk-sensitive asset management criterion (2003), extended the asset management model into an intertemporal CAPM (2004), worked on transaction costs (2000), numerical methods (2002) and considered factors driven by a CIR model (2005).

A major contribution was made by Kuroda and Nagai (2002) who introduce an elegant solution method based on a change of measure argument that transforms the risk-sensitive control problem into a linear exponential of the quadratic regulator. They solve the associated Hamilton-Jacobi-Bellman (HJB) PDE over a finite time horizon and then study the properties of the ergodic HJB PDE. Davis and Lleo (2008) apply this change of measure technique to solve a benchmarked investment problem in which an investor selects an asset allocation to outperform a given financial benchmark and analyzes the link between optimal portfolios and fractional Kelly strategies.

The risk-sensitive control theory is also used by Davis and Lleo (2020) to optimize the investment strategy after the unknown factors have been estimated. They apply the mathematical results derived by Kuroda and Nagai (2002) and Davis and Lleo (2014) to solve the risk-sensitive stochastic control problem. The mathematics behind the process will be presented more in detail in the next chapter. Before that, a mention of the Kelly criterion and the fractional Kelly strategies is required since they hold an important place in investment management theory.

1.4.1. Kelly Criterion and Fractional Kelly Strategies

The Kelly criterion was originally developed by John Kelly, a researcher at Bell Labs, to analyze long-distance telephone signal noise. The Kelly criterion is a mathematical formula for bet sizing, which is frequently used by investors to decide how much money they should allocate to each investment or bet through a predetermined fraction of assets. More specifically, the Kelly criterion maximizes the log return on invested wealth and is therefore related to the seminal work of Bernoulli (1738) reported by the Journal of the Econometric Society (1954). Early contributions to the theory and application of the Kelly criterion to gambling and investment include Kelly (1956), Latané (1959), Breiman (1961), Thorp (1971) or Markowitz (1976). From a practical investment management perspective, several of the most successful investors, including Keynes, Buffett and Gross have used Kelly-style strategies in their funds (Ziemba, 2005; Thorp, 2006; Ziemba, 2007).

The Kelly criterion has a number of good as well as bad properties, as discussed by MacLean, Thorp and Ziemba (2010). Its ‘good’ properties extend beyond practical asset management and into asset pricing theory, as the Kelly portfolio is the numéraire portfolio associated with the physical probability measure. In terms of ‘bad’ properties, Samuelson was a long-time critic of the Kelly criterion (1969, 1971, 1979) showing that it is inherently a very risky investment.

According to Davis and Lleo (2013), the objective of a Kelly investor with a fixed time horizon T is to maximize:

$$J(t; h; T) = \boldsymbol{E}[U(V_T)] = \boldsymbol{E}[lnV_T]$$

where $h(t)$ is a control process and V_T is the wealth at time T.
A pointwise maximization of the criterion J produces the Kelly portfolio:

$$h^* = (\Sigma\Sigma')^{-1} (\mu - r\mathbf{1})$$

A more detailed explanation of the Kelly criterion based on David and Lleo (2013) will be provided in Chapter 2.

To reduce the risks associated with the Kelly investment strategy, MacLean, Ziemba, and Blazenko (1992) suggest a modified approach called the fractional Kelly strategy. This strategy involves investing a fraction (f) of one's wealth in the Kelly portfolio and the remaining proportion $(1-f)$ in a risk-free asset. According to MacLean, Sanegre, Zhao, and Ziemba (2004), the fractional Kelly strategy has two main advantages. Firstly, it is significantly less risky than the full Kelly portfolio but still allows for significant potential gains. Second, in a continuous time setting where asset prices follow a geometric Brownian motion, a fractional Kelly strategy is optimal with respect to Value at Risk and a Conditional Value at Risk criteria. Indeed, fractional Kelly strategies correspond to the optimal investment of a power utility investor seeking to maximize the terminal utility of his/her wealth. MacLean, Ziemba and Li (2005) further prove that fractional Kelly strategies are efficient when asset prices are lognormally distributed.

According to Davis and Lleo, this result is a corollary to Merton's Fund Separation theorem. In particular, the optimal control h^* is given by:

$$h^* = \frac{1}{1-\gamma}\,(\Sigma\Sigma')^{-1}\,(\mu - r\mathbf{1})$$

which represents a position of $\frac{1}{1-\gamma}$ in the Kelly criterion portfolio.

Therefore, fractional Kelly strategies represent a consequence of a classical Fund Separation Theorem which states that an investor can separate his asset allocation between the Kelly (log-utility) portfolio and the risk-free rate. Moreover, if an investor has a risk sensitivity γ, the proportion of the Kelly portfolio will equal $\frac{1}{1-\gamma}$.

Unfortunately, this implies that fractional Kelly strategies are no longer optimal when the basic assumptions of the Merton model, such as the lognormality of asset prices, are removed. Most recently, many authors have tried to extend the definition of fractional Kelly strategies to guarantee their optimality.

Davis and Lleo (2011) show how the definition of fractional Kelly strategies can be extended to guarantee optimality. In their paper, they present an overview of the Kelly investment strategies in an incomplete market environment where asset prices are not lognormally distributed. The key idea of their paper is to get the definition of fractional Kelly strategies to coincide with the fund separation theorem related to the problem at hand. In these instances, "fractional Kelly investment strategies appear as the natural solution for investors seeking to maximize the terminal power utility of their wealth".

Most of this literature has viewed the Kelly portfolio as objective and universal. Davis and Lleo (2020) show that this is not necessarily true. In their paper, they show that the Kelly portfolio is a personal portfolio that depends on the views. As a result, even if two investors have the same investment universe and data, their Kelly portfolios could differ if their forecasts are not the same. From this result, Davis and Lleo show that the universal portfolio aggregation results, such as mutual funds theorems, hold only within a restricted set of model assumptions. Thus, the preference-free nature of the Personal Intertemporal Hedging Portfolio (PIHP) and the universality of the Kelly portfolio vanish when more general problems are taken into consideration, such as including stochastic state variables and interest rates or moving to a personal decision framework with expert forecasts. However, according to Davis and Lleo the Kelly portfolio's properties still hold, despite the loss of universality. The Kelly portfolio is still growth optimal. Hence, it is still a numéraire portfolio (Long, 1990; Becherer, 2001) but associated with a subjective probability measure.

Once the stochastic problem is solved and the initial portfolio allocation is set and implemented, monitoring begins. To understand the impact of analyst views and behavioral bias on the portfolio's structure, Davis and Lleo (2020) compare several strategies in the same investment universe. This requires utilizing some portfolio performance measures.

1.5. Portfolio Performance Measures

Measurement and evaluation of portfolio performance is a key step in the investment management process. The most used portfolio performance metrics based on the simulated weekly excess returns include summary statistics (mean, standard deviation and semi-deviation), tail risk measures (VaR and CVaR), and portfolio efficiency measures (Sharpe and Sortino ratios).

In chapter 3, we evaluate the portfolio weekly return distribution for a certain horizon of time with the classical statistics metrics as mean (μ), standard deviation (σ) and semi-deviation.

The standard deviation indicates the dispersion of returns for a given security or market index. However, standard deviation penalizes both the upside and the downside potential of portfolio return. Thus, it is not a very appropriate choice as a measure of performance.

An alternative measurement to standard deviation is semi-deviation. However, unlike the previous measure, semi-deviation looks only at negative price fluctuations and is most often used to evaluate the downside risk of an investment.

The formula for semi-deviation is:

$$Semi - deviation = \sqrt{\frac{1}{n} \sum_{r_t < Average}^{n} (Average - r_t)^2}$$

where n is the total number of observations below the mean, r_t is the observed value and $Average$ is the mean or target value of the data set. Semi-deviation can be used to evaluate an investor's entire portfolio showing the worst-case performance that can be expected from a portfolio, compared to the losses in an index or whatever comparable is selected.

The measures just presented can be very unsatisfactory risk measures when we are dealing with seriously non-normal distributions. We, therefore, use valid risk measures in the face of more general distributions, such as Value at Risk (VaR) and Conditional Value at Risk (CVaR).

In the most general form, VaR measures the maximum expected potential loss on a portfolio over a given time horizon for a given confidence interval. If it is assumed, for example, that the VaR of a portfolio over a one-week period is equal to $200 million with a 99% confidence level (α), this implies that the investor could expect the portfolio to exceed this loss with a probability of 1% ($p = 1 - \alpha$). Alternatively, there is only a 1% chance that the value of the asset will drop more than $200 million over the next week.

VaR depends on two arbitrarily chosen parameters: the confidence level (α), which indicates the likelihood of an outcome no worse than the VaR, and which might be any value between 0 and 1; and the holding or horizon period, which is the period of time over which the portfolio's profit or loss is measured.

VaR has several significant attractions over traditional risk measures. It provides a common consistent measure of risk because it enables investors to aggregate the risks of sub-positions into an overall measure of portfolio risk, taking account of the correlation between risk factors. It is a holistic measure because it considers all risk factors affecting the portfolio, and a probabilistic measure because it provides information on the probabilities associated with specific loss amounts. On the other hand, VaR has also some limitations. It is uninformative of tail losses because provides the higher loss at a certain probability but doesn't explain what happens if the tail event does occur. It creates perverse incentives because doesn't take into consideration low-probability, high-impact events. Finally, VaR is not sub-additive which means that in adding risks together we might create an extra residual risk that someone must bear, and that didn't exist before.

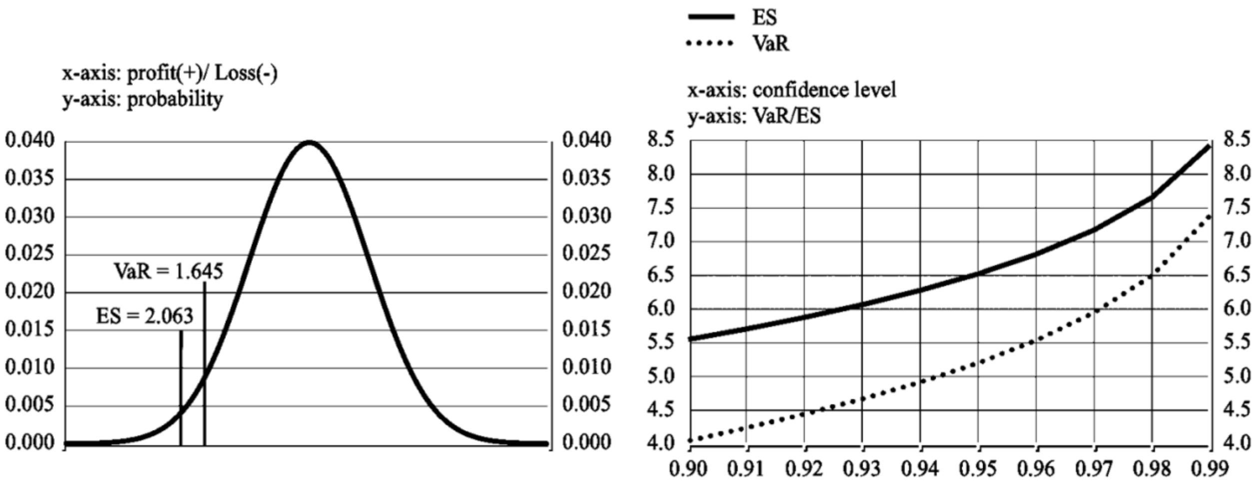
Given these problems with the VaR, a new risk measure has been developed, following the theory of coherent risk measures.

Expected shortfall (ES), or Conditional VaR (CVaR), is indicated as a coherent alternative risk measure. The ES is the average of the worst $100(1-\alpha)$ % of losses of a portfolio's profit and loss distribution. This measure provides the amount an investor expects to lose if a tail event does occur, while VaR only indicates the most the investor can lose if a tail event does not occur. Alternatively, ES is the expected loss, given a loss exceeding VaR:

$$ES = E(L|L > VaR)$$

The ES has many of the same uses as the VaR but it is considered a better measure because it is coherent and always satisfies subadditivity, while the VaR does not.

The figure below shows the value of the ES measure and VaR for a return distribution based on a hypothetical stock whose price is normally distributed with mean 0 and standard deviation equal to 1. While the second chart shows that the ES measure, like VaR, tends to rise with the confidence level.



Source: ECB calculations.
Note: This assumes a 95% confidence level, a one-day holding period and that the daily profit and loss is distributed as standard normal (i.e. mean 0 and standard deviation 1).

Source: ECB calculations.
Note: This assumes a ten-day holding period, and that daily profit and loss is distributed as standard normal (i.e. mean 0 and standard deviation 1).

Figure 2 - VaR and ES

In chapter 3, after the calculation of various risk measures, some risk-adjusted performance measures have been used to gauge the trade-off between the risks and returns of the different portfolios.

The first ratio is the Sharpe ratio which measures the relationship between the excess return and the standard deviation of the portfolio. It is simply the risk premium per unit of risk, which is quantified by the standard deviation of the portfolio. Therefore, the Sharpe ratio is used to evaluate the performance of a portfolio in relation to the level of risk that was assumed.

The formula used for the Sharpe ratio is:

$$Sharpe\ Ratio = \frac{R_P - r_f}{\sigma_A}$$

where R_P represents the expected return of the portfolio, σ_A is the associated volatility and r_f indicates the risk-free rate of the asset return on investment with zero risk. Since it utilizes the volatility of the portfolio, the Sharpe ratio focus on total volatility. Thus, the Sharpe ratio does not distinguish between upside and downside volatility.

The second ratio is the Sortino ratio. It is a variation of the Sharpe ratio that only takes into account negative returns, rather than all returns. It measures the performance of a portfolio relative to the downside risk taken. This ratio is useful for investors who are more concerned with avoiding losses than maximizing gains. Indeed, the Sortino ratio uses downside deviation (i.e., semi-deviation) rather than standard deviation as a measure of risk, thus only those returns falling below a user-specific target or required rate of return are considered risky. The Sortino ratio determines the excess returns for each unit of downside risk.

The formula for the Sortino ratio is:

$$Sortino\ Ratio = \frac{R_P - r_f}{\sigma_d}$$

where R_P is the expected return of the portfolio, r_f is the risk-free rate and σ_d is the standard deviation of the downside.

The next chapter will provide the mathematics behind our model based on the BLCT model presented by David and Lleo (2013).

Chapter 2

Mathematics behind the Behavioral BLCT Model

In this chapter, we aim to present the mathematics behind our model following the structure provided by Davis and Lleo (2013, 2020). We start by introducing the three elements composing the model. More specifically, we parametrize the financial market, collect expert opinions and address the impact of behavioral biases. Then, we present the Kalman filter and its role in combining data with opinions to estimate the unobservable factors. Since the filtering step is separable from the stochastic control problem the mathematics of the latter remains unchanged despite the addition of debiased expert views. Finally, once all the information is collected, we derive a solution to the risk-sensitive stochastic control problem and analyze the implication of the optimal investment strategy.

In the next chapter, we will implement the model and analyze the impact of behavioral biases and expert opinions on portfolio construction.

2.1. Constructing the Stochastic Model

In this section, we construct the financial market model and describe the elements composing it. We then analyze the expert views and how to incorporate them into the model addressing the impact of behavioral biases.

Our investment model consists of three types of market inputs: $\mathrm{n} \geq 0$ factors, $\mathrm{m} > 0$ risky financial securities, and $\mathrm{k} \geq 0$ expert forecasts.

The n risk factors are the underlying sources of risk driving the return of an asset class. For example, the return profile of a bond is influenced by risk elements like duration, credit spreads, and default risk, whereas the return profile of a stock is determined by elements like size, value, and momentum. Examples of market risk factors are volatility and inflation.

Risk factors were initially presented in academic financial models such as the capital asset pricing model (CAPM), which expresses the relationship between expected return and risk for stocks. Subsequently, the Fama-French model extended the capital asset allocation model by adding size risk and value risk factors to the market risk factor in CAPM. Even though risk factors affect the return, they are not directly observable. In the next section, we use filtering techniques to estimate these risk factors.

The m financial securities are either used in our model to estimate the risk factors and as an asset for the portfolio. We split them into m_1, the securities that the investor is able to invest, and m_2, the securities that the investor can only observe. Indeed, not all the securities available in the market are tradable.

Finally, we incorporate k expert forecasts to estimate the risk factors. Each forecast incorporates three pieces of information: the risk factor forecasted, a central view, and a confidence level. The central view expresses the experts' best forecast of the risk factor(s)' trajectory while the confidence interval measures how spread-out experts expect the distribution of forecasting errors to be. Before including those expert forecasts in the model we need to address the behavioral biases that may affect them.

Once they have been debiased, financial securities and expert forecasts are combined with the filter. The Kalman filter treats both securities prices and expert forecasts as one single joint observation. However, they differ because financial securities provide an online observation of the risk factors while expert forecasts provide a time 0 subjective forecast of the evolution of the risk factors over the period. Indeed, as time moves forward, new securities prices become available in real-time, and the realized return enters the Kalman filter at the time t. On the other end, the expert provides a trajectory for a risk factor, which will be used to calibrate the central view, but typically the expert does not provide a value for each t ∈ [0, T], which is what the Kalman filter needs. More details are provided in the next subsections.

Finally, the estimated risk factors derived from the Kalman filter are used to solve the risk-sensitive asset management problem.

We assume that the three inputs just presented evolve on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0}^{T}, \mathbb{P})$, on which we define a $\mathbb{R}^d$-valued $(\mathcal{F}_t)$-standard Brownian motion with elements $Wj(t)$, j = 1, ..., d, $d := n + m + k$. In this section, we construct the stochastic models for each type of input and estimate the hidden factors using a Kalman filter.

2.1.1. Financial Market

The financial market is composed of $n \geq 0$ risk factors and $m > 0$ risky financial securities that are related with each other. Indeed, the n risk factors are financial and economic factors driving the returns of financial securities. These risk factors influence the growth rate of securities prices, but they are not directly observable. Therefore, the number of risk factors is typically lower than the one of financial securities. In this subsection, we define and analyze the processes driving the evolution of these market components.

We model the n risk factors $X_1(t)$, ..., $X_n(t)$ as a time-dependent vector Ornstein-Uhlenbeck process which evolves according to the stochastic differential equation (SDE):

$$dX(t) = \big(b(t) + B(t)X(t)\big)dt + \Lambda(t)dW(t), \quad X(0) \sim N(\mu_0, P_0) \tag{2.1}$$

where the random starting value $X(0)$ is independent of the Brownian motion and the coefficients $b : [0, T] \to \mathbb{R}^n$, $B : [0, T] \to \mathbb{R}^{nxn}$, and $\Lambda : [0, T] \to \mathbb{R}^{nxd}$ are bounded[1], C^1, and Lipschitz[2] continuous. The class C^1 consists of all differentiable functions whose derivative is continuous; such functions are called continuously differentiable.

Davis and Lleo (2020) estimate the parameters *b, B* using standard econometric techniques by considering (2.1) as the continuous-time analog to a vector autoregressive process of order 1; the matrix Λ through a discretization of the quadratic variation; and the initial mean μ_0 and covariance P_0 applying econometric methods, expert opinions, or a mix of both. More details about the model parametrization are provided in the next chapter.

We model the m risky financial securities via their discounted prices. These discounted prices are computed using the money market $S_0(t)$ as numéraire. Computationally, the price $S_0(t)$ is a strictly positive, $(\mathcal{F}_t)$-measurable stochastic process. The m risky securities' discounted price follows a geometric process:

$$\frac{dS_i(t)}{S_i(t)} = (a(t) + A(t)X(t))_i dt + \sum_{j=1}^{d} \sigma_{ij} dW_j(t), \qquad S_i(0) = s_i, i = 1, \dots, m, \tag{2.2}$$

where the time-dependent coefficients $a : [0, T] \to \mathbb{R}^m$, $A : [0, T] \to \mathbb{R}^{mxn}$, and $\Sigma = (\sigma_{ij})$, $i =$ 1, ..., m; j = 1, ..., $d : [0, T] \to \mathbb{R}^{mxd}$are bounded, C^1, and Lipschitz continuous. Because discounted financial securities prices have a geometric dynamic, they are not suitable observations for the (linear) Kalman filter. Instead, we use excess log returns (or risk premiums). Indeed, the relation between discounted prices $S_i(t)$ and risk premium $\pi_i(t)$ is simply π_i(t) = ln(S_i(t)), i = 1, ..., m. Thus, we define the excess log return vector as $\mathfrak{s}_i$(t) = ln(S_i(t)), i = 1, ..., m. Then, by Itô's lemma (see Appendix B), $\mathfrak{s}$(t) solves the SDE

$$d\mathfrak{s}(t) = \Big[\Big(a(t) - \tfrac{1}{2}\, d_\Sigma\Big) + A(t)X(t)\Big]\, dt + \Sigma dW(t), \quad \mathfrak{s}(0) = ln(s), \tag{2.3}$$

[1] f is bounded if there is M > 0 such that for all x, |f(x)| ≤ M

[2] A function f : X → Y is called Lipschitz continuous with constant C if, for each $x_1, x_2 \in$ X one has d(f (x_1), f (x_2)) ≤ C · d(x_1, x_2), where d stands for the distance.

where $d_\Sigma = ((\Sigma \Sigma ')_{11} \quad (\Sigma \Sigma ')_{22} \quad \dots \quad (\Sigma \Sigma ')_{mm})'$.

As in David and Lleo (2020), the coefficients of (2.3) are generally estimated from financial markets using standard econometric methods: a and A are the (possibly shrunk) coefficients of a regression of the log return $\mathfrak{s}_i(t) = ln(S(t))$, on the risk factor $X(t)$, while Σ is computed using a discretization of the quadratic variation (for more details see next chapter).

As in David and Lleo (2020) (Assumptions 3.2.), we make the assumption that no two securities have the same risk profile:

Assumption 2.1. The matrix $\Sigma\Sigma'$ is positive definite.

This model for assets and factors dates back at least to Merton's ICAPM (Merton, 1973) and is a continuous time equivalent of Sharpe's multifactor model. As we said in the previous chapter, Merton's model has become a cornerstone of the risk-sensitive investment management literature (see Bielecki and Pliska, 1999; Kuroda and Nagai, 2002; and Davis and Lleo, 2014).

The model has three main advantages. First, it is straightforward enough to generate closed-form solutions. Second, it incorporates risk factors that create a dependence structure at the level of the assets' expected returns, in addition to the correlation structure already present in the assets' volatility. Finally, we can use standard statistical and econometric techniques from linear regression and time-series analysis to estimate the parameters of the model.

Importantly, the risk factors are not directly observable. As a result, we must rely on observable data such as market data and expert opinions to estimate the current level of the factors.

2.1.2. Expert Views

In this subsection, we introduce the dynamic of expert opinions and address the impact of behavioral biases to obtain unbiased views to use in the filtering process.

As anticipated in the previous chapter, Black and Litterman were not the first to propose using expert opinion in quantitative decision models. Indeed, the question of expert opinion arises naturally in the context of personal probabilities and personal utility (Savage,1971b). Regarding expert opinions, two issues come up: how to find out an expert's true opinions (elicitation), and whether expert forecasts are accurate.

In our method, experts provide their opinions on how factors will evolve over time. As explained before, this forecast contains three pieces of information: specific risk factor(s); a central view; and a confidence interval. As in David and Lleo (2020) (Assumption 3.3.), we make the assumption that:

Assumption 2.2. Experts formulate their forecasts, comprised of a central view and a confidence interval, on the evolution of the risk factors at t = 0.

The expert forecast process is constructed, calibrated, and debiased using all available information about the risk factor(s) and the confidence interval.

To be able to use the *k* expert forecasts as an observation in the filtering step, we model their dynamics, as in Davis and Lleo (2020), via a $(\mathcal{F}_t)$-adapted vector process $Z(t)$ which offers a noisy estimate of a functional transformation of the state variable $X(t)$:

$$Z(t) = Z(0) + \int_0^t f\big(s, X(s)\big)ds \; + \; \varepsilon(t) \tag{2.4}$$

The functional link between the risk factors and the expert's forecasts is modeled by the sensor function f. The forecast noise process $\varepsilon(t)$ introduces an artificial noise in the Kalman filter to model the uncertainty around the experts' central views.

We choose $f\big(\mathrm{t}, \mathrm{X(t)}\big) \coloneqq a_Z(t) + A_Z(t)X(t)$ and $\varepsilon(\mathrm{t}) \coloneqq \int_0^t \Psi_Z(\mathrm{s})dW(s)$, so Z(t) is an affine process solving the SDE:

$$dZ(t) = \big(a_Z(t) + A_Z(t)X(t)\big)dt + \; \Psi_Z(t)dW(t), \quad Z(0) = z. \tag{2.5}$$

The time-dependent coefficients $a_Z : [0, T] \to \mathbb{R}^k$, $A_Z : [0, T] \to \mathbb{R}^{kxn}$, and $\Psi_Z : [0, T] \to \mathbb{R}^{kxd}$ are bounded, C^1, and Lipschitz continuous.

We calibrate the functions $a_Z(t)$ and $A_Z(t)$ to the risk factor(s) or relation between risk factors that the experts are forecasting, and the matrix Ψ_Z to the confidence interval. Three pieces of information are embedded in expert forecasts: a risk factor or relation between risk factors, a central view, and a confidence interval. For a given $0 < \alpha < 1$, the expert's confidence level is given as a $(1 - \alpha)$ level. According to David and Lleo (2020), when we assume that the forecasting error between the expert's central view and the realized value of the factor(s) is normally distributed, we traditionally ask for a two-tailed confidence level. This model is flexible and can be tailored to fit various types of forecasts, including single factors, spreads between factors, and real-time nowcasting models as shown by Davis and Lleo (2020).

The final step before filtering is debiasing. As already described in the previous chapter, this process corrects expert forecasts by removing the effect of behavioral biases such as overconfidence, excessive optimism, conservatism, confirmation bias, and groupthink. We address overconfidence by increasing the magnitude of the diffusion term $\Psi_Z(t)$ to widen the confidence interval. We address excessive optimism by widening the confidence interval to reflect the possibility that the actual realization may differ significantly from the prediction. Conservatism is solved by the fact that our model does not require experts to update their forecasts; once formulated at the initial stage, expert forecasts are processed by the Kalman filter in a Bayesian manner, gradually lowering the influence of conservatism. We address confirmation bias through the Kalman filter that weights forecasts based on their accuracy, as measured by the magnitude of $\Psi_Z(t)$. Finally, to reduce the impact of groupthink, we bring in dissenting experts whose forecasts differ markedly from the majority, to ensure heterogeneity of the expert pool. We also add a stress test scenario to broaden the range of forecasts.

According to Davis and Lleo (2020), once we have debiased, possibly augmenting the forecast vector with dissenting expert forecasts and stress test scenarios, we have a K-dimensional forecast vector modeled by the SDE:

$$dZ(t) = \big(a_Z(t) + A_Z(t)X(t)\big)dt + \Psi_Z(t)dW(t), \quad Z(0) = z. \tag{2.6}$$

where W(t) is now a $\mathbb{R}^{n+m+K}$-valued $(\mathcal{F}_t)$-Brownian motion on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0}^{T}, \mathbb{P}), K \geq k$.

We assume that the time-dependent coefficients $a_Z(t)$, $A_Z(t)$ and $\Psi_Z(t)$ have already been debiased following the process proposed by Davis and Lleo (2020).

The next section introduces the Kalman filter and explains how filtering techniques are used to combine market data with expert views to estimate the unobservable risk factors.

2.2. Filtering to Estimate the Risk Factors

Once we have collected market data and expert opinions, we combine them to estimate the current values of the risk factors. In our dynamic setting, we apply linear filtering, an established estimation technique, to perform the aggregation and to combine data with views.

In this section, we introduce the Kalman filter and its application to solve the discrete data linear filtering problem. We then apply the filter to our model as an instrument to estimate the unobservable risk factors by combining market data with expert opinions.

The filtering problem deals with the estimation of a stochastic process $X(t)$ that is unobserved based on the past and current measurement of a related process Y, $\{Y(s): 0 \leq s \leq$

$t\}$. The information coming out of the measurement process up to time t is conveniently represented by the sigma-algebra $\mathcal{F}$, generated by $\{Y(s): 0 \leq s \leq t\}$.

The Kalman filter and its generalizations have been the main tools for estimating the unobserved variables from the observed ones in econometrics, engineering and finance for a long time. The filter can be defined as a conditional moment estimator for linear Gaussian systems. It is used in the calibration of time series models, forecasting of variables, and data smoothing applications. Despite the length and degree of noise in the input data, the Kalman Filter is quick and simple to implement. The idea is to proceed in two steps: first considering all the data available up to that time step, we estimate the hidden state or previous distribution. Then we derive a conditional estimation of the posterior distribution of the state using the predicted value together with the new observation.

A Kalman filter is a specific application of a Bayes filter in which the dynamics and sensory model are linear Gaussian. A Bayes filter is used to estimate the probability density function of states over time using observations, utilizing Bayes' theorem which gives a mathematical formula for conditional probability. Bayes' theorem expresses the probability of an event based on prior knowledge of conditions related to the event.

An excellent introduction to the filtering problem is given by Davis and Marcus (1981). Besides, see Lipster and Shiryayev (1978) for a more mathematical and complete presentation. On the other hand, Jazwnski (1970) gives a more applied and practical analysis of the filtering issues.

The filtering technique has long been used in signal processing and control engineering. In economics and finance, filtering theory has developed considerably since the seminal work of Kalman (1960) and Kalman and Bucy (1961). Since then, filtering techniques found applications in stochastic optimization (see for example Davis, 1977; Bucy and Joseph, 1987; or Bensoussan, 2004), and in finance where Brennan (1998), Xia (2001), Nagai and Peng (2002), and Davis and Lleo (2011, 2013) used linear filtering to estimate the parameters of their portfolio selection models.

2.2.1. Kalman Filter

Kalman presented a recursive solution to the discrete data linear filtering problem for the first time in 1960. We initially introduce the discrete time version of the Kalman filter to give a better understanding of how the filter works and to prepare the reader for the subsequent presentation of the continuous time version.

In this section, we describe the linear Gaussian filter and use a general specification of the model based on Welch and Bishop (2002) to have an overview of the filter and its application. See Harvey (1989) or Welch and Bishop (2002) for a detailed introduction.

Given a dynamic process $x(t)$ following a transition equation

$$x(t) = f(x(t-1), w(t)) \tag{2.7}$$

we suppose we have a measurement $z(t)$ such that

$$z(t) = h(x(t), u(t)) \tag{2.8}$$

where $w(t)$ and $u(t)$ are two mutually uncorrelated sequences of temporally-uncorrelated Normal random variables with zero means and covariance matrices $Q(t), R(t)$ respectively. Moreover, $w(t)$ is uncorrelated with $x(t-1)$ and $u(t)$ is uncorrelated with $x(t)$.

We denote the dimension of $x(t)$ as $n(x)$, the dimension of $w(t)$ as $n(w)$ and so on.

The *a priori* process estimate can be defined as

$$\hat{x}(t)^{-} = E[x(t)] \tag{2.9}$$

which is the estimation at time step $t = 0$ prior to the step t measurement. Similarly, the *a posteriori* estimate can be defined as

$$\hat{x}(t) = E[x(t)|z(t)] \tag{2.10}$$

which is the estimation at time step t after the measurement.

We also have the corresponding estimation errors $e(t)^{-} = x(t) - \hat{x}(t)^{-}$ and $e(t) = x(t) - \hat{x}(t)$ and the estimate error covariances

$$\begin{aligned} P(t)^{-} &= E\left[e(t)^{-}e(t)^{-^{T}}\right] \\ P(t) &= E[e(t)e(t)^{T}] \end{aligned} \tag{2.11}$$

where the superscript *T* corresponds to the transpose operator.

In deriving the equations for the Kalman filter, Welch and Bishop (2002) begin with the goal of finding an equation that computes an a posteriori state estimate $\hat{x}(t)$ as a linear combination of an a priori estimate $\hat{x}(t)^-$ and a weighted difference between an actual measurement $z(t)$ and a measurement prediction $H\hat{x}(t)^-$ as shown below in (2.12). Some justification for (2.12) is given in Welch and Bishop (2002).

$$\hat{x}(t) = \hat{x}(t)^- + K(z(t) - H\hat{x}(t)^-) \tag{2.12}$$

The difference $(z(t) - H\hat{x}(t)^-)$ in (2.12) is called the measurement innovation, or the residual. The residual reflects the discrepancy between the predicted measurement $H\hat{x}(t)^-$ and the actual measurement $z(t)$. The two are in complete agreement if the residual is zero.

The $n \times m$ matrix K in (2.12) is chosen to be the gain or blending factor that minimizes the a posteriori error covariance (2.11). This minimization can be accomplished by first substituting (2.12) into the above definition for $e(t)$, substituting that into (2.11), taking the indicated expectations, taking the derivative with respect to K, setting that result equal to zero, and then solving for K (for more details see Maybeck, 1979; Brown, 1992; or Jacobs, 1993). One form of the resulting K that minimizes (2.11) is given by[3]

$$\begin{aligned} K(t) &= P(t)^- H^T (HP(t)^- H^T + R)^{-1} \\ &= \frac{P(t)^- H^T}{HP(t)^- H^T + R} \end{aligned} \tag{2.13}$$

Looking at (2.13), as the measurement error covariance R approaches zero, the gain K weights the residual more heavily. Specifically,

$$\lim_{R(t) \to 0} K(t) = H^{-1}.$$

On the contrary, as the a priori estimate error covariance $P(t)^-$ approaches zero, the gain K weights the residual less heavily. Specifically,

[3] All of the Kalman filter equations can be algebraically manipulated into to several forms. Equation (2.13) represents the Kalman gain in one popular form.

$$\lim_{P(t)^- \to 0} K(t) = 0$$

Another perspective on the weighting by K is that as the measurement error covariance R approaches zero, the actual measurement $z(t)$ is given greater credibility, while the predicted measurement $H\hat{x}(t)^-$ is given less credibility. On the other hand, as the a priori estimate error covariance $P(t)^-$ approaches zero the actual measurement $z(t)$ is given less credibility, while the predicted measurement $H\hat{x}(t)^-$ is given greater credibility.

2.2.2. Risk Factors Estimation

The previous sections introduced the discrete Kalman filter. Now, we apply this linear filtering technique to our model to estimate the unobservable risk factors.

In our approach, market data and expert views are treated as a single set of observations that the filter will process to determine the factors' values. The Kalman-Bucy filter is a continuous time equivalent of the discrete-time Kalman Filter. Similar to the Kalman Filter, the Kalman-Bucy filter is used to estimate unmeasured states of a process typically with the intention of controlling one or more of those states. The Kalman–Bucy linear filter is simple, robust to misspecification of the noise term, and closed-form solvable. Another benefit of the Kalman–Bucy linear filter is that it naturally provides more weight to more reliable observations such as lower confidence interval or lower volatility, and less weight to less reliable ones. Hence, the opinion of an asset manager with an impeccable track record will have more weight than that of a rookie analyst with a checkered record.

Computationally, since the factor process is not observable, we estimate its current value $X(t)$ from an observation process $Y(t)$ using a filter. In this paper, the observations consist of data on the m risky financial securities and the k expert forecasts.

We start by combining the securities' excess returns and expert forecasts into a single $(m+k)-$ dimensional observation vector Y(t) = ($\mathfrak{s}$(t), Z(t))′. This observation process is linear in the state with Gaussian noise, and solves the SDE:

$$dY(t) = \big(a_Y(t) + A_Y(t)X(t)\big)dt + \Gamma(t)dW(t), \; Y(0) = y_0 \tag{2.14}$$

where

$$a_Y(t) = \begin{pmatrix} a(t) - \frac{1}{2}\Sigma\Sigma' \\ a_Z(t) \end{pmatrix}, \quad A_Y(t) = \begin{pmatrix} A(t) \\ A_Z(t) \end{pmatrix}, \quad \Gamma(t) = \begin{pmatrix} \Sigma \\ \Psi_Z(t) \end{pmatrix}.$$

Now, let $\mathcal{F}_t^Y = \sigma\{Y(u), 0 \leq u \leq t\}$ be the filtration generated by the observation process only. Since the risk factor process $X(t)$ and the observation process $Y(t)$ take the form of the 'signal' and 'observation' processes in a Kalman filter system, the conditional distribution of the factor process $X(t)$ is normal $N\left(\hat{X}(t), P(t)\right)$ where $\hat{X}(t) = \mathbb{E}[X(t)|\ \mathcal{F}_t^Y]$ satisfies the Kalman filter equation and $P(t)$ is a deterministic matrix-valued function. The Kalman filter replaces the initial state process $X(t)$ with an estimate $\hat{X}$(t).

Next, define the processes $Y^1(t)$, $Y^2(t) \in \mathbb{R}^m$ as the solution to:

$$dY^1(t) = A_Y(t)X(t)dt + \Gamma(t)dW(t), \qquad Y^1(0) = 0, \tag{2.15}$$

$$dY^2(t) = a_Y(t) \cdot dt, \quad Y^2(0) = y_0 \tag{2.16}$$

then $Y(t) = Y^1(t) + Y^2(t)$. Therefore, (2.1) and (2.14) constitute a Kalman filtering system, and we can apply the following theorem to solve the filtering problem:

Theorem 2.1 Kalman Filter (Davis, 1979; Davis & Lleo, 2011, 2020).

1. *The Kalman estimate $\hat{X}(t)$ is the unique solution of the SDE:*

$$d\hat{X}(t) = \left(b(t) + B(t)\hat{X}(t)\right)dt + \hat{\Lambda}(t)dU(t),\ \hat{X}(0) = \mu_0, \tag{2.17}$$

   where $\widehat{\Lambda}(t) = (\Lambda(t)\Gamma(t)' + P(t)A_Y')(\Gamma(t)\Gamma(t)')^{-\frac{1}{2}}$.

2. *The variance $P(t)$ is the unique non-negative definite symmetric solution of the matrix Riccati*

$$\begin{aligned}\dot{P}(\mathrm{t}) = {} & \Lambda(t)Y^{\perp}(t)\Lambda'(t) - P(t)A_Y'(t)(\Gamma(t)\Gamma(t)')^{-1}A_Y(t)P(t) \\ & + (B(t) - \Lambda(\mathrm{t})\Gamma(t)'(\Gamma(t)\Gamma(t)')^{-1}A_Y(t))P(t) \\ & + P(t)(B(t)' - A_Y'(t)(\Gamma(t)\Gamma(t)')^{-1}\Gamma(t)\Lambda'(t)), \quad P(0) = P_0 \end{aligned} \tag{2.18}$$

   with $Y^{\perp}(t) := I - \Gamma(t)'(\Gamma(t)\Gamma(t)')^{-1}\Gamma(t)$.

The matrix $P(t)$ represents the conditional variance-covariance matrix with respect to all available observations up to time t. Thus, the filter distribution is a normal distribution having as its mean the filtered expected value (the filtered estimate of the unobserved factor) and as its variance-covariance matrix precisely $P(t)$.

3. *The innovation process* $U(t)$ *defined by*

$$dU(t) = (\Gamma(t)\Gamma(t)')^{-\frac{1}{2}}\big(dY^1(t) - A_Y(t)\hat{X}(t)dt\big), \quad U(0) = 0 \tag{2.19}$$

is a R^{m+k}*- valued*$(\mathcal{F}_t^Y)$*-Brownian motion on* $(\Omega, \mathcal{F}, \mathcal{F}_t^Y)_{t=0}^T, \mathbb{P})$.

The Kalman filter has replaced our initial stage process $X(t)$ by an estimate $\hat{X}(t)$ with dynamics given in equation (2.17). To recover the components of the observation process, we use (2.14) together with (2.19) to express the dynamics of $Y(t)$ as:

$$dY(t) = dY_1(t) + dY_2(t) = \left(a_Y(t) + A_Y(t)\hat{X}(t)\right)dt + (\Gamma(t)\Gamma(t)')^{\frac{1}{2}}dU(t), \quad Y(0) = y_0 \tag{2.20}$$

Then, we decompose the $(m+k)\times(m+k)$ matrix $(\Gamma\Gamma(t)')^{\frac{1}{2}}$ as $(\Gamma\Gamma(t)')^{\frac{1}{2}} := \left(\hat{\Sigma}'(t)\hat{\Psi}'_Z(t)\right)'$, where $\hat{\Sigma}'(t)$ and $\hat{\Psi}'_Z(t)$ are respectively a $m\times(m+k)$ matrix and a $k\times(m+k)$ matrix such that $\hat{\Sigma}\hat{\Sigma}' = \Sigma\Sigma'(t)$ and $\hat{\Psi}_Z\hat{\Psi}_Z(t)' = \Psi_Z\Psi_Z(t)'$.

Finally, $s(t), Z(t),$ and $S(t)$ respectively solve the following SDEs (for more details see Davis and Lleo 2020):

$$d\mathfrak{s}(t) = \left[\left(a(t) - \tfrac{1}{2}\,d_\Sigma\right) + A(t)\hat{X}(t)\right]dt + \hat{\Sigma}dU(t),$$

$$dZ(t) = \left(a_Z(t) + A_Z(t)\hat{X}(t)\right)dt + \hat{\Psi}_Z(\mathrm{t})dU(t), \quad Z(0) = z.$$

$$\frac{dS_i(t)}{S_i(t)} = (a(t) + A(t)\hat{X}(t))_i dt + \sum_{j=1}^{m+k}\Sigma_{ij}dU_j(t), \qquad S_i(0) = s_i(0) \tag{2.21}$$

Once the risk factors have been estimated, we have all the information required to solve the risk-sensitive stochastic control problem and to derive the optimal asset allocation.

2.3. The Risk-Sensitive Stochastic Control Problem

In this section, we express the investor's portfolio selection problem as a risk-sensitive stochastic control problem based on the Kalman filter estimate $\hat{X}(t)$. We derive the optimal asset allocation and show that the value function is the unique $C^{1,2}$ solution to the Hamilton-Jacobi-Bellman partial differential equation.

Investors can monitor any security traded on the financial markets, but many times they are unable to invest in certain securities due to investment restrictions. For instance, it is typically disallowed for U.S. large-cap equity managers to trade small-cap stocks or government bonds. Therefore, we split the financial universe into $m_1 > 0$ securities that investors are authorized to hold in their portfolio and $m_2 = m - m_1 > 0$ securities that investors can only observe.
When $m_2 = 0$ then $m_1 = m$ and investors can trade all the securities. In general, $m_2 > 0$ and we decompose the securities price vector $S(t)$ and associated parameter vectors and matrices a, A, Σ as:

$$S(t) = \begin{pmatrix} S_1(t) \\ S_2(t) \end{pmatrix}, \qquad a(t) = \begin{pmatrix} a_1(t) \\ a_2(t) \end{pmatrix}, \qquad A(t) = \begin{pmatrix} A_1(t) \\ A_2(t) \end{pmatrix}, \qquad \Sigma(t) = \begin{pmatrix} \Sigma_1(t) \\ \Sigma_2(t) \end{pmatrix}$$

Here, as in Davis and Lleo (2020), $S_1(t)$ is the m_1 −dimensional process of investable securities prices and $S_2(t)$ is the m_2 −dimensional process of non-investable, but observable, securities prices. We adopt a similar notation for the parameters a, A, Σ.

Next, we introduce the investor's strategy $h(t)$ as a m_1-element vector asset allocation process representing the proportion of wealth invested in the financial securities. The risk-sensitive stochastic control problem and its solution are well-known in the literature (see Kuroda and Nagai, 2002; or Davis and Lleo, 2014 for details). The investor's strategy is defined mathematically in *definition 2.3.1* below.

The discounted wealth process $V(t)$ is the market value of the investment portfolio corresponding to the investment strategy $h(t)$. It solves the SDE:

$$\frac{dV(t)}{V(t)} = h'(t)\big(a_1(t) + A_1(t)X(t)\big)dt + h'(t)\Sigma_t(t)dW(t), \qquad V(0) = v \tag{2.22}$$

When the factors are observable, Davis and Lleo (2013) express the investors' objective as the maximization of the risk-sensitive asset management criterion, as explained in the previous chapter and equation (20) in their paper:

$$J(t,x,h,T,\theta) = -\frac{1}{\theta} ln\mathbb{E}_{t,x}\left[e^{-\theta lnV(t)}\right] = -\frac{1}{\theta} ln\mathbb{E}_{t,x}\left[V^{-\theta}(T)\right] \tag{2.23}$$

where $\theta \in (-1, 0) \cup (0, \infty)$ is the risk sensitivity parameter and $T < \infty$ is a fixed time horizon. $E_{t,x}[\cdot]$ denotes the expectation with initial conditions (t, x).

Maximizing the risk-sensitive asset management criterion is tantamount to selecting the asset allocation h that maximizes the risk-adjusted excess log return of the asset portfolio. Davis and Lleo (2020) show this by performing a Taylor expansion of the risk-sensitive criterion J around $\theta = 0$:

$$J(t,x,h,T,\theta) = \mathbb{E}_{t,x}[lnV(t)] - \frac{\theta}{2} var_{t,x}[lnV(T)] + O(\theta^2)$$

Thus, the risk-sensitive asset management criterion is akin to a dynamic "mean-variance" optimization on the log excess return of the portfolio over the money market instrument. Alternately, the criterion is comparable to maximizing the Hyperbolic Absolute Risk Aversion (HARA) utility. The expectation $\mathbb{E}\left[e^{-\theta lnV(t)}\right] =: \mathbb{E}\left[U\left(V(t)\right)\right]$ defines the expected utility of wealth at time t, subject to a HARA utility function.

In our model, investors cannot observe the factor value $X(t)$. So, they use a modified risk-sensitive criterion based on the Kalman filter estimate $\hat{X}(t)$.

By Itô's lemma,

$$e^{-\theta lnV(t)} = v^{-\theta} exp\,\{\theta \int_0^t g\left(\hat{X}_s, h(s); \theta\right) ds\} X_t^h\,, \tag{2.24}$$

where

$$g(x,h;\theta) = \tfrac{1}{2}(1+\theta) h'\hat{\Sigma}_1\hat{\Sigma}_1' h - h'\left(\hat{a}_1 + \hat{A}_1 x\right) \tag{2.25}$$

and the exponential martingale $\mathcal{X}_T^h$ is

$$\mathcal{X}_T^h \coloneqq \exp\,\{-\theta \int_0^T h'(t)\hat{\Sigma}_1(t) dU(t) - \frac{1}{2}\theta^2 \int_0^T h'(t)\hat{\Sigma}_1(t)\,\hat{\Sigma}_1'(t) h(t) dt\} \tag{2.26}$$

We also assume that the investor's strategy $h(t)$ is in class $\mathcal{A}(T)$ defined below.

Definition 2.1. (Definition 5.1. Davis and Lleo, 2020) An $\mathbb{R}^{m_1}$ -valued control process $h(t)$ is in class $\mathcal{A}(t)$ if the following conditions are satisfied:

i. $h(t)$ is progressively measurable with respect to $\{ \mathcal{F}_t^Y \}_{t\geq 0}$ and is càdlàg[4];

ii. $P(\int_0^t |h(s)|^2 ds < +\infty) = 1$;

iii. The Doléans exponential[5] $\mathcal{X}_T^h$ given at (2.26) is an exponential martingale, thus $\mathbb{E}[\mathcal{X}_T^h] = 1$.

Remark. The control process $h(t)$ is adapted to the filtration $\mathcal{F}_t^Y = \sigma\{ s_i(u), Z_j(u), 0 \leq u \leq t, i = 0, \ldots, m, j = 1, \ldots, k\}$ generated by the observations.

We solve the stochastic control problem by expanding the change of measure argument proposed by Davis and Lleo (2008) and take into consideration the case $\theta > 0$ that leads to a maximization over a concave function.

Let $\mathbb{P}_h$ be the measure on $(\Omega, \mathcal{F}_T)$ defined via the Radon-Nikodym derivative[6] $\frac{d\mathbb{P}_h}{d\mathbb{P}} := \mathcal{X}_T^h$. For $h(t) \in \mathcal{A}(T)$, $U^h(t) := U(t) + \theta \int_0^t \hat{\Sigma}_1' (s) h(s) ds$ is a standard Brownian motion under the measure $\mathbb{P}_h$. As in Proposition (5.6) in Davis and Lleo (2020), the control criterion, under this new measure, is

$$I(t, x, h; T, \theta) = -\frac{1}{\theta} ln E_{t,x}^h [exp \{\theta \int_t^T g(\hat{X}_s, h(s); \theta) ds\}] \tag{2.27}$$

where $E_{t,x}^h[\cdot]$ denotes the expectation taken with respect to the measure $\mathbb{P}_h$ and represents also the expectation conditionally on time t and on the value at time t of the filtered factor process. The dynamics of the state variable $\hat{X}(t)$ under the new measure are obtained from (2.17) and (2.26)

[4] A cadlag function is a function, defined on $\mathbb{R}$ or a subset of $\mathbb{R}$, that is right continuous and has a left limit. The acronym *cadlag* comes from the French "continue à droite, limite à gauche," which translates to the English "right-continuous with left limits" (sometimes abbreviated "RCLL"). All continuous functions are "cadlag."

[5] Let X_t be a measurable process adapted to the filtration. Doleans Dade exponential is the solution to: $dY_t = Y_t dX_t$ which is the form of density process $dQ = \eta_t dP$.

[6] Let X be a measurable space and let μ and ν be measures on X, valued in the real numbers or in the complex numbers. Let f be a measurable function f (with real or complex values) on X. The function f is a Radon–Nikodym derivative of μ with respect to ν if, given any measurable subset A of X, the μ-measure of A equals the integral of f on A with respect to ν: $\mu(A) = \int_A f\nu = \int_{x\in A} f(x) d\nu(x)$.

$$d\hat{X}(t) = \left(b(t) + B\hat{X}(t) - \theta\hat{\Lambda}(t)\hat{\Sigma}_1' h(t)\right) dt + \hat{\Lambda}(t) dU^h(t),\ \ t \in [0, T] \tag{2.28}$$

The value function Φ for the auxiliary criterion $I(t, \hat{x}; \text{h}; T, \theta)$ is defined as

$$\Phi(t, \hat{x}) \ := \ sup_{h \in \mathcal{A}(t)} I(t, x;\ h;\ T, \theta) \tag{2. 29}$$

The Hamilton-Jacobi-Bellman partial differential equation linked to the control problem is

$$\frac{\partial \Phi}{\partial t}(t, x) + sup L_t^h(t, x, D\Phi, D^2\Phi) = 0 \tag{2.30}$$

where $D\Phi = \left(\frac{\partial \Phi}{\partial X_1}, \dots, \frac{\partial \Phi}{\partial X_i}, \dots, \frac{\partial \Phi}{\partial X_n}\right)'$, $\quad D^2\Phi = \left[\frac{\partial^2 \Phi}{\partial x_i \partial x_j}\right], i, j = 1, \dots, n,$

and

$$L_t^h(t, x, p, M) = \left(b(t) + B(t)x - \theta\hat{\Lambda}(t)\hat{\Sigma}_1'(t)h\right)' p + \frac{1}{2} tr\left(\hat{\Lambda}\hat{\Lambda}(t)' p\right)$$
$$-\frac{\theta}{2} p' \hat{\Lambda}\hat{\Lambda}(t)' p - g(t; x, h; \theta) \tag{2.31}$$

for p $\in \mathbb{R}^n$ and subject to terminal condition $\Phi(T, \hat{x}) = 0$.

The term inside the sup is quadratic in h. Its unique maximizer refers to the candidate's optimal control

$$\hat{h}(t, \hat{x}, p) = \frac{1}{1+\theta}\left(\hat{\Sigma}_1\hat{\Sigma}_1'(t)\right)^{-1}\left[\hat{a}_1(t) + A_1(t)\hat{x} - \theta\hat{\Sigma}_1(t)\hat{\Lambda}(t)' p\right)] \tag{2.32}$$

where $(t, \hat{x}, p)$ stands in for $(t, \hat{X}(t), D\Phi(\text{t}, \hat{X}(t)))$. Moreover, the value function $\Phi(t, \hat{x})$ is $\Phi(t, \hat{x}) = \frac{1}{2}\hat{x}' Q(t)\hat{x} + \hat{x}' q(t) + k(t)$, where $Q(t)$ is the unique symmetric non-negative solution to the matrix Riccati equation, $q(t)$ solves a linear ODE, and $k(t)$ is found by integration. Specifically, $Q(t)$ solves

$$\dot{Q}(t) - Q(t)K_0(t)Q(t) + K_1(t)Q(t) + Q(t)K_1(t) + \frac{1}{\theta+1} A_1(t)(\hat{\Sigma}_1\hat{\Sigma}_1(t))^{-1} A_1(t) = 0, \tag{2.33}$$

where

$$K_0(t) = \theta[\widehat{\Lambda}(t)\left(I - \frac{\theta}{\theta+1}\hat{\Sigma}_1(t)\left(\hat{\Sigma}_1\hat{\Sigma}_1(t)\right)^{-1}\hat{\Sigma}_1(t)\right)\widehat{\Lambda}(t)\Big],$$

$$K_1(t) = B(t) - \frac{\theta}{\theta+1}\widehat{\Lambda}(t)\hat{\Sigma}_1(t)(\hat{\Sigma}_1\hat{\Sigma}_1(t))^{-1}A_1(t),$$

and I is the $n \times n$ identity matrix. The vector-valued function $q(t)$ solves

$$\dot{q}(t) + \big(K_1(t) - Q(t)K_0(t)\big)q(t) + Q(t)\left(b + \theta\widehat{\Lambda}(t)\right)$$
$$+\frac{1}{\theta+1}(A_1(t) - \theta Q(t)\widehat{\Lambda}(t)\hat{\Sigma}_1(t))(\,\hat{\Sigma}_1\hat{\Sigma}_1(t))^{-1}\left(a_1 + \theta\hat{\Sigma}_1(t)\right) - C(t) = 0$$

and $k(t) = \int_s^T \ell(s)ds$, where

$$\ell(s) = \frac{1}{2}tr\left(\widehat{\Lambda}\widehat{\Lambda}(t)Q(t)\right) - \frac{\theta}{2}q(t)\widehat{\Lambda}\widehat{\Lambda}(t)q(t) + b(t)q(t)$$

$$+\frac{1}{2}\frac{1}{\theta+1}a_1(t)(\,\hat{\Sigma}_1\hat{\Sigma}_1(t))^{-1}a_1(t) + \frac{1}{2}\frac{\theta^2}{\theta+1}q(t)\widehat{\Lambda}(t)\hat{\Sigma}_1(t)(\,\hat{\Sigma}_1\hat{\Sigma}_1(t))^{-1}\hat{\Sigma}_1(t)\widehat{\Lambda}(t)q(t)$$

$$-\frac{\theta}{\theta+1}q(t)\,\widehat{\Lambda}(t)\hat{\Sigma}(t)(\,\hat{\Sigma}_1\hat{\Sigma}_1(t))^{-1}a_1 - \frac{\theta^2}{\theta+1}q(t)\widehat{\Lambda}(t)\hat{\Sigma}_1(t)(\,\hat{\Sigma}_1\hat{\Sigma}_1(t))^{-1}\hat{\Sigma}_1(t)$$

$$+\,\theta\widehat{\Lambda}(t)q(t) - \frac{1}{2}(\theta-1) + \frac{\theta}{\theta+1}a_1(t)\,(\,\hat{\Sigma}_1\hat{\Sigma}_1(t))^{-1}\hat{\Sigma}_1(t)$$

$$+\frac{1}{2}\frac{\theta^2}{\theta+1}\hat{\Sigma}_1(t)(\,\hat{\Sigma}_1\hat{\Sigma}_1(t))^{-1}\hat{\Sigma}_1(t).$$

The resolution of the stochastic control problem is completed by a standard verification argument that shows that the optimal investment strategy h^* is:

$$h^*\left(t,\hat{X}(t)\right) = \frac{1}{\theta+1}(\,\hat{\Sigma}_1\hat{\Sigma}_1'(t))^{-1}[a_1(t) + A_1(t)\hat{X}(t)$$
$$-\,\theta\hat{\Sigma}_1\hat{\Lambda}(t)'(q(t) + Q(t)\hat{X}(t))]. \quad (2.34)$$

2.3.1. Implications of the Optimal Investment Strategy

The previous section culminated with the derivation of a solution to the risk-sensitive stochastic problem. This section explores the implications of this solution starting with a detailed study of the optimal investment strategy given at (2.34).

The following Proposition (see Proposition 5.2 in Davis and Lleo, 2020) establishes that the optimal investment strategy is a fractional Kelly strategy with allocations to two constituent portfolios: the Kelly or log-optimal portfolio, and an intertemporal hedging portfolio.

Proposition 2.3.1. (Proposition 5.2 Davis and Lleo, 2020) (Fractional Kelly Strategy (PFKS)). The optimal investment strategy $h^*\left(t, \hat{X}(t)\right)$ consists of an allocation between two funds: $h^K$and h^{PIHP}.

(i) The fund h^K is a personal Kelly portfolio with a factor-dependent allocation

$$h^K(t, \hat{X}(t)) = (\hat{\Sigma}_1 \hat{\Sigma}_1{}'(t))^{-1}(a_1(t) + A_1(t)\hat{X}(t)) \tag{2.35}$$

(ii) The fund h^{PIHP} is a Personal Intertemporal Hedging Portfolio (PIHP) with factor-dependent allocation

$$h^{PIHP}(t, \hat{X}(t)) = (\hat{\Sigma}_1 \hat{\Sigma}_1'(t))^{-1}\hat{\Sigma}_1(t)\hat{\Lambda}(t)'(q(t) + Q(t)\hat{X}(t)) \tag{2.36}$$

Moreover, the relative allocation of each fund is constant at $f := \frac{1}{\theta+1}$ for h^K and f-1 for h^{PIHP}.

The Personal Intertemporal Hedging Portfolio (PIHP) portfolio is a portfolio that takes positions in financial securities to protect the investors' optimal utility of future consumption from variations in the risk factors (Merton, 1973). Due to the similarity between this portfolio and risk management hedging procedures, it is called "intertemporal hedging portfolio". The PIHP has a small impact on the optimal investment strategy and is heavily dependent on both factors and expert forecasts. Financial securities noise is generally larger than factor noise and the correlations between factors and securities are not perfect, so the slope term $(\hat{\Sigma}_1\hat{\Sigma}_1'(t))^{-1}\Sigma_1(t)\hat{\Lambda}(t)$ is usually small. Moreover, $\hat{X}(t)$, $q(t)$, and $Q(t)$ tend to be small as well.

We observe from Proposition 2.3.1 that fractional Kelly strategies extend and replace Mutual Fund Theorems when the constituent portfolios are not universal funds, meaning that they are not identical for every investor. In this situation, the PIHP and Kelly portfolio are not universal

(Davis & Lleo, 2020). The PIHP is not universal because its asset allocation depends on both subjective expert forecasts, via the estimated $\hat{X}(t)$, and risk sensitivity θ, via the functions $q(t)$ and $Q(t)$. Thus, two investors with identical preferences, investment universe and dataset, but different risk-sensitivities, will construct two different PIHPs. Similarly, the Kelly portfolio is not universal because it is a function of expert forecasts via $\hat{X}(t)$. Therefore, two investors who have access to different experts, but the same factors and securities may identify different Kelly portfolios. However, the Kelly portfolio is not dependent from the risk aversion coefficient and can be then defined as a "more" universal portfolio compared to the PIHP.

These results also contribute to the understanding of why investors with a Kelly-like track record, such as Buffett, Gross, Thorp, or Keynes, maintain different portfolios, as stated in Davis and Lleo (2020). Although their risk tolerance is the same, their investment universe, investment horizon, and personal views are not.

2.4. Extensions to the Model

We may identify several extensions to this basic setup. The most significant modification is the use of Lévy processes to complement Gaussian distributions. Lévy processes offer a more accurate description of the distribution of asset returns and new techniques for determining the confidence interval of expert opinions. In addition, these processes address new biases, such as the following:

1. *Narrow framing* is the tendency to break down a complicated or multidimensional problem into smaller, simpler problems without considering how these smaller problems interact. A silo-centric[7] perspective of the world is frequently the result of narrow framing.
2. *Opaque framing* relates to the degree of transparency or opacity in the explanation of a task or choice.
3. *Extrapolation bias* prompts analysts to base their forecasts mostly on current patterns. This bias can be shown, for instance, in the "hot hand fallacy" and the "gambler's fallacy." In contrast to the "gambler's fallacy," which holds that a recent trend will reverse, the "hot hand fallacy" is characterized by the assumption that a recent trend will continue in the near future.
4. '*Affect heuristic'* relates to decisions based on intuition and feelings rather than on an analysis of the facts and circumstances.

[7] Silo mentality is when different teams or team members in the same company purposely don't share valuable information with other members of the company. This silo mindset hurts the unified vision of a business and deters long term goals from being accomplished.

Second, the stochastic control and filtering problems can actually be separated. We have already combined data from the financial market with expert opinions using this information. We can go a step further by integrating "big data" analytics, or confidence indices as observations in the filter, even if they are not included in the portfolio optimization

Finally, by considering how an investment benchmark or liabilities may affect the estimating process and the final portfolio allocation, we can broaden the variety of investment objectives and constraints (see Davis and Lleo, 2021).

In the next chapter, we will implement the model described in the previous sections to analyze the impact of behavioral biases on portfolio selection.

Chapter 3

Model Implementation

In this chapter, we aim to develop an implementation of the model we have described in the previous chapter. Following our discussion of the optimal investment strategy, we present a simulation that provides a simple computational example. This simulation has two objectives. The initial goal is to demonstrate the functioning of the model. The second, and more important objective, is to investigate the impact of behavioral biases on portfolio construction.

To study the effect of behavioral biases on the optimal asset allocation, we compare the performance of the Black-Litterman model in continuous-time against its behavioral version where the impact of behavioral biases on expert forecasts has been addressed.

First, we introduce the dataset and the parameters composing the models such as risk factors, financial securities, and expert opinions. Second, we describe how to obtain the parameters of the models to perform the simulation. We then present the differences between the two models to analyze the effect of behavioral biases on asset allocations. Finally, we monitor and analyze the models based on some portfolio performance metrics that we have already presented in the first chapter.

The Conclusion chapter summarizes the content presented in this thesis and underlines the results that we find by running the model implementation.

3.1. Dataset

In this section, we introduce the dataset to construct our model implementation underlying the main elements of the model.

We consider $n = 3$ risk factors. We select the three risk factors from the Fama & French model. The Fama-French Three-Factor Model, developed in 1992, improves on the Capital Asset Pricing Model (CAPM) by including size and value risk factors. This model takes into account the tendency of small-cap and value stocks to outperform the market, and by including these extra factors it aims to provide a more accurate evaluation of manager performance. The formula is:

$$R_{it} - R_{ft} = \alpha_{it} + \beta_1(R_{Mt} - R_{ft}) + \beta_2 SMB_t + \beta_3 HML_t + \epsilon_{it}$$

where $R_{it} - R_{ft}$ is the expected return on a stock or portfolio i at time t, $R_{Mt} - R_{ft}$ is the excess return on the market portfolio (index), SMB_t is size premium (small minus big), HML_t is the value premium (high minus low), $\beta_{1,2,3}$ are the factor coefficients.

In our model we select the following risk factors: X_1 is the market risk premium (MrktminRF); X_2 is the risk premium between stocks with a small and big market capitalization (SMB); X_3 captures the risk premium between stocks with a high and low book-to-market ratio (HML). We chose the Fama-French factor model since it is a simple multi-factor model and the data are easily accessible.

The investment universe consists of $m = 13$ U.S. Exchange Traded Funds (ETFs). The first 11 ETFs track the performance of S&P500 sectors, as per the latest Global Industry Classification Standard (GICS). Hence, the investor can replicate the S&P500 by holding these 11 ETFs according to the weights indicated in Table 1. Fund 12 is the iShares Core S&P 400 Mid-Cap ETF (ticker: IJH), which tracks 400 mid-cap stocks. Fund 13 is the iShares Core S&P 600 Small-Cap ETF (IJR), a small-cap index with 600 stocks. The holdings of the 13 ETFs do not have any overlap.

The Global Industry Classification Standard (GICS) was developed in 1999 by MSCI in collaboration with S&P Dow Jones Indices to provide an efficient, detailed and flexible tool for use in the investment process. It was created to address the worldwide financial industry's requirement for a globally accepted, accurate, and complete method of defining industries and classifying securities by industry. Its universal industry classification approach aims to enhance transparency and effectiveness in the investment process. The GICS has 11 sector classifications:

- *Energy Sector*: includes companies that are involved in the exploration and production, refining, marketing, storage, and transportation of oil, gas, coal, and consumable fuels, as well as those that provide equipment and services for the oil and gas industry.
- *Materials Sector*: encompasses companies that produce chemicals, construction materials, glass, paper, forest products and packaging, metals, minerals, and mining companies, such as steel producers.
- *Industrials Sector*: encompasses manufacturers and distributors of capital goods like aerospace and defense, building products, electrical equipment, machinery, and companies that provide construction and engineering services. Additionally, it includes companies that offer commercial and professional services such as printing, environmental and facilities services, office services, security and alarm services, human resource and employment services, research and consulting services. It also includes transportation service providers.

Fund	Sector	ETF Ticker	Weight in the S&P500
1	Materials Select Sector SPDR	XLB	2.69%
2	Communication Services Select Sector SPDR	XLC	7.20%
3	Energy Select Sector SPDR	XLE	5.01%
4	Financials Select Sector SPDR	XLF	11.46%
5	Industrial Select Sector SPDR	XLI	8.61%
6	Technology Select Sector SPDR	XLK	26.09%
7	Consumer Staples Select Sector SPDR	XLP	7.23%
8	Real Estate Select Sector SPDR	XLRE	2.71%
9	Utilities Select Sector SPDR	XLU	3.16%
10	Health Care Select Sector SPDR	XLV	15.81%
11	Consumer Discretionary Select Sector SPDR	XLY	10.03%
	Total		100.00%

Table 1 - S&P500 Sector Indexes ETFs. A portfolio allocated to the 11 ETFs with weights indicated in column 4 replicates the S&P500.

- *Consumer Discretionary Sector*: encompasses businesses that are typically more sensitive to economic cycles. Its manufacturing segment includes automotive, household durable goods, leisure equipment, and textiles and apparel. The services segment includes hotels, restaurants, and other leisure facilities, media production and services, and consumer retailing and services.
- *Consumer Staples Sector*: includes companies whose operations are less affected by economic cycles. It encompasses manufacturers and distributors of food, beverages, and tobacco, as well as producers of non-durable household goods and personal products. It also includes retailers of food and drugs, and hypermarkets and consumer super centers.
- *Health Care Sector*: includes companies involved in health care services, health care equipment, and supplies manufacturers and distributors, and health care technology companies. It also encompasses companies involved in researching, developing, producing and marketing pharmaceutical and biotechnology products.
- *Financials Sector*: encompasses companies involved in banking, thrifts, mortgage finance, specialized finance, consumer finance, asset management, custody banks, investment banking, brokerage, and insurance. It also includes financial exchanges and data, and mortgage real estate investment trusts (REITs).
- *Information Technology Sector*: includes companies that provide software and information technology services, manufacturers and distributors of technology hardware and equipment such as communications equipment, cellular phones, computers and peripherals, electronic equipment and related instruments, and semiconductors.

- *Communication Services Sector*: encompasses companies that enable communication and provide related content and information through various mediums. It includes telecommunications and media and entertainment companies, such as those involved in interactive gaming and those engaged in creating or distributing content and information through their own platforms.
- *Utilities Sector*: includes utility companies such as electric, gas, and water utilities, independent power producers and energy traders and companies that generate and distribute electricity using renewable sources.
- *Real Estate Sector*: encompasses companies involved in real estate development and operation, as well as those offering real estate-related services and equity real estate investment trusts (REITs).

Finally, we simulate forecasts from six experts $k = 6$, as reported in Table 2. Expert forecasts focus on a single risk factor, whereas ETFs provide observations on all three risk factors simultaneously. Thus, simulating six expert forecasts ensures that the effect of the forecast is not dominated by the effect of the ETFs. The first two experts forecast the market risk premium X_1, the next two forecast the market capitalization premium X_2, and the last two forecast the book-to-market premium X_3. For each view, we choose a 90% confidence level as in Davis and Lleo (2020).

View #	Risk Factor	Central View	Confidence Level
1	Mkt-RF	10%	90%
2	Mkt-RF	8%	90%
3	SMB	5%	90%
4	SMB	3%	90%
5	HML	6%	90%
6	HML	4%	90%

Table 2 – Expert views simulation

As anticipated in the previous chapter, three pieces of information are included into each expert forecast: the specific risk factor(s), the central view, and the confidence interval. An example of an expert forecast could be:

“I am 90% confident that the value premium will reach 6% over the next year”.

In this hypothetical forecast, the risk factor is the value premium; the central view is the trajectory of the risk premium, 'reach 6% over the next year'; and the confidence interval is expressed as a 90% subjective probability.

Each expert is subject to a different combination of behavioral biases: overconfidence, excessive optimism, conservatism, confirmation bias, and groupthink. We then remove the effect of behavioral biases to produce unbiased forecasts. To account for excessive optimism and overconfidence biases, Davis and Lleo (2020) simulate the experts' "true" confidence level. Indeed, both biases lead experts to state confidence bounds that are too narrow. However, we use a different approcach explained in the next section. Conservatism does not need to be modeled explicitly because our model does not require experts to update their forecasts. Experts formulate their forecasts at the initial stage when the model is parameterized. The Kalman filter also helps to address the effect of confirmation bias similarly, by confining the effect of this bias to the initial data. Finally, we address groupthink by adding forecasts from dissenting experts and/or stress test scenarios that help to broaden the range of observations and scenarios that enters the Kalman filter, ultimately translating into a less polarized asset allocation.

The next section presents a computational explanation of how we estimate the parameters of the model. This represents a fundamental step to implementing our model.

We estimate the parameters of the risk factors and financial securities prices from weekly prices and returns over the period June 19, 2018 to June 30, 2022 (T = 209 observations). June 19, 2018, is the earliest date for which all data series are available because a revision to the GICS changed the sectors' definition and composition. Nevertheless, this dataset reflects a volatile period on the U.S. stock market, with a sharp rise in 2019, the COVID crash in March 2020, and a substantial part of the subsequent rally. Factors come from Kenneth French's database[8], ETF prices from Yahoo Finance and the return of the money market instrument is the daily rate of a 1-month Treasury Bill[9]. We use this rate to discount securities prices.

3.2. Model Parametrization

To get the parameters for the simulation model, we discretize the SDE for $X(t)$ and $\mathfrak{s}(t)$. Next, we estimate these discretized models' coefficients, which we then use to compute the coefficients for the SDEs. To simplify the modeling process, we assume that all the parameters are constant. As we are working with weekly data, we set the time step $\Delta t := \frac{1}{52}$.

[8] The data is available online at http://mba.tuck.dartmouth.edu/pages/faculty/ ken.french/Data_Library/f-f_factors.html

[9] Source: Kenneth French

The entire implementation has been performed using the coding program R and we refer the reader to Appendix C to have a more detailed presentation of the model implementation.

3.2.1. Drift parameters

First, we estimate the drfit paramters.

The SDE (2.1) for $X(t)$ discretizes into the first order vector autoregressive VAR(1) process

$$X_{t+1} = b_1 + B_1 X_t + \Lambda Z_t \sqrt{\Delta t} \tag{3.1}$$

where $b_1 := b\Delta t$ and $B_1 := I_m + B\Delta t$. We estimate the parameters b_1 and B_1 separately via ordinary least square regression (OLS). Then, we use the OLS estimates to compute the parameters b and B for the factor SDE (2.1).

Next, we discretize the log discounted asset price SDE (2.3) into the linear model

$$\Delta \mathfrak{s}_{it} = (a_i - \frac{1}{2}\Sigma\Sigma'_{ii})\Delta t + (AX_t)_i \Delta t + (\Sigma Z_t^{\mathfrak{s}})_i \sqrt{\Delta t} \tag{3. 2}$$

where $\Delta \mathfrak{s}_t = \mathfrak{s}_{t+1} - \mathfrak{s}_t$, and $Z_t^{\mathfrak{s}}$ is a *d*-dimensional standard normal random variable for every *t.* According to Davis and Lleo (2020), X_t should be independent from the time discretization scheme. However, the factor values in the Fama-French dataset depend on the discretization scheme. To address this inconsistency, we considered instead the linear model

$$y_{it} = \alpha_i + \beta x_t + \varepsilon_{it} \tag{3. 3}$$

where $y_{it} = \frac{\Delta \mathfrak{s}_t}{\Delta t}, \alpha_i = a_i - \frac{1}{2}\Sigma\Sigma'_{ii}$, $\beta_i = A[i,\cdot]'\Delta t$ is the ith row of matrix *A* transposed, $x_t = \frac{x_t}{\Delta t}$ and ε_{it} is an error term. We use ordinary least squares (OLS) to calculate the vector α and matrix β. We then use the OLS estimates for α and β to find the parameters *a* and *A* for the stochastic differential equation (2.3) related to securities prices.

3.2.2. Diffusion Parameters

Second, we estimate the diffusion parameters. More specifically, we estimate the historical diffusion matrix $\Lambda\Lambda'$ by applying the definition of the quadratic variation of $X(t)$, that is

$$\langle X, X\rangle_t = \lim_{\Delta t_k \to 0} \sum_{t_k \le t} \big(X(t_{k+1}) - X(t_k)\big)\big(X(t_{k+1}) - X(t_k)\big)' = \Lambda\Lambda' t$$

Similarly, we estimate the historical diffusion matrix $\Sigma\Sigma'$ using the quadratic variation of $\mathfrak{s}_t = lnS(t)$,

$$\langle \mathfrak{s}, \mathfrak{s}\rangle_t = \lim_{\Delta t_k \to 0} \sum_{t_k \le t} \big(\mathfrak{s}(t_{k+1}) - \mathfrak{s}(t_k)\big)\big(\mathfrak{s}(t_{k+1}) - \mathfrak{s}(t_k)\big)' = \Sigma\Sigma'\mathrm{t}$$

and the cross-variation term $\Sigma\Lambda'$ is estimated via

$$\langle \mathfrak{s}, X\rangle_t = \lim_{\Delta t_k \to 0} \sum_{t_k \le t} \big(\mathfrak{s}(t_{k+1}) - \mathfrak{s}(t_k)\big)\big(X(t_{k+1}) - X(t_k)\big)' = \Sigma\Lambda'\mathrm{t}$$

Next, we compute the $n \times d$ and $m \times d$ diffusion matrices Λ and Σ from the estimates for $\Lambda\Lambda'$, $\Sigma\Sigma'$ and $\Sigma\Lambda'$. Note that

$$\begin{pmatrix} \Lambda \\ \Sigma \end{pmatrix} (\Lambda' \quad \Sigma') = \begin{pmatrix} \Lambda\Lambda' & \Sigma\Lambda' \\ \Sigma\Lambda' & \Sigma\Sigma' \end{pmatrix} =: \mathcal{M}$$

Following Davis & Lleo (2020) we make two assumptions to simplify the estimation process:

Assumption 3.1. The noise in the factor process is generated by the first *n* Brownian motions only.

Assumption 3.2. The noise generated by the securities prices process and the dynamic confidence interval around the forecasts are independent, that is, $\Sigma\Psi'_Z(t) = 0$.

As Davis & Lleo (2020) noted, neither assumption is essential to the argument.
Assumption 3.1 allows us to define $\Lambda := (\lambda \quad 0)$, where λ is a $n \times n$ matrix and 0 is the $\mathrm{n} \times (d - n)$ zero matrix. Additionally, Assumption 3.2 lets us set $\Sigma := (\, \sigma_n \; \sigma_m \; 0)$,where σ_n is a $m \times n$ matrix, σ_m is a $m \times m$ matrix and 0 is the $m \times K$ zero matrix. As a result,

$$\mathcal{M} = \begin{pmatrix} \lambda & 0 & 0 \\ \sigma_n & \sigma_m & 0 \end{pmatrix} \begin{pmatrix} \lambda' & {\sigma'}_n \\ 0 & {\sigma'}_m \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \lambda\lambda' & \lambda{\sigma'}_n \\ \sigma_n\lambda' & \sigma_n{\sigma'}_n + \sigma_m{\sigma'}_m \end{pmatrix}$$

A block Cholesky decomposition of the $(n + m) \times (n + m)$ matrix $\mathcal{M}$ yields the factorization

$$\mathcal{M} = FF', \quad F = \begin{pmatrix} (\Lambda\Lambda')^{1/2} & 0 \\ \Sigma\Lambda'(\Lambda\Lambda')^{-1/2} & \Sigma\Sigma' - \Sigma\Lambda'(\Lambda\Lambda')^{-1}\Lambda\Sigma')^{1/2} \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ \sigma_n & \sigma_m \end{pmatrix}$$

As in Davis and Lleo (2020), we conclude that:

(i) λ is the $n \times n$ matrix resulting from a Cholesky decomposition of the matrix $\Lambda\Lambda'$;

(ii) $\sigma_n = \Sigma\Lambda'(\Lambda\Lambda')^{1/2} = \Sigma\Lambda'\lambda$;

(iii) σ_m is the $m \times m$ matrix resulting from a Cholesky decomposition of the matrix

$$\Sigma\Sigma' - \Sigma\Lambda'(\Lambda\Lambda')^{-1}\Lambda\Sigma' = \Sigma(I - \Lambda'(\Lambda\Lambda')^{-1}\Lambda)\Sigma'.$$

3.2.3. Prior Distribution of the Risk Factors

Once we obtain the parameters' values, we compute the prior distribution of the risk factors. The starting value $X(0)$ follows a normal distribution with a mean of μ_0 and a covariance of P_0. According to Davis & Lleo (2020), we calculate μ_0 econometrically, with a maximum likelihood estimation of a discrete $VAR(1)$ model. The estimated value is a forecast of the risk factors using the discrete $VAR(1)$ process as outlined in (3.1). The result is as follows:

$$\mu_0^{VAR} = b_1 + B_1 X_{209} = (0.2565\%, 0.009\%, -0.1603\%)'$$

The prior covariance matrix P_0 is also an important aspect to consider. According to Davis & Lleo (2020), while determining the prior covariance matrix is crucial for the one-period Black-Litterman model, it is not as crucial for continuous-time models. In line with Davis & Lleo (2020), we interpret P_0 as the mean squared error of the estimate, and as such, any reasonable approximation of the mean squared error of estimates can be used as a proxy for P_0. In this study, we use the average of the covariance matrix $\frac{1}{T}\Lambda\Lambda'$ for simplicity.

Once we estimate all the parameters, we can run the simulation. Figure 3 and Figure 4 below displays the simulation of the risk factors according to SDE (2.1) and the simulation of the financial securities according to SDE (2.3), respectively.

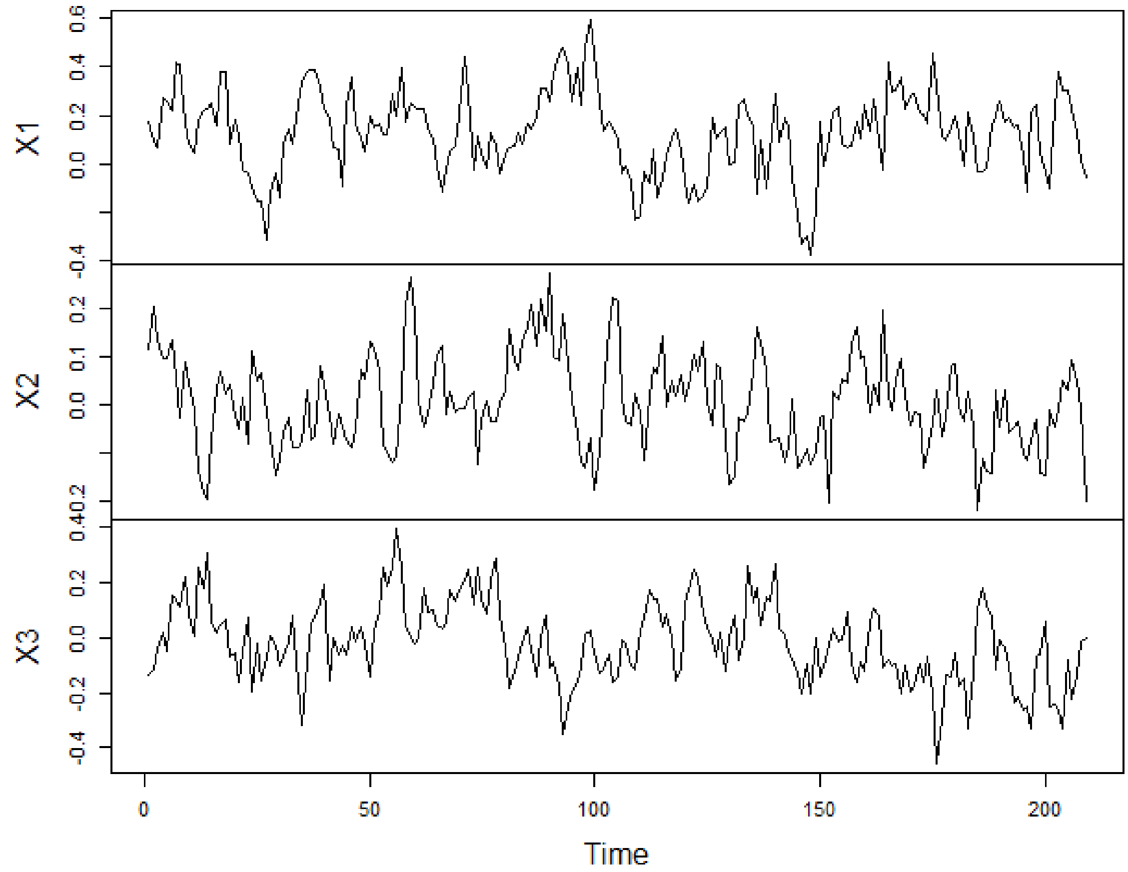
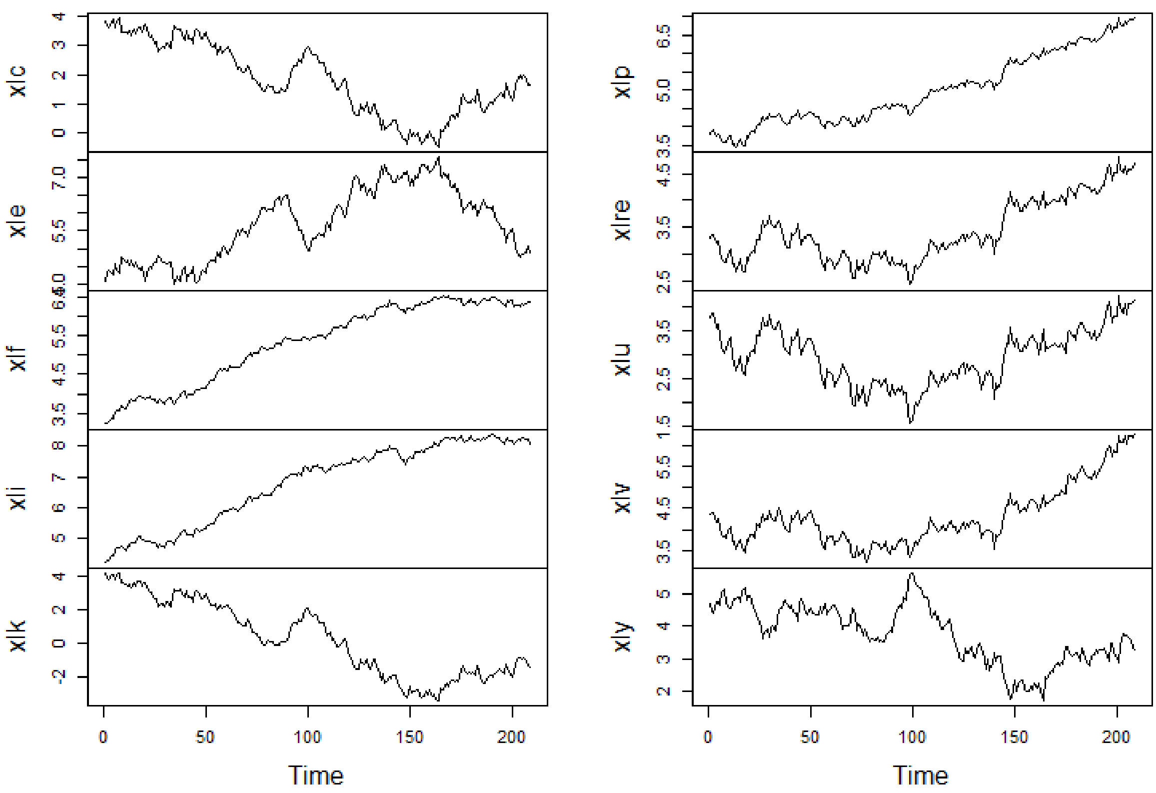


Figure 3 – Risk factors simulation according to SDE 2.1



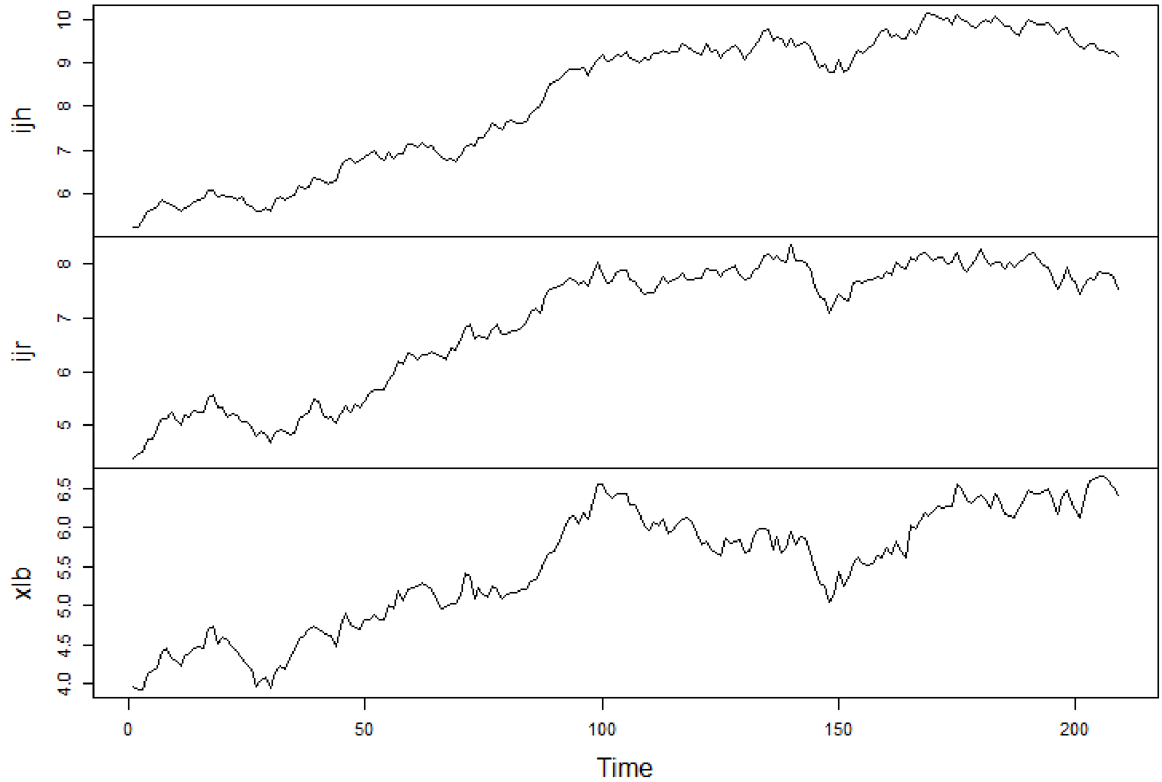


Figure 4 - Financial securities simulation according to SDE 2.3

3.2.4. Parametrizing and Debiasing Expert Forecasts

Finally, we parametrize and debias the expert forecasts. The SDE describing the evolution of expert forecasts is given at (2.5) by

$$dZ(t) = \big(a_Z(t) + A_Z(t)X(t)\big)dt + \Psi_Z(t)dW(t), \quad Z(0) = z.$$

We have 6 expert forecasts: two for the market risk premium X_1, two for the market capitalization premium X_2 and two for the book-to-market premium X_3, so we parametrize A_Z as:

$$A_Z^{Experts} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

We also assume that there is no systematic bias in the drift, so a_Z is the 6-element zero vector.

To parametrize the confidence matrix Ψ_B, we consider that each expert has associated a forecast error, for example, 0.02. This means that the first expert believes that in 90% of cases, the true value (unknown) of factor $X_1(t)$ will be within the interval $[Z_1(t) - 0.02, Z_1(t) + 0.02]$. The forecast error is given by $\Psi_B(t)W(t)$, therefore $\Psi_B(t)$ should be calibrated in such a way that $|\Psi_B(t)W(t)| < 0.02$ in 90% of cases, which corresponds to calculating a quantile of the normal distribution, knowing that $W(t)$ is normal with mean 0 and variance t. Thus, since experts are formulating forecasts over an annual horizon, we take a reference time horizon of 1 year and calculate Ψ_B in such a way that $|\Psi_B(t)W(1)| < 0.02$ in 90% of cases.

Importantly, this confidence matrix is inferred directly from the expert forecasts. It does not address behavioral biases. The BLCT model will be implemented using Ψ_B.

Figure 5 displays the simulated expert forecasts according to SDE (2.5). These expert forecasts have not been debiased, thus they include the effect of behavioral biases.

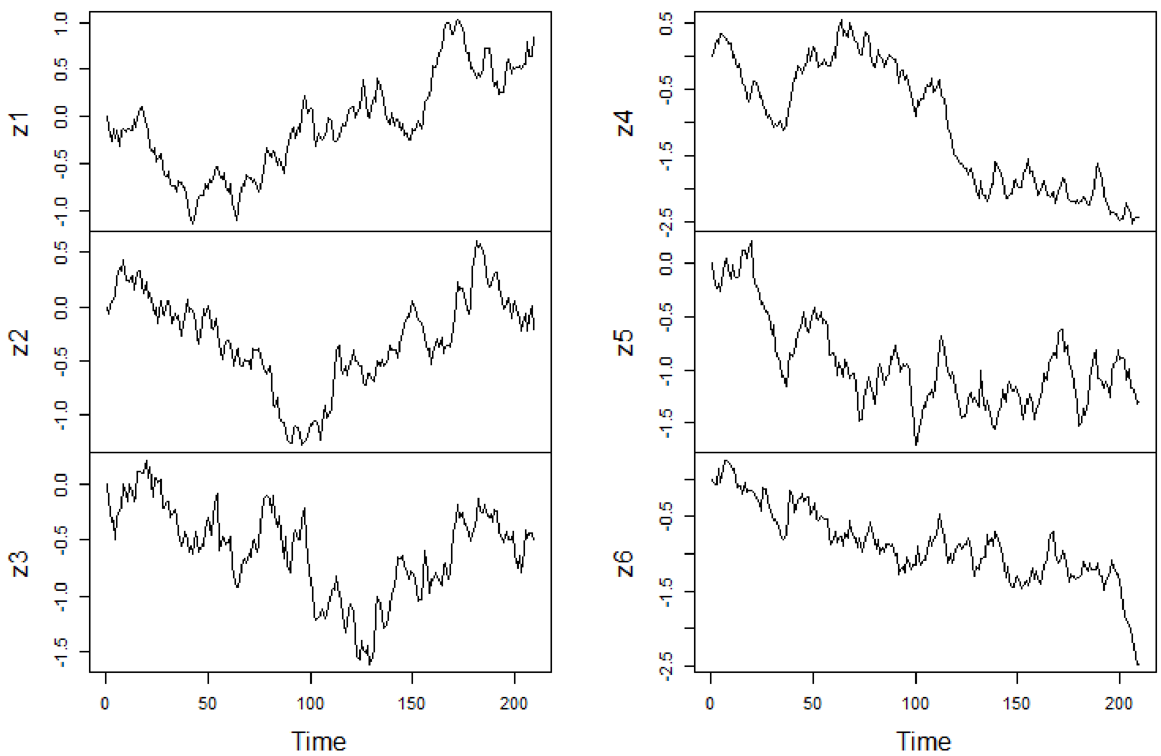


Figure 5 – Simulated expert forecasts (biased) according to SDE 2.5

Next, we need to debias the expert forecasts. Excessive optimism and overconfidence directly affect Ψ_Z. To find the matrix Ψ_Z, we use the estimated value of Ψ_B. Since the experts are overconfident, the forecast errors are greater than stated. Therefore, they need to be amplified to obtain the debiased matrix. To do this, Ψ_B must be multiplied by a given factor, chosen more or less arbitrarily (see Appendix C for details).

Our model does not require explicit modeling of conservatism as experts already address it in the initial stage when formulating their forecasts. The parameters of the prior distribution, μ_0 and P_0, take into account conservatism during the elicitation process. The Kalman filter then incorporates all available information in a Bayesian manner, reducing the effect of conservatism over time. Confirmation bias can also be addressed by including dissenting experts in the pool of experts, as the Kalman filter weighs forecasts based on their accuracy, correcting for any bias.

Then, we model groupthink by adding a correlation structure to the confidence intervals around the forecasts. In this paper, we assume that the correlation is ρ = 0.5. As noted in Davis & Lleo (2020), estimating the correlation between the experts' forecasting errors is challenging in reality. To estimate this correlation, we would need to analyze past expert forecasts to understand the dependence structure between the forecasting errors as well as the dependence between forecasting errors and factors.

To include correlation in our model, we introduce the block matrix

$$\begin{pmatrix} \Lambda \\ \Psi_D \end{pmatrix} \begin{pmatrix} \Lambda' & \Psi'_D \end{pmatrix} = \begin{pmatrix} \Lambda\Lambda' & \Lambda\Psi'_D \\ \Psi_D\Lambda' & \Psi_D\Psi'_D \end{pmatrix}$$

By Assumptions 3.1 and 3.2, we ignore the financial securities prices in this derivation and express Λ as a $n \times (n+k)$ matrix and Ψ_D as a $k \times (n+k)$ matrix. By Assumption 3.1, the noise in the factor $X(t)$ solely comes from the first n Brownian motions, so we can write $\Lambda = (\lambda \quad 0_{nk})$ where λ is a $n \times n$ matrix and 0_{nk} is a $n \times k$ matrix of zeros.

The next step is to partition Ψ_D as $\Psi_D = (\psi_n \quad \psi_k)$ where ψ_n and ψ_k are respectively a $k \times n$ and a $k \times k$ matrix. Then

$$\Psi_D\Lambda' = (\psi_n \quad \psi_k) \begin{pmatrix} \lambda' \\ 0_{nk} \end{pmatrix} = \psi_n\lambda'$$

Alternatively, we express the covariance of the factors *X(t)* and views *Z(t)* as:

$$\begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \begin{pmatrix} R_1 & R_{12} \\ R'_{12} & R_2 \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} = \begin{pmatrix} \sigma_1 R_1 \sigma_1 & \sigma_1 R_{12} \sigma_2 \\ \sigma_2 R'_{12} \sigma_1 & \sigma_2 R_2 \sigma_2 \end{pmatrix}$$

where σ_1 and σ_2 are a $n \times n$ and $k \times k$ diagonal matrices of standard deviations and R_{11}, R_{22} and R_{12} are $n \times n$, $k \times k$ and $n \times k$ correlation matrices.

Equating the two expressions, we get $\Psi_D \Lambda' = \psi_n \lambda' = \sigma_2 R'_{12} \sigma_1$, and deduce that $\psi_n = \sigma_2 R'_{12} \sigma_1 (\lambda')^{-1}$.

Another way to counteract groupthink is to include stress test scenarios in forecasts, with the understanding that protecting against large losses during difficult times is more important than maximizing gains in favorable market conditions. These stress tests should have a wide, heavily skewed range of possible outcomes, as it is unlikely that the realized value of $X(t)$ will be as extreme as the test scenario predicts, but it could potentially be worse. In a Gaussian model, it is crucial to establish a wide range of possible outcomes around the stress test scenario.

For this implementation, we develop a stress test scenario using data from the one year representing the Covid-19 crisis. We use data from March 1, 2020 to March 1, 2022 for three risk factors and define the stress test scenario as a set of three "observations" with

$$dZ^{ST}(t)(t) = \left(a_Z^{ST} + A_Z^{ST} X(t)\right)dt + \Psi_Z^{ST} dW(t), \qquad Z(0) = z \tag{3.4}$$

$$a_Z^{ST} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \qquad A_Z^{ST} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Since the stress test scenario is not a prediction of how the risk factors will evolve, the uncertainty in the forecast is not related to the uncertainty in the factors or the securities. However, the three observations in the stress test are related to each other, so we express Ψ_Z^{ST} as $\Psi_Z^{ST} = (0_{n+m} \quad \psi^{ST})$ where 0_{n+m} is a $3 \times (n \times m)$ zero matrix and Ψ_Z^{ST} is a 3×3 matrix. To get ψ^{ST}, we started by estimating $\Psi_Z^{ST}(\Psi_Z^{ST})' = \psi^{ST}(\psi^{ST})'$ as the quadratic variation of $X(t)$ over the duration of the Covid-19 crisis. We then performed a SVD decomposition of $\psi^{ST}(\psi^{ST})'$ and computed ψ^{ST} as the square root matrix implied by the SVD decomposition.

Figure 6 below shows the simulation of the three observations in the stress test scenario according to SDE (3.4). We can see that the three observations present extreme value to address the effect of groupthink on expert forecasts.

Once we take into consideration the impact of behavioral biases on expert views, we can simulate the debiased expert forecast according to SDE (2.6) as displayed in Figure 7. These values will then be used together with financial data to estimate the unobservable risk factors through the Kalman filter.

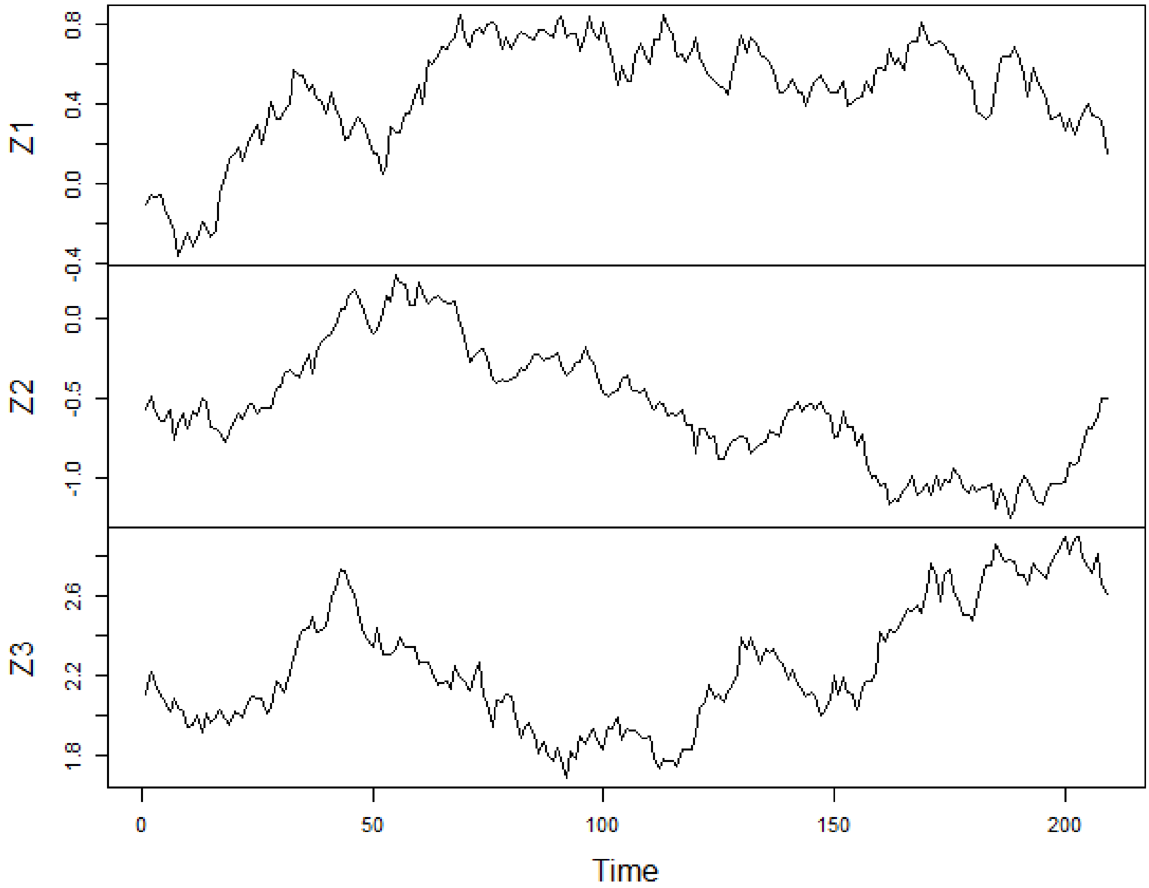


Figure 6 – Simulated observation in the stress test scenario according to SDE 3.4

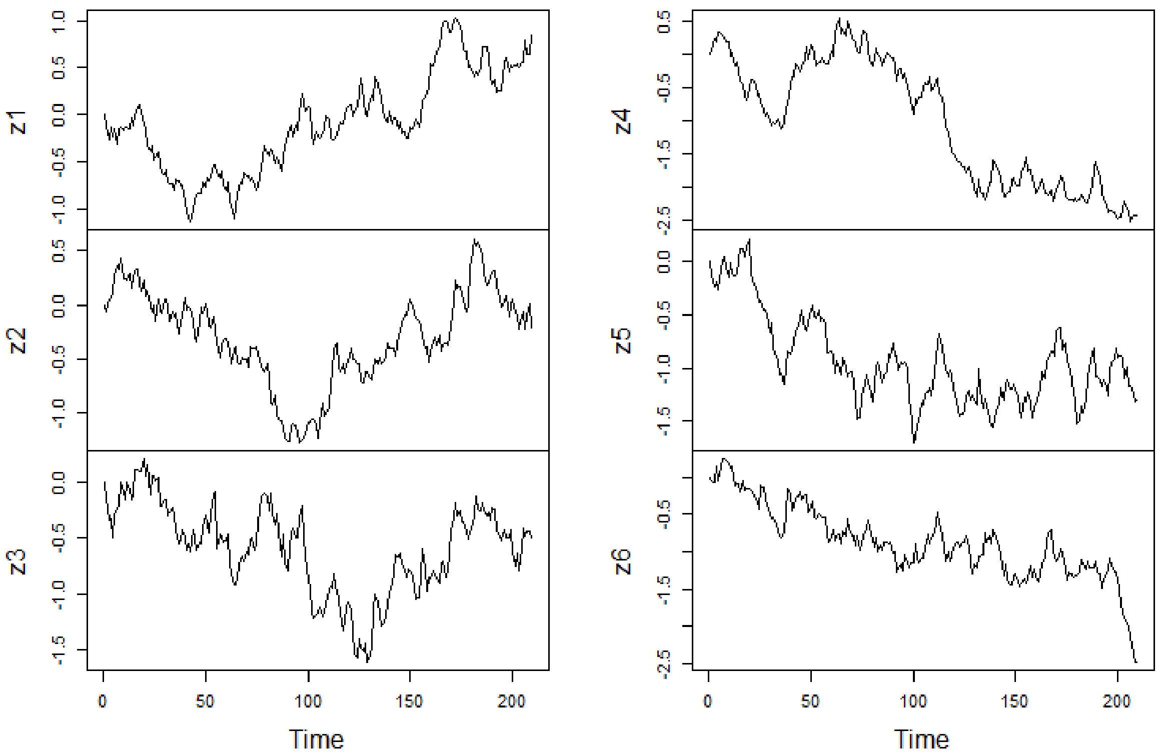


Figure 7 - Simulated debiased expert forecasts according to SDE 2.6

We compare the simulated biased expert forecasts from Figure 5 with the debiased one from Figure 7. We notice that expert forecasts maintain the same path over time but debiasing slightly increases the variability of the forecasts. However, it is not necessarily true that debiased expert forecasts are more volatile than biased expert forecasts. The volatility of a forecast depends on many factors, such as the accuracy of the forecasting method, the stability of the underlying data and trends, and the level of uncertainty involved in the forecast. The extent to which a forecast is biased or debiased may also play a role, but it is not a determining factor. In some cases, removing bias from a forecast may result in greater volatility, but in other cases it may result in more accurate and stable forecasts. The relationship between bias and volatility in forecasts is complex and context-dependent.

3.3. Model Selection

In this section, to examine the effect of behavioral biases on portfolio optimization, we compare two implementations of the model we presented in the previous chapter, with one variation. The main features of the implementations are summarized in Table 3 and described below.

Model name	Partial or Full observation?	Number of securities observed	Number of forecasts observed	Are the forecasts debiased?
BLCT	Partial	13	6	no
BB	Partial	13	6	yes

Table 3 - Main features of the five implementations considered

The first model is the Black-Litterman model in continuous-time (BLCT) as in Davis and Lleo (2013). The investor has access to ETF prices and 6 expert forecasts, but the expert forecasts are not debiased. This model maintains the same structure of the one introduced in the previous chapter with the only difference that in the BLCT we do not address the effect of behavioral biases.

The second model is the behavioral Black-Littermna model in continuous-time (BB) proposed in this paper and presented by Davis and Lleo (2020). The investor observes both ETF prices and debiased expert forecasts. Because we address the effect of behavioral biases, we expect the BB model to be less aggressive and more efficient than the BLCT.

Comparing the BLCT model with biased expert forecasts and the BB model with debiased expert forecasts allows us to analyze the impact of behavioral biases on asset allocation by simulating the effect of different biases on asset prices and returns. In the BLCT model, the biased expert forecasts represent the investor's belief about the expected returns of the risk factors, which may be influenced by various behavioral biases as explained in the previous chapters. On the other hand, the BB model represents the scenario where the effect of the biases has been addressed. By comparing the results of these two models, we can quantify the impact of behavioral biases on asset allocation and determine how they may affect the overall performance of an investment portfolio.

3.4. Simulation

We consider the following example of a U.S. equity portfolio manager with 2-year investment horizon (104 weeks) and an initial wealth of $10,000. As said before, the manager uses n = 3 risk factors, comprised of the three factors of the Fama-French model (Fama & French, 2015). The investment universe consists of m = 13 U.S. Exchange Traded Funds (ETFs), tracking the 11 S&P500 sectors, plus a mid-cap and a small-cap index. Thus, all 13 ETFs are available for investment $m_1 = m$. The manager has access to k = 6 expert opinions, exhibiting a mix of behavioral biases. To simplify the implementation, the models are optimized for a risk aversion $\theta = 0$ corresponding to the Kelly criterion.

In both our models, we use the Kalman filter to determine the parameters of the distribution of the factors using financial securities and expert forecasts. In BLCT, the expert opinions are biased, whereas in the BB model, they are debiased. The estimated value of the factor process is reflected on the optimal asset allocation.

We perform 3,000 simulations partitioning the 2-year investment horizon into 104 weekly intervals, and pool the resulting 312,000 simulated excess returns to compute sample distributions and performance metrics. We study the output by comparing the portfolio performance of the two models introduced before.

Figure 7 displays the simulation of 10 trajectories of the BLCT model. Even though we simulate 3,000 trajectories, we present just 10 trajectories to make the chart clear to the reader.

Figure 8 below shows the simulation of 10 trajectories of the BB model. For the same reason described above, we just present 10 trajectories to make the chart clear to read.

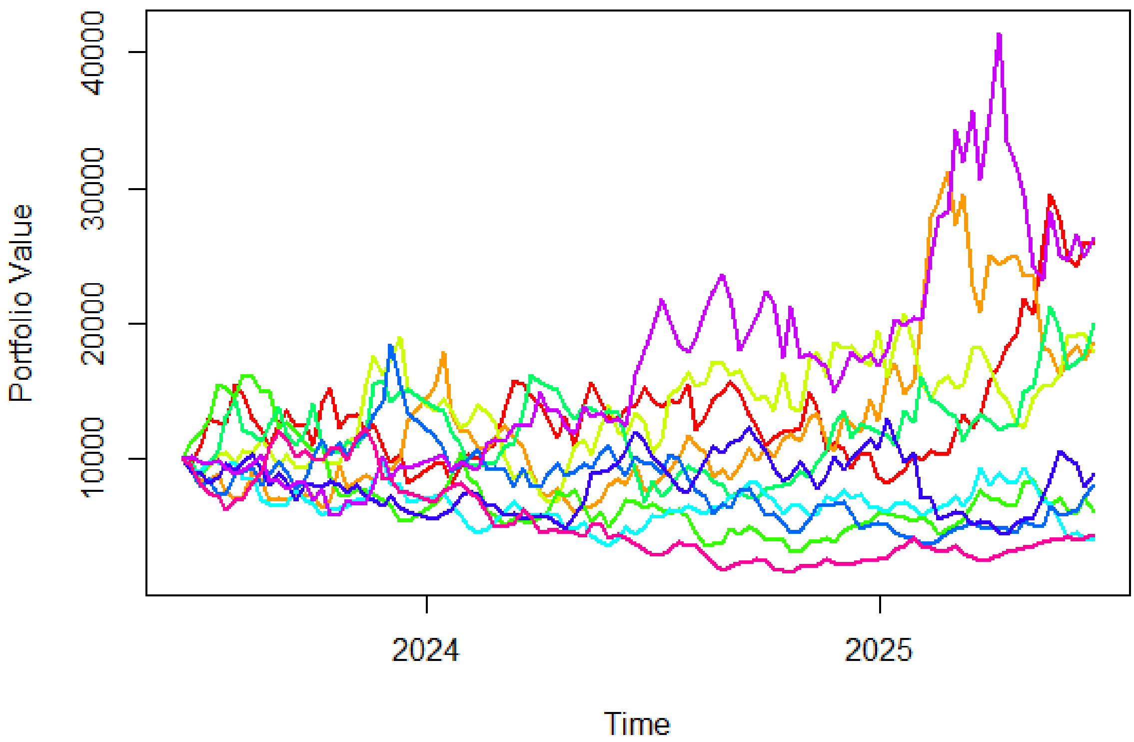


Figure 8 – Simulation of 10 trajectories BLCT

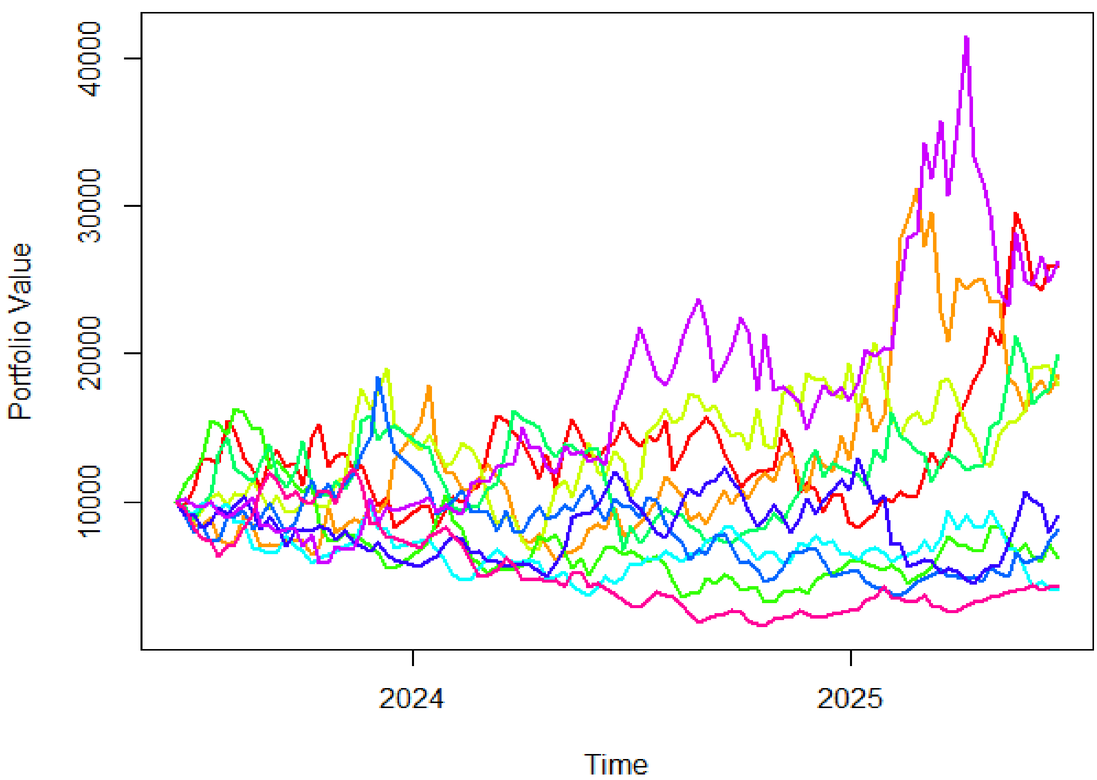


Figure 9 - Simulation of 10 trajectories BB

The two figures show similar paths considering a 2-year forecast period and present only slight differences.

3.5. Examining Weekly Portfolio Performance

We examine the effect of behavioral biases on asset allocations by comparing the weekly performance measures of the models.

Using portfolio performance measures is a good practice in investment analysis because it provides a comprehensive evaluation of the portfolio's overall performance. Portfolio performance measures, such as the ones described in the first chapter, provide valuable insights into the portfolio's behavior, allowing investors to make informed investment decisions. These measures can help investors assess the effectiveness of their investment strategy, determine if they are meeting their investment goals, and evaluate the impact of market movements on their portfolios. By using portfolio performance measures, investors can make informed investment decisions and potentially maximize returns while minimizing risk.

Table 4 reports the portfolio performance metrics based on simulated weakly excess returns for an investor with a Kelly fraction f = 100% (equivalent to θ = 0). As explained in the first chapter, these metrics include: summary statistics, tail risk measures, and portfolio efficiency measures. The mean and standard deviation are the summary statistics on the (log) excess return. The 99th percentile Value-at-Risk (VaR) and the Conditional Value-at-Risk (CVaR) are the two tail risk metrics and are expressed as a percentage return relative to the portfolio's mean return. Finally, the Sortino ratio and the Sharpe ratio are examples of portfolio efficiency metrics.

Weekly portfolio performance metrics are computed by pooling into a single sample the 312,000 excess returns across time intervals and simulation paths (3,000 paths × 104 weekly intervals = 312,000 excess return data). For example, the mean weekly excess return is computed as

$$Mean = \frac{1}{3{,}000x104}\sum_{i=1}^{3{,}000}\sum_{j=1}^{104}\bar{r}_{ij}$$

,

where $\bar{r}_{ij}$ is the portfolio return at time j for simulation i. All the other performance metrics follow this methodology.

The mean excess (log) return, the first summary statistic, increase from the BLCT to the BB model by around 4% in relative terms. In our model, the introduction of debiased expert forecasts results in substantially better returns.

Performance Metrics	BLCT	BB
θ = 0		
Mean	0.48%	0.52%
Standard deviation	11.10%	11.00%
VaR 99	24.80%	24.70%
CVaR 99	28.70%	24.80%
Sharpe	0.0440	0.0473
Sortino	0.0703	0.0755

Table 4 – Weekly portfolio performance measures

By looking at the standard deviation, we notice that the BLCT has the highest deviation illustrating the complete impact of expert forecasts on portfolio performance. The additional risk in the BLCT model over the BB model is almost entirely due to behavioral factors.

Additionally, VaR and CVaR have a comparable pattern. Indeed, BB shows a lower value for both measures compared to BLCT.

Risk and return are related by portfolio efficiency, which is measured using the Sharpe and Sortino ratios. Since BLCT uses ETF prices and raw expert forecasts, we expect it to perform worst compared to the BB where the impact of behavioral biases is addressed. The outcomes confirm our expectations. Indeed, the Sharpe ratio of BLCT is 0.0440 per week. Debiasing helps improve efficiency. BB has a Sharpe ratio of 0.0473 which is 7.5% higher than BLCT.

Overall, our findings are consistent with Davis and Lleo (2020), that is, adding debiased expert forecasts improves the portfolio’s performance. However, we find that expert forecasts do not influence portfolio performance as sharply as reported in Davis and Lleo (2020).

Our analysis shows that behavioral biases have a significant impact on asset allocation and it is important to consider them when constructing investment portfolios. These biases can lead to suboptimal investment decisions and negatively impact portfolio performance. By taking into account the effects of biases such as overconfidence, herding, and anchoring, investors can make more informed investment decisions and potentially improve their portfolio's performance. It is therefore important to address these biases in the investment process in order to obtain better optimal portfolios and achieve better investment outcomes.

Conclusion

Behavioral biases are systematic patterns of deviation from normative decision-making that can lead to suboptimal decisions. They are caused by cognitive and emotional factors such as the ones we analyzed in chapter one.

The study conducted by CXO Advisory Group LLC1 analyzed 6,582 forecasts for the U.S. stock market made by 68 experts between 2005 and 2012. The average accuracy among the experts was 47.4%, with some experts having accuracy as low as 21% and others as high as 68%. Among the factors contributing to these results, the most relevant was the lack of consideration of the impact of behavioral biases on expert forecasts.

Behavioral biases can have a significant impact on asset allocation decisions because they can influence the way experts to form their forecasts and the way investors make their investment decisions. These biases can lead to poor investment decisions, such as overinvestment in certain assets and underinvestment in others. This can result in poor portfolio performance and increased risk. For example, if investors are overconfident in their ability to predict the future, they may make too many high-risk investments, which can lead to significant losses if those investments do not perform as well as expected. Similarly, if investors are influenced by herding behavior, they may invest in assets that are popular among other investors, regardless of whether those assets are undervalued or overvalued.

Therefore, investors and experts need to identify and address their behavioral biases to make more informed and effective asset allocation decisions. This can be achieved by using debiasing techniques or by seeking out unbiased expert forecasts.

In this thesis, we have demonstrated that behavioral biases can significantly impact expert forecasts and, subsequently, on asset allocation decisions. Using the Continuous-Time Black-Litterman model proposed by Davis and Lleo (2013), we have shown that carefully formulated debiased forecasts improve portfolio efficiency while biased forecasts produce the opposite effect. The proposed model is an integrated behavioral continuous-time portfolio selection model which can be solved in closed form and can also be used to identify and reduce the impact of five main behavioral biases.

Our procedure involved six distinct steps. First, we parameterized the financial market and defined the input of the models. Second, we collected expert opinions and views. Third, we addressed the impact of behavioral bias. Fourth, we combined data with opinions and applied the Kalman filter to estimate the current level of the risk factors. Fifth, we optimized the portfolio, and finally, we analyzed the effect of biased expert forecasts.

To investigate the impact of behavioral biases on optimal asset allocation, we evaluated the performance of the Behavioral Black-Litterman model in continuous time against the BLCT model where the impact of behavioral biases is not addressed. Finally, we analyzed the weekly performance of the portfolio showing that BB perform better than BLCT.

Our empirical analysis confirms that identifying and addressing behavioral biases in expert forecasts is essential in controlling risks in asset allocation. By recognizing and addressing these biases, investors and experts can make more informed and effective decisions that lead to better portfolio performance and reduced risk.

Appendix A

A.1. Probability Spaces

We start by introducing the mathematical concept of a probability space, which has three components $(\Omega, \mathcal{F}, \mathbb{P})$, where

- Ω is the set of all states of the world.
  An element $\omega \in \Omega$ represents a specific state of the world.
- $\mathcal{F}$ is a sigma-algebra on Ω.
  $\mathcal{F}$ contains all events that can be described. It is defined as sigma-algebra and it represents the historical but not future information available on our stochastic process. A sigma-algebra $\mathcal{F}$ on Ω is a family of subsets of Ω such that $\Omega \in \mathcal{F}$; if $A \in \mathcal{F}$, then $A^C \in \mathcal{F}$; if $(A_n)_{n \in N}$ is a countable family of elements of $\mathcal{F}$, then $\bigcup_{n \in N} A_n \in \mathcal{F}$.
- $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$, i.e. a function $\mathbb{P} : \mathcal{F} \to [0, 1]$.
  For $A \in \mathcal{F}$, then $\mathbb{P}$ (A) measures how likely is the realization of event A. On $(\Omega, \mathcal{F}, \mathbb{P})$, a probability measure $\mathbb{P}$ is a function $\mathbb{P} : \mathcal{F} \to [0,1]$ such that $\mathbb{P}(\Omega) = 1$ and, if $(A_n)_{n \in N}$ is a family of disjoint events belonging to $\mathcal{F}$ (in the sense that $A_i \cap A_j = \emptyset$ for all $i \neq j$), then

$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mathbb{P}(A_n)$$

A.2. Filtration

The notion of filtration, introduced by Doob, has become a fundamental feature of the theory of stochastic processes. Most basic objects, such as martingales, semimartingales, stopping times, or Markov processes, involve the notion of filtration.

Definition. Let Ω be the set of states of the world. A filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in [0,T]}$ is an increasing family of sigma-algebras on Ω, so that $\mathcal{F}_s \subseteq \mathcal{F}_t$, for all $0 \leq s \leq t \leq T$.

$\mathbb{F}$ represents the information flow of the market, as time passes you collect more and more information (from market prices and other sources).

- For a random variable ξ and $t \in [0, T]$, the conditional expectation
  $$E[\xi | \mathcal{F}_t]$$
  represents the expectation of ξ given the information available at date t.

- We say that ξ is independent of $\mathcal{F}_t$ whenever the information collected up to date t is useless to forecast the value of ξ. In this case

$$E[\xi|\mathcal{F}_t] = E[\xi].$$

- If ξ is $\mathcal{F}_t$ -measurable (i.e., fully determined by $\mathcal{F}_t$), then

$$E[\xi|\mathcal{F}_t] = \xi.$$

- For any dates 0 ≤ s ≤ t ≤ T, we have the tower property:

$$E[\mathbb{E}[\xi|\mathcal{F}_t]|\mathcal{F}_s = E[\mathbb{E}[\xi|\mathcal{F}_s]|\mathcal{F}_t = E[\xi|\mathcal{F}_s].$$

A.3. Stochastic Processes

In deterministic processes, we study a phenomenon that depends on time, of which we are able to predict the exact evolution over time. However, in order to describe those phenomena whose evolution is influenced by random events classical analysis is no longer adequate and it is necessary to introduce the stochastic processes, based on probability theory (Gallagher, 2013).

Definition. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A stochastic process $X = (X_t)_{t\in[0,T]}$ is a family of random variables $X_t : \Omega \to \mathbb{R}$, indexed with respect to t ∈ [0, T].

Interpretation:

- $X = (X_t)_{t\in[0,T]}$ describes the random evolution of a phenomenon over time;
- X_t is the value of the process X at time t, for t ∈ [0,T];
- for each ω ∈ Ω, the map $t \to X_t(\omega)$ denotes the trajectory (or path) of X associated to a specific state of the world ω.

A.4. Brownian Motion

The term Brownian derives from the name of the botanist Robert Brown that in 1827 observed that the movement of a pollen grain suspended in a liquid (e.g. water) follows chaotic and irregular movements. Albert Einstein in 1905 formulated a mathematical model of Brownian Motion. But already in 1900, L. Bachelier had used Brownian Motion to describe the movement of stock prices and other financial indices on the Paris stock market (Borrelli, 2012). The random motion of a particle in a fluid subject to collision and influence of other particles is called Brownian Motion. One of the mathematical models of this motion is the Wiener Process (Krishnan, 2006).

Definition. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space endowed with a filtration $\mathcal{F}$. A stochastic process $W = (W_t)_{t\in[0,T]}$ starting from $W_0 = 0$ is a Brownian motion if

1. $W_t - W_s$ is independent of $\mathcal{F}_s$, for all $0 \leq s \leq t \leq T$;
2. $W_t - W_s \sim N(0, t-s)$, for all $0 \leq s \leq t \leq T$;
3. W has continuous trajectories.

A process with property (1) is called a process with independent increments. Property (2) implies that the distribution of the increment $W_t - W_s$ only depends on t−s. This is called the stationarity of the increments. A stochastic process that has property (3) is called a continuous process. Similarly to this, if almost all of a stochastic process' sample paths are right-continuous functions, the process is said to be right-continuous. For processes where sample pathways have right-continuous left-hand limits at every time point, we frequently abbreviate them as cadlag (continu 'a droite, limites 'a gauche).

The name Wiener process is also now more common when discussing technical details. The term Brownian motion is often used for the phenomenon of diffusion but both seem to be used interchangeably for the process.

A.5. Ornstein-Uhlenbeck Process

The Ornstein-Uhlenbeck process is a stochastic process that satisfies the following stochastic differential equation:

$$dX_t = \kappa(\theta - X_t)dt + \sigma dW_t \tag{5.1}$$

where W_t is a standard Brownian motion on $t \in [0, \infty) t \in [0, \infty)$.

The constant parameters are:

- $\kappa > 0$ is the rate of mean reversion;
- θ is the long-term mean of the process;
- $\sigma > 0$ is the volatility or average magnitude, per square-root time, of the random fluctuations that are modeled as Brownian motions.

If we ignore the random fluctuations in the process due to dW_t, then we see that X_t has an overall drift towards a mean value θ. The process X_t reverts to this mean exponentially, at rate κ, with a magnitude in direct proportion to the distance between the current value of X_t and θ.

This can be seen by looking at the solution to the ordinary differential equation $dx_t = \kappa(\theta - X_t)dt$ which is

$$\frac{\theta - x_t}{\theta - x_0} = e^{-\kappa(t-t_0)}\ , \ or\ x_t = \theta + (x_0 - \theta)e^{-\kappa(t-t_0)} \qquad (5.1)$$

For this reason, the Ornstein-Uhlenbeck process is also called a mean-reverting process, although the latter name applies to other types of stochastic processes exhibiting the same property as well.

The solution to the stochastic differential equation (5.1) defining the Ornstein-Uhlenbeck process is, for any $0 \leq s \leq t_0 \leq s \leq t$, is

$$X_t = \theta + (X_s - \theta)e^{-\kappa(t-s)} + \sigma \int_s^t e^{-\kappa(t-u)} dW_u$$

where the integral on the right is the Itô integral.

For any fixed s and t, the random variable X_t, conditional upon X_s, is normally distributed with

$$mean = \theta + (X_s - \theta)e^{-\kappa(t-s)}, \qquad variance = \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa(t-s)}\right)$$

Observe that the mean of X_t is exactly the value derived heuristically in the solution (5.2) of the ODE.

The Ornstein-Uhlenbeck process is a time-homogeneous Itô diffusion.

Appendix B

This section provides the mathematical computation to obtain equation (2.3) starting from equation (2.2).

To apply Ito's lemma to the function $\mathfrak{s}_i(\mathrm{t}) = \ln(S_i(\mathrm{t}))$ we use the chain rule and the fact that the derivative of $ln(x)$ is $\frac{1}{x}$.

$$d\mathfrak{s}_i(\mathrm{t}) = dln(S_i(t)) = \left(\frac{1}{S_i(t)}\right) dS_i(t) - \left(\frac{1}{2S_i^2(t)}\right) d[S_i(t)]^2$$

By substituting in the original equation for $dS_i(t)$ and $d[S_i(t)]^2$ we get:

$$d\mathfrak{s}_i(\mathrm{t}) = \left(\frac{1}{S_i(t)}\right)\left[S_i(t)(a(t) + A(t)X(t))_i dt + S_i(t)\sum_{j=1}^{d}\sigma_{ij}dW_j(t)\right] - \left(\frac{1}{2S_i^2(t)}\right)\sum_{j=1}^{d}\left[\sigma_{ij}dW_j(t)\right]^2$$

Simplifying the equation we get:

$$d\mathfrak{s}_i(\mathrm{t}) = \left(a(t) + A(t)X(t) - \frac{1}{2}\sum_{j=1}^{d}\sigma_{ij}^2\right)dt + \sum_{j=1}^{d}\sigma_{ij}dW_j(t)$$

So the final equation for the derivative of the stochastic process $\mathfrak{s}_i(\mathrm{t}) = \ln(S_i(\mathrm{t}))$ is:

$$d\mathfrak{s}_i(\mathrm{t}) = \left[\left(a(t) - \frac{1}{2}\sum_{j=1}^{d}\sigma_{ij}^2\right) + A(t)X(t)\right]dt + \sum_{j=1}^{d}\sigma_{ij}dW_j(t)$$

with initial condition $\mathfrak{s}(0) = ln(s)$

It's worth noting that the difference between the two equations is the second term of the first equation, which is the volatility term that comes from the quadratic variation of the process, it's a non-stochastic term and it's subtracted in this case as it's squared.

Appendix C

R Code

C.1. Parametrizing the Model

The following lines of code are used to parametrize the model and estimate the parameters necessary for the model implementation. All the data for risk factors, financial securities, and expert vires are retrieved from the excel file “Input data”.

```
library(dplyr)
library(lubridate)
library(dynlm)
library(lmtest)

# Import and clean data

etf_data = readxl::read_xlsx('data/Input data.xlsx', sheet = 1) %>% janitor::clean_names()

risk_factors = readxl::read_xlsx('data/Input data.xlsx', sheet = 2, skip = 4) %>%
  janitor::clean_names() %>%
  mutate(date = as.Date(as.character(date), "%Y%m%d"))


# filter for weekly data
all_data = inner_join(etf_data, risk_factors, by = 'date') %>%
  filter(weekdays(date) == 'martedì')

# transforms to time series data to be managed by time series modeling functions like VAR()
ts_all_data = ts(all_data %>% dplyr::select(-date), freq = 365.25/7, start =
 decimal_date(ymd("2018-06-19")))

# Calculate Drift parameters for risk factors (A.1)
# define delta_t as stated in the paper, this has the effect of consider weekly dynamics
dt = 1/52

risk_factors_names = c('mkt_rf', 'smb', 'hml')
#securities_cols = c(2:9, 11:12)
securities_cols = c(2:14)

# define risk matrix and securities matrix (plus s = logS matrix)
X_ts = ts_all_data[, risk_factors_names]
X = all_data[, risk_factors_names] %>% as.matrix
S_ts = ts_all_data[, securities_cols]
S = all_data[, securities_cols] %>% as.matrix
s = log(S)
k_ST = 3

dim(s)
```

```
dim(X)

# dimensions
n = ncol(X)
m = ncol(S)
k = 2*3 # confirm that k is the number of forecasts (2expert x 3 risk factors)
K = k #s till not clear what K is,
d = n + m + k
t = nrow(X)

VAR_data = window(X_ts, start = decimal_date(ymd("2018-06-19")))

# estimate model coefficients using VAR(), it uses OLS in the background to estimate
 coefficients
VAR_est = vars::VAR(y = VAR_data, p = 1, type = 'const')


# extract coefficients from the varest object
mkt_rf_coef = VAR_est$varresult$mkt_rf$coefficients
smb_coef = VAR_est$varresult$smb$coefficients
hml_coef = VAR_est$varresult$hml$coefficients

# vector b1 of constants
b1 = c(mkt_rf_coef['const'], smb_coef['const'], hml_coef['const'])
length(b1) == n

# matrix B1 of coefficients: the coefficient in position 1,2 is the coef for smb.l1 in the model
 where y = mkt_rf
B1 = matrix(c(mkt_rf_coef[1:n], smb_coef[1:n], hml_coef[1:n]),
         nrow = n, ncol = n, byrow = T)
colnames(B1) = risk_factors_names
rownames(B1) = risk_factors_names # to be read as y of the model
dim(B1) == c(n,n)

# deducing b and B for the SDE 3.1
In = base::diag(nrow = n, ncol = n) # identity matrix of dimension n = 3

dim(In) == c(n,n)

b = b1/dt # a vector 1xn = 1x3 of constant of the drift part of the SDE
B = (B1-In)/dt # a nxn (3x3) matrix of risk factor coefs, the multiplicative effect of the drift in
 the SDE 3.1

length(b) == n
dim(B) == c(n,n)

# calculate log returns for regression and adding a row of 0 logret for dimension purposes

matrix_lm = lm(as.matrix(s) ~ X)

summary(matrix_lm)
# deriving a and A (3.3): A should be mxn (11x4) and a should have length m (11)
```

```
a = coef(matrix_lm)[1,]
A = t(coef(matrix_lm)[-1,])

length(a) == m
dim(A) == c(m,n)


# estimate diffusion matrix Lambda using quadratic variation of risk factor matrix
quadratic_variation_X = (t(diff(X)) %*% diff(X))
quadratic_variation_s = (t(diff(s)) %*% diff(s))
quadratic_variation_sX = (t(diff(s)) %*% diff(X))

#quadratic_variation_X_prof = apply(X, 2, get_quadratic_variation)

dim(quadratic_variation_X) == c(n,n)
dim(quadratic_variation_s) == c(m,m)
dim(quadratic_variation_sX) == c(m,n)

# Cholesky factorization of quadratic variation to obtain lambda (Appendix A.2)
lambda = chol(quadratic_variation_X/t)
Lambda = cbind(lambda, matrix(rep(0, n*(d-n)), n, d-n))

dim(lambda) == c(n,n)
dim(Lambda) == c(n,n+(d-n))
dim(Lambda) == c(n,d)

# Obtain Sigma

sigma_n = (quadratic_variation_sX) %*% lambda #m x n

sigma_m_helper = quadratic_variation_s - quadratic_variation_sX %*%
 solve(quadratic_variation_X) %*% t(quadratic_variation_sX)
sigma_m = chol(sigma_m_helper)

(Sigma = cbind(sigma_n, sigma_m, matrix(0, m, k)))

dim(Sigma) == c(m, n+m+k)
```

C.2. Computing the Prior Distribution of the Risk Factors

This part of the code is related to the computation of the prior distribution of the risk factors. More specifically, μ_0 and P_0.

```
# Estimate mu_0
forecast = predict(VAR_est, n.ahead = 1)
mu_0 = b1 + B1 %*% X[209, ]

# Estimate P_0
tau = 1/t
P_0 = tau*(Lambda%*%t(Lambda))
```

```
# C1.1
expert_errors = seq(from = 0.015, to = 0.06, by = 0.005)
sampled_lb = sample(expert_errors, k)

#lower_bound_matrix = matrix(sampled_errors, 2, 3, byrow = F)
# sampled lb represent the 1-sided confidence bound
# around the central view (that for expert correspond tom 90% cl)

expert_central_view = c(0.1, 0.09, 0.06, 0.07, 0.1, 0.08)

set.seed(123)

Psi_B = matrix(rnorm(k*(n+m+k), 0.1, 0.3), k, n+m+k)
#quantile(abs(rnorm(100, 0.1, 0.5)*rnorm(100, 0, 1)), probs = seq(0, 1, 0.1))['90%']
#quantile(abs(as.numeric(Psi_B)*rnorm(132, 0, 1)), probs = seq(0,1, 0.1))['90%']

W_1 = matrix(rnorm(132, 0, 1),  nrow = 22, ncol = 6)

quantile(abs(as.numeric(Psi_B%*%W_1)), probs = seq(0,1, 0.01))['90%']

# Verify that 90% of value of Psi_B*W(1) are < 0.02
```

C.3. Parametrizing and Debiasing Expert Forecasts

This part of the code aims to parametrize the expert forecasts and to address the effect of behavioral biases on expert views through a process called debiasing.

```
#parametrizing the drift parameters a and A for expert forecasts
az_expert = matrix(0, k)

Az_expert = rbind(c(1,1,rep(0,4)),
              c(0,0,1,1,0,0),
              c(rep(0,4),1,1)) %>% t()


confidence_levels = seq(from = 0.3, to = 0.9, by = 0.1)
sampled_levels = paste0(sample(confidence_levels, k)*100, '%')

# the paper keeps calling Lambda a matrix that is now different
Lambda_2 = cbind(lambda, matrix(0, n,k))

# sd of risk factors actual values
sigma_1 = diag(apply(all_data[risk_factors_names], 2, sd))
sigma_2 = diag(sampled_lb, k, k)

dim(sigma_1) == c(n,n)
dim(sigma_2) == c(k,k)
```

```
R11 = cor(all_data[risk_factors_names])
R22 = cor(Z)
R12 = cor(all_data[risk_factors_names],Z)


# #overconfidence: solve overconfidence
# Z_overconfidence = matrix(NA, t, k)
#
# for(i in 1:k){
#   print(i)
#   Z_overconfidence[,i] =
#     create_expert_forecast(quantile(Z[, i], seq(0,1, by = 0.1))[sampled_levels[i]],
#                            0.11,
#                            expert_central_view[i], 209)
#
# }
# Z_overconfidence[is.na(Z_overconfidence)] = 0.1
#
# psi_Z = t(Z_overconfidence)%*%Z_overconfidence
#
# Psi_Z = diag(sampled_lb, K, n+m+K)


# Debias from overconfidence by increasing the magnitude of Psi_B by 20%
Psi_Z = 1.2*Psi_B

dim(R11) == c(n,n)
dim(R22) == c(k,k)

# Derive psi_n
psi_n = sigma_2%*%t(R12)%*%sigma_1%*%solve(t(lambda))

dim(psi_n) == c(k,n)


# Stress Test Scenario
a_z_ST = matrix(0, 3)
A_z_ST = diag(1, 3)
zero_matrix_ST = matrix(0,n,n+m)

# stress test data (covid financial crisis)
ST_data = all_data %>% filter(date >= '2020-04-01', date <= '2021-03-01' )

X_ST = ST_data[risk_factors_names]
quadratic_variation_X_ST = (t(diff(ts(X_ST))) %*% diff(ts(X_ST)))

psi_ST = svd(quadratic_variation_X_ST)$u
Psi_ST = cbind(matrix(0, k_ST, n+m), psi_ST)
```

C.4. Kalman Filter

This section provides the computation of the Kalman filter. The next sections present its implementation for both the BLCT and BB models.

```
library(yuima)

### simulation of securities price

x_helper_s = paste0('*x', 1:m)
drift_character_s = character(m)

SSt = Sigma%*%t(Sigma)
diag_Sigma = diag(SSt)



for(i in 1:m){
  drift_character_s[i] = paste0('(',a[i], '-0.5*', diag_Sigma[i], ')'  ,'+', paste0(paste0(A[i,],
 x_helper_s), collapse = '+'))
}

diffusion_character_s = apply(Sigma, 2, as.character)


sol = c(paste0('x', 1:m))
mod_s = setModel(drift = drift_character_s, diffusion = diffusion_character_s, solve.variable
 = sol)

set.seed(123)
s_sim = simulate(mod_s,
              sampling = setSampling(Initial = 0, Terminal = 1, n = 208),
              xinit = s[1,])

sim_data = apply(s_sim@data@original.data, 2, as.numeric) %>% data.frame() %>%
 as_tibble()
colnames(sim_data) = paste0(colnames(s), '')
securities_sim = tibble(all_data[1:209,1], sim_data)
stats::plot.ts(securities_sim[-c(1:4)],
            main = 'Securities simulation according to SDE 3.3')

stats::plot.ts(securities_sim[c(2:4)],
            main = 'Securities simulation according to SDE 3.3')

#Real values
# stats::plot.ts(ts_all_data[,4:13],
#              main = 'Securities Real Valus (not log)')


# Simulated expert forecasts using Psi_Z as diffusion matrix (debiased)
```

```
x_helper_Z = paste0('*x', 1:k)
drift_character_Z = character(k)

for(i in 1:k){
  drift_character_Z[i] = paste0(az_expert[i], '+', paste0(paste0(Az_expert[i,], x_helper_Z),
 collapse = '+'))
}

diffusion_character_Z = apply(Psi_Z, 2, as.character)

sol = c(paste0('z', 1:k))
mod_Z = setModel(drift = drift_character_Z, diffusion = diffusion_character_Z,
              solve.variable = sol,
              xinit = 0)

set.seed(123)
Z_sim = simulate(mod_Z, sampling = setSampling(Initial = 0, Terminal = 1, n = 208))
Zsim_data = apply(Z_sim@data@original.data, 2, as.numeric) %>% data.frame() %>%
 as_tibble()
expert_sim = tibble(all_data[1:209,1], Zsim_data)
colnames(expert_sim)[-1] = paste0('z', 1:K)
plot.ts(expert_sim[-1], main='Simulated Expert Forecasts from SDE', xlab = 'Time')


# Simulated expert forecasts using Psi_B as diffusion matrix (biased)

x_helper_Zbias = paste0('*x', 1:k)
drift_character_Zbias = character(k)

for(i in 1:k){
  drift_character_Zbias[i] = paste0(az_expert[i], '+', paste0(paste0(Az_expert[i,],
 x_helper_Zbias), collapse = '+'))
}

diffusion_character_Zbias = apply(Psi_B, 2, as.character)

sol = c(paste0('z', 1:k))
mod_Zbias = setModel(drift = drift_character_Zbias, diffusion = diffusion_character_Zbias,
              solve.variable = sol,
              xinit = 0)

set.seed(123)
Zbias_sim = simulate(mod_Zbias, sampling = setSampling(Initial = 0, Terminal = 1, n =
 208))
Zbias_sim_data = apply(Z_sim@data@original.data, 2, as.numeric) %>% data.frame() %>%
 as_tibble()
expert_bias_sim = tibble(all_data[1:209,1], Zbias_sim_data)
colnames(expert_bias_sim)[-1] = paste0('z', 1:K)
plot.ts(expert_bias_sim[-1], main='Simulated Expert Forecasts (Bias) from SDE',
      xlab = 'Time')
```

```
#Risk Factors simulation
set.seed(123)
starting_points_X = MASS::mvrnorm(1, mu_0, P_0)

x_helper_X = paste0('*x', 1:n)
drift_character_X = character(n)

for(i in 1:n){
  drift_character_X[i] = paste0(b[i], '+', paste0(paste0(B[i,], x_helper_X), collapse = '+'))
}

diffusion_character_X = apply(Lambda, 2, as.character)

sol = c(paste0('x', 1:n))
mod_X = setModel(drift = drift_character_X, diffusion = diffusion_character_X,
 solve.variable = sol)

set.seed(123)
X_sim = simulate(mod_X, xinit = starting_points_X, sampling = setSampling(n= 208))
plot(X_sim, ylab = c('rf1', 'rf2', 'rf3'),
     main = 'Risk factor simulation from SDE', xlab = 'Time')


#Kalman
#stacking data rows (s and z) (13 dim for s and 6 for z)
y1=t(matrix(c(s_sim@data@zoo.data$`Series 1`)))
y2=t(matrix(c(s_sim@data@zoo.data$`Series 2`)))
y3=t(matrix(c(s_sim@data@zoo.data$`Series 3`)))
y4=t(matrix(c(s_sim@data@zoo.data$`Series 4`)))
y5=t(matrix(c(s_sim@data@zoo.data$`Series 5`)))
y6=t(matrix(c(s_sim@data@zoo.data$`Series 6`)))
y7=t(matrix(c(s_sim@data@zoo.data$`Series 7`)))
y8=t(matrix(c(s_sim@data@zoo.data$`Series 8`)))
y9=t(matrix(c(s_sim@data@zoo.data$`Series 9`)))
y10=t(matrix(c(s_sim@data@zoo.data$`Series 10`)))
y11=t(matrix(c(s_sim@data@zoo.data$`Series 11`)))
y12=t(matrix(c(s_sim@data@zoo.data$`Series 12`)))
y13=t(matrix(c(s_sim@data@zoo.data$`Series 13`)))

BIAS = F

y14 = if_else(BIAS == F,
          t(matrix(c(Z_sim@data@zoo.data$`Series 1`))),
          t(matrix(c(Zbias_sim@data@zoo.data$`Series 1`))))

y15 = if_else(BIAS == F,
          t(matrix(c(Z_sim@data@zoo.data$`Series 2`))),
          t(matrix(c(Zbias_sim@data@zoo.data$`Series 2`))))


y16 = if_else(BIAS == F,
          t(matrix(c(Z_sim@data@zoo.data$`Series 3`))),
```

```
          t(matrix(c(Zbias_sim@data@zoo.data$`Series 3`))))

y17 = if_else(BIAS == F,
          t(matrix(c(Z_sim@data@zoo.data$`Series 4`))),
          t(matrix(c(Zbias_sim@data@zoo.data$`Series 4`))))

y18 = if_else(BIAS == F,
          t(matrix(c(Z_sim@data@zoo.data$`Series 5`))),
          t(matrix(c(Zbias_sim@data@zoo.data$`Series 5`))))

y19 = if_else(BIAS == F,
          t(matrix(c(Z_sim@data@zoo.data$`Series 6`))),
          t(matrix(c(Zbias_sim@data@zoo.data$`Series 6`))))
yy = rbind(y1,y2,y3,y4,y5,y6,y7,y8,y9,y10,y11,y12,y1,y14,y15,y16,y17,y18,y19)
yy_d = matrix(nrow = 19, ncol = 208)
del_t = 1/208;

#data_matrix from A.3
for(ii in 1:208){
  yy_d[,ii] = yy[,ii+1] - yy[,ii]
  yy_d[,ii]/del_t
}

#defining variables for FKF library

#from A.1 discretizing state equation
dt = matrix(c(b))*del_t
Tt = diag(3) + matrix(c(B), nrow = 3, ncol = 3)*del_t

Zt = rbind(matrix(c(A), ncol = 3), Az_expert)
ct = rbind(((a[1:13])-.5*matrix(diag_Sigma)), az_expert)
GGt=0.01*diag(19)
HHt=0.01*diag(3)

a0 = (c(mu_0))
p0 = matrix(5*P_0, nrow = 3)

# apply Kalman Filter
library(FKF)
fkf_obj = fkf(a0, p0, dt, ct, Tt, Zt, HHt, GGt, yy_d)

#x_hat=fkf_obj$att
x11 = X_sim@data@zoo.data$`Series 1`
par(mfrow=c(3,1))

#tracking
plot(fkf_obj$att[1, ], col = 'green', ylim = c(-1,1), type = 'l')
lines((1:208), x11[2:209], col = 'black',ylim = c(-1,1))
#lines(ts(fkf.obj$att[1, ], start = start(yy_d), frequency = frequency(yy_d)), col = "blue")

x12 = X_sim@data@zoo.data$`Series 2`
```

```
#tracking
plot(fkf_obj$att[2, ], col = 'green', ylim = c(-1,1), type = 'l')
lines((1:208), x12[2:209], col = 'black', ylim = c(-1,1))

x13 = X_sim@data@zoo.data$`Series 3`
#tracking
plot(fkf_obj$att[3, ], col = 'green', ylim = c(-1,1), type = 'l')
lines((1:208), x13[2:209],col='black', ylim = c(-1,1))

#Kalman results in fkf_.obj

X_hat = fkf_obj$att[1:3, ]

plot.ts(ts(t(X_hat)), order_by(all_data$date), main = 'KM X_hat')

# Obtain Sigma_hat
quadratic_variation_sX_hat = (t(diff(s)) %*% rbind(rep(0,3),diff(t(X_hat))))

#sigma_n_hat = (quadratic_variation_sX_hat) %*% lambda #m x n

sigma_m_hat_helper = quadratic_variation_s - quadratic_variation_sX %*%
 solve(quadratic_variation_X) %*% t(quadratic_variation_sX_hat)
sigma_m_hat = chol(sigma_m_hat_helper)

(Sigma_hat = cbind(sigma_m_hat, matrix(0, m, k)))

dim(Sigma_hat) == c(m, m+k)
```

C.5. BLCT Model Implementation

This part of the code presents the implementation of the BLCT model where expert forecasts are not debiased.

```
# Simulate securities price (needed for kalman filter only)

x_helper_s = paste0('*x', 1:m)
drift_character_s = character(m)

# SSt_hat = Sigma_hat%*%t(Sigma_hat)
# diag_Sigma_hat = diag(SSt_hat)


set.seed(123)
for(i in 1:m){
  drift_character_s[i] = paste0('(',a[i], '-0.5*', diag_Sigma[i], ')'  ,'+', paste0(paste0(A[i,],
 x_helper_s), collapse = '+'))
}

diffusion_character_s = apply(Sigma, 2, as.character)
```

```
sol = c(paste0('x', 1:m))
mod_s = setModel(drift = drift_character_s, diffusion = diffusion_character_s, solve.variable
 = sol)


s_sim = simulate(mod_s,
            sampling = setSampling(Initial = 0, Terminal = 1, n = 208),
            xinit = s[1,])

sim_data = apply(s_sim@data@original.data, 2, as.numeric) %>% data.frame() %>%
 as_tibble()
colnames(sim_data) = paste0(colnames(s), '_sim')

### End simulation


#Kalman
#stacking data rows (s and z) (13 dim for s and 6 for z)
y1=t(matrix(c(s_sim@data@zoo.data$`Series 1`)))
y2=t(matrix(c(s_sim@data@zoo.data$`Series 2`)))
y3=t(matrix(c(s_sim@data@zoo.data$`Series 3`)))
y4=t(matrix(c(s_sim@data@zoo.data$`Series 4`)))
y5=t(matrix(c(s_sim@data@zoo.data$`Series 5`)))
y6=t(matrix(c(s_sim@data@zoo.data$`Series 6`)))
y7=t(matrix(c(s_sim@data@zoo.data$`Series 7`)))
y8=t(matrix(c(s_sim@data@zoo.data$`Series 8`)))
y9=t(matrix(c(s_sim@data@zoo.data$`Series 9`)))
y10=t(matrix(c(s_sim@data@zoo.data$`Series 10`)))
y11=t(matrix(c(s_sim@data@zoo.data$`Series 11`)))
y12=t(matrix(c(s_sim@data@zoo.data$`Series 12`)))
y13=t(matrix(c(s_sim@data@zoo.data$`Series 13`)))

BIAS = T

y14 = if_else(BIAS == F,
          t(matrix(c(Z_sim@data@zoo.data$`Series 1`))),
          t(matrix(c(Zbias_sim@data@zoo.data$`Series 1`))))

y15 = if_else(BIAS == F,
          t(matrix(c(Z_sim@data@zoo.data$`Series 2`))),
          t(matrix(c(Zbias_sim@data@zoo.data$`Series 2`))))


y16 = if_else(BIAS == F,
          t(matrix(c(Z_sim@data@zoo.data$`Series 3`))),
          t(matrix(c(Zbias_sim@data@zoo.data$`Series 3`))))

y17 = if_else(BIAS == F,
          t(matrix(c(Z_sim@data@zoo.data$`Series 4`))),
          t(matrix(c(Zbias_sim@data@zoo.data$`Series 4`))))
```

```
y18 = if_else(BIAS == F,
          t(matrix(c(Z_sim@data@zoo.data$`Series 5`))),
          t(matrix(c(Zbias_sim@data@zoo.data$`Series 5`))))

y19 = if_else(BIAS == F,
          t(matrix(c(Z_sim@data@zoo.data$`Series 6`))),
          t(matrix(c(Zbias_sim@data@zoo.data$`Series 6`))))


yy = rbind(y1,y2,y3,y4,y5,y6,y7,y8,y9,y10,y11,y12,y1,y14,y15,y16,y17,y18,y19)
yy_d = matrix(nrow = 19, ncol = 208)
del_t = 1/208;

#data_matrix from A.3
for(ii in 1:208){
  yy_d[,ii] = yy[,ii+1] - yy[,ii]
  yy_d[,ii]/del_t
}

#defining variables for FKF library

#from A.1 discretizing state equation
dt = matrix(c(b))*del_t
Tt = diag(3) + matrix(c(B), nrow = 3, ncol = 3)*del_t

Zt = rbind(matrix(c(A), ncol = 3), Az_expert)
ct = rbind(((a[1:13])-.5*matrix(diag_Sigma)), az_expert)
GGt=0.01*diag(19)
HHt=0.01*diag(3)

a0 = (c(mu_0))
p0 = matrix(5*P_0, nrow = 3)

# apply Kalman Filter
library(FKF)
fkf_obj = fkf(a0, p0, dt, ct, Tt, Zt, HHt, GGt, yy_d)

# #x_hat=fkf_obj$att
# x11 = X_sim@data@zoo.data$`Series 1`
# par(mfrow=c(3,1))
#
# #tracking
# plot(fkf_obj$att[1, ], col = 'green', ylim = c(-1,1), type = 'l')
# lines((1:208), x11[2:209], col = 'black',ylim = c(-1,1))
# #lines(ts(fkf.obj$att[1, ], start = start(yy_d), frequency = frequency(yy_d)), col = "blue")
#
# x12 = X_sim@data@zoo.data$`Series 2`
#
# #tracking
# plot(fkf_obj$att[2, ], col = 'green', ylim = c(-1,1), type = 'l')
# lines((1:208), x12[2:209], col = 'black', ylim = c(-1,1))
#
```

```
# x13 = X_sim@data@zoo.data$`Series 3`
# #tracking
# plot(fkf_obj$att[3, ], col = 'green', ylim = c(-1,1), type = 'l')
# lines((1:208), x13[2:209],col='black', ylim = c(-1,1))

#Kalman results in fkf_.obj

X_hat_bias = fkf_obj$att[1:3, ]


h_K = solve(Sigma_hat%*%t(Sigma_hat))%*%(a+(A%*%X_hat_bias))

#h_PIHP = solve(SSt_hat)%*%(Sigma_hat%*%Lambda_hat[, 1:13])


# portfolio allocation strategy h*_t using fractional kelly
theta = 0

f = 1/(theta +1)

h_hat = f*h_K


#simulate portfolios (sostituendo X con X_hat (output Kalman Filter) si ottiene la simulazione
 BBLX)
# x_helper_s = paste0('*x', 1:m)
# drift_character_s = character(m)
#
# SSt = Sigma%*%t(Sigma)
# diag_Sigma = diag(SSt)



# for(i in 1:m){
#   drift_character_s[i] = paste0('(',a[i], '-0.5*', diag_Sigma[i], ')'  ,'+', paste0(paste0(A[i,],
 x_helper_s), collapse = '+'))
# }
#
# diffusion_character_s = apply(Sigma, 2, as.character)
#
#
# sol = c(paste0('x', 1:m))
# mod_s = setModel(drift = drift_character_s, diffusion = diffusion_character_s,
 solve.variable = sol)
#

library(PerformanceAnalytics)

W0 = 10000
pt_sim_list = list()
avg_return = as.numeric()
sd_return = c()
```

```
VaR = c()
CVaR = c()
sharpe = c()
sortino = c()

set.seed(123)
nsim = 3000
for(i in 1:nsim){

 #derive weights for the securities
 # weights_helper = runif(m)
 # weights = weights_helper/sum(weights_helper)
 dim(h_K)
 weights = t(apply(t(h_K), 1, function(x) x/sum(x)))


 #simulate one trajectories for the 13 securities over 5 years
 s_sim = simulate(mod_s,
                  sampling = setSampling(Initial = 0, Terminal = 1, n = 207),
                  xinit = s[1,])

 sim_data = apply(s_sim@data@original.data, 2, as.numeric) %>% data.frame() %>%
 as_tibble()
 #portfolio_trajectory = as.matrix(sim_data)%*%matrix(weights)
 portfolio_trajectory = apply(as.matrix(sim_data)*(weights), 1, sum)
 pt_sim_list[[i]] = exp(portfolio_trajectory)*(W0/exp(portfolio_trajectory)[1]) # transform to
 real prices
 avg_return[i] = mean(diff(portfolio_trajectory))
 sd_return[i] = sd(diff(portfolio_trajectory))
 VaR[i] = -VaR(as.vector(diff(portfolio_trajectory)), p = 0.99)
 CVaR[i] = -CVaR(as.vector(diff(portfolio_trajectory)), p = 0.99)

 diff_zoo_pt = as.zoo(diff(portfolio_trajectory), order.by = all_data$date[-1])
 sharpe[i] = SharpeRatio(diff_zoo_pt)[1]
 sortino[i] = SortinoRatio(diff_zoo_pt)

}

plot(pt_sim_list[[1]], type = 'l')

tbl_result_helper = tibble(avg_return, sd_return, VaR, CVaR, sharpe, sortino)

tbl_result_BLCT = tibble(measure = names(tbl_result_helper), BLCT =
 apply(tbl_result_helper, 2, mean))

zoo_pt = as.zoo(portfolio_trajectory, order.by = all_data$date)

trajectories_tibble = bind_cols(pt_sim_list) %>% as.data.frame()
colnames(trajectories_tibble) = paste0('sec_', 1:10)


zoo_trajectories = as.zoo(trajectories_tibble[1:10], order.by = all_data$date + years(5))
```

```
tsRainbow <- rainbow(10)

# Plot the overlayed series
plot(x = zoo_trajectories[1:104,], ylab = "Portfolio Value", main = "Simulation of 10
 trajectories of the BLCT",
    col = tsRainbow, screens = 1, lwd = 2, xlab = 'Time')

tbl_result_BLCT
```

C.6. BB Model Implementation

The last part of the code presents the implementation of the BB model where expert forecasts are debiased.

```
h_K = solve(Sigma_hat%*%t(Sigma_hat))%*%(a+(A%*%X_hat))

#h_PIHP = solve(SSt_hat)%*%(Sigma_hat%*%Lambda_hat[, 1:13])


# portfolio allocation strategy h*_t using fractional kelly
theta = 0

f = 1/(theta +1)

h_hat = f*h_K


#simulate portfolios (sostituendo X con X_hat (output Kalman Filter) si ottiene la simulazione
 BBLX)
x_helper_s = paste0('*x', 1:m)
drift_character_s = character(m)

SSt = Sigma%*%t(Sigma)
diag_Sigma = diag(SSt)



for(i in 1:m){
  drift_character_s[i] = paste0('(',a[i], '-0.5*', diag_Sigma[i], ')'  ,'+', paste0(paste0(A[i,],
 x_helper_s), collapse = '+'))
}

diffusion_character_s = apply(Sigma, 2, as.character)


sol = c(paste0('x', 1:m))
mod_s = setModel(drift = drift_character_s, diffusion = diffusion_character_s, solve.variable
 = sol)
```

```
library(PerformanceAnalytics)

W0 = 10000
pt_sim_list = list()
avg_return = as.numeric()
sd_return = c()
VaR = c()
CVaR = c()
sharpe = c()
sortino = c()

set.seed(123)
nsim = 3000
for(i in 1:nsim){

  #derive weights for the securities
  # weights_helper = runif(m)
  # weights = weights_helper/sum(weights_helper)
  dim(h_K)
  weights = t(apply(t(h_K), 1, function(x) x/sum(x)))


  #simulate one trajectories for the 13 securities over 5 years
  s_sim = simulate(mod_s,
                sampling = setSampling(Initial = 0, Terminal = 1, n = 207),
                xinit = s[1,])

  sim_data = apply(s_sim@data@original.data, 2, as.numeric) %>% data.frame() %>%
 as_tibble()
  #portfolio_trajectory = as.matrix(sim_data)%*%matrix(weights)
  portfolio_trajectory = apply(as.matrix(sim_data)*(weights), 1, sum)
  pt_sim_list[[i]] = exp(portfolio_trajectory)*(W0/exp(portfolio_trajectory)[1]) # transform to
 real prices
  avg_return[i] = mean(diff(portfolio_trajectory))
  sd_return[i] = sd(diff(portfolio_trajectory))
  VaR[i] = -VaR(as.vector(diff(portfolio_trajectory)), p = 0.99)
  CVaR[i] = -CVaR(as.vector(diff(portfolio_trajectory)), p = 0.99)

  diff_zoo_pt = as.zoo(diff(portfolio_trajectory), order.by = all_data$date[-1])
  sharpe[i] = SharpeRatio(diff_zoo_pt)[1]
  sortino[i] = SortinoRatio(diff_zoo_pt)

}

plot(pt_sim_list[[10]], type = 'l')

tbl_result_helper = tibble(avg_return, sd_return, VaR, CVaR, sharpe, sortino)

tbl_result_BBLX = tibble(measure = names(tbl_result_helper), BBLX =
 apply(tbl_result_helper, 2, mean))
```

```
zoo_pt = as.zoo(portfolio_trajectory, order.by = all_data$date)

trajectories_tibble = bind_cols(pt_sim_list) %>% as.data.frame()
colnames(trajectories_tibble) = paste0('sec_', 1:nsim)


zoo_trajectories = as.zoo(trajectories_tibble[, 1:10], order.by = all_data$date + years(5))

tsRainbow <- rainbow(ncol(zoo_trajectories))

# Plot the overlayed series
plot(x = zoo_trajectories[1:104], ylab = "Portfolio Value", main = "Simulation of 10
 trajectories of the BBL - X",
    col = tsRainbow, screens = 1, lwd = 2, xlab = 'Time')

tbl_result_BBLX
```

References

BERNOULLI, D., 1954. *Exposition of a new theory on the measurement of risk*. Econometric. 22:22–36.

BIELECKI, T.R., HERNANDEZ-HERNANDEZ D., PLISKA, S.R., 2002. *Recent Developments in Mathematical Finance*. Singapore: World Scentific. 127-138

BIELECKI, T.R., PLISKA, S.R., 1999. *Risk-sensitive dynamic asset management. Applied Mathematics and Optimization*. 39:337–360.

BIELECKI, T.R., PLISKA, S.R., 2000. *Risk sensitive asset management with transaction costs. Finance and Stochastics*. 4:1–33.

BIELECKI, T.R., PLISKA, S.R., 2003. *Economic properties of the risk sensitive criterion for portfolio management. The Review of Accounting and Finance*. 2(2):3–17.

BIELECKI, T.R., PLISKA, S.R., 2004. *Risk sensitive intertemporal CAPM. IEEE Transactions on Automatic Control*. 49(3):420–432.

BIELECKI, T.R., PLISKA, S.R., SHEU, S.J., 2005. *Risk sensitive portfolio management with Cox-Ingersoll-Ross interest rates: the HJB equation*. SIAM Journal of Control and Optimization. 44:1811–1843.

BREIMAN L., 1961. *Optimal gambling system for favorable games*. Los Angeles: University of California. 63–68.

COSTA, D.F., DE MELO CARVALHO, F., DE MELO MOREIRA, B.C. et al., 2017. *Bibliometric analysis on the association between behavioral finance and decision making with cognitive biases such as overconfidence, anchoring effect and confirmation bias.* Scientometrics. 111, 1775–1799.

DAVIS, M., LLEO, S., 2008. *Risk-sensitive benchmarked asset management*. Quantitative Finance. 8(4):415–426.

DAVIS, M., LLEO, S., 2009. *Jump-Diffusion Risk-Sensitive Asset Management.*

DAVIS, M., LLEO, S., 2013. *Black–Litterman in continuous time: the case for filtering.* Quantitative Finance Letters. 1:1, 30-35.

DAVIS, M., LLEO, S., 2016. *A simple procedure for combining expert opinion with statistical estimates to achieve superior portfolio performance.* Journal of Portfolio Management. 42(4), 38–48.

DAVIS, M., LLEO, S., 2020. *Debiased expert forecasts in continuous-time asset allocation.* Journal of Banking & Finance. 113, 105759.

DOWD K., 2005. *Measuring Market Risk*. John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England.

FREY, R., GABIH, A., WUNDERLICH, R., 2012. *Portfolio optimization under partial information with expert opinions.* International Journal of Theoretical and Applied Finance. 15(1).

FREY, R., GABIH, A., WUNDERLICH, R., 2014. *Portfolio optimization under partial information with expert opinions: A dynamic programming approach.* Communications on Stochastic Analysis , Vol. 8, No. 1. 49-79.

GABIH, A., KONDAKJI, H., SASS, J., WUNDERLICH, R., 2014. *Expert opinions and logarithmic utility maximization in a market with gaussian drift.* Communications on Stochastic Analysis. 8(1), 27–47.

KELLY J., 1956. *A new interpretation of the information rate.* Bell System Technical Journal, 35:917–926.

KURODA K., NAGAI H., 2002. *Risk-sensitive portfolio optimization on infinite time horizon*. Stochastics and Stochastics Reports, 73:309–331.

LATANÉ H., 1959. *Criteria for choice among risky ventures.* Journal of Political Economy, 67:144–155, 1959.

LINTNER J., 1965. *The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets*. Review of Economics and Statistics, 47(1):13–37.

MARKOWITZ H., 1976. *Investment for the long run: New evidence for an old rule.* Journal of Finance, 36(5):495–508.

MACLEAN L., ZIEMBA W.T., LI Y., 2005. *Time to wealth goals in cap- ital accumulation and the optimal trade-off of growth vesrus security.* Quantitative Finance, 5(4):343–357.

MACLEAN L.C., SANEGRE R., ZAO Y., ZIEMBA W.T., 2004. *Capital growth with security.* Journal of Economic Dynamics and Control, 28(5):937– 954.

MACLEAN L., THORP E., ZIEMBA W.T., 2010. *Long-term capital growth: The good and bad properties of the kelly criterion*. Quantitative Finance, 10(7):681–687.

MACLEAN L.C., ZIEMBA W.T., BLAZENKO G., 1992. *Growth versus security in dynamic investment analysis.* Management Science, 38(11):1562– 1585.

SAMUELSON P., 1969. *Lifetime portfolio selection by dynamic stochastic pro-gramming*. Review of Economics and Statistics, 51(3):239–246.

SAMUELSON P., 1979. *Why we should not make mean log of wealth big though years to act are long*. Journal of Banking and Finance, 3:305–307.

SAMUELSON P.A., 1971. *The "fallacy" of maximizing the geometric mean in long sequences of investing or gambling.* In Proceedings of the National Academy of Sciences of the United States of America, volume 68, pages 2493–2496, 1971.

SASS, J., WESTPHAL, D., WUNDERLICH, R., 2017. *Expert opinions and logarithmic utility maximization for multivariate stock returns with gaussian drift*. International Journal of Theoretical and Applied Finance, 20(04), 1750022.

SASS, J., WESTPHAL, D., WUNDERLICH, R., 2021. *Diffusion approximations for expert opinions in a financial market with gaussian drift.* Journal of Applied Probability, 58(1), 197–216.

SHARPE, W., 1964. *Capital asset prices: A theory of market equilibrium under conditions of risk*. Journal of Finance, 19(3):425–442.

THORP E., 1971. *Portfolio choice and the kelly criterion.* In Proceedings of the Business and Economics Section of the American Statistical Association. 215–224.

THORP E., 2006. *The kelly criterion in blackjack, sports betting and the stock market.* In S.A. Zenios and W.T. Ziemba, editors, Handbook of Asset and Liability Management, volume 1 of Handbook in Finance.

WELCH, G., BISHOP, G., 2001. *An introduction to the Kalman filter.* University of North Carolina at Chapel Hill.

ZIEMBA W.T. (2005). *The symmetric downside-risk sharpe ratio.* Journal of Portfolio Management, 32(1):108–122.

ZIEMBA R., ZIEMBA W.T., 2007. *Scenarios for Risk Management and Global Investment Strategies.* John Wiley & Sons, 2007