Università degli Studi di Padova

Dipartimento di Scienze Statistiche
Dipartimento di Scienze Economiche

Corso di Laurea Magistrale in
Scienze Statistiche

INVESTING FOR THE LONG RUN: PREDICOR VARIABLES AND LOSS AVERSION

Relatore: Ch.mo Prof. Massimiliano Caporin
Dipartimento di Scienze Economiche

Laureando: Alessandra Porra

Anno Accademico 2012/2013

Contents

Introduction ... vi
1. Short run portfolio allocation ... 1
1.1Introduction ... 1
1.1 Financial market returns from 1802 ... 1
1.2 Asset returns ... 4
1.2.1 Portfolio returns ... 6
1.2.2 Excess returns and risk-free asset ... 7
1.3 Expected Utility Theory ... 9
1.4 Mean-Variance Analysis ... 12
1.4.1 The form of the utility function ... 16
1.4.2 Limitations of the Mean-Variance Model ... 17
1.5 The holding period ... 19
1.5.1 Long-run portfolio choice ... 20
2. Portfolio allocation with parameter uncertainty ... 22
2.1 Introduction ... 22
2.2 Parameter uncertainty ... 22
2.3 Data set ... 24
2.3.1 Preliminary analysis ... 25
2.4 Long horizon portfolio allocation ... 32
2.4.1 Sampling process ... 36
2.5 Results ... 37
2.6 Resampling ... 41
2.6.1 Results ... 42
3 Portfolio allocation with predictable returns ... 44
3.1 Introduction ... 44
3.2 Returns predictability ... 44
3.3 Predictor variable: dividend yield ... 45
3.3.1 Preliminary Analysis ... 46
3.4 Long horizon predictability and parameter uncertainty ... 49
3.5 Predictability analysis model ... 51

3.6 Long horizon portfolio allocation 53
3.6.1 Sampling process 57
3.7 Results 58
3.8 The role of the predictor variable 64
4 Portfolio allocation with parameter uncertainty: two risky assets 67
4.1 Introduction 67
4.2 An extra risky asset: the bond index 67
4.2.1 Preliminary analysis 69
4.3 Model with two risky assets 72
4.4 Long horizon portfolio allocation 74
4.4.1 Sampling process 78
4.5 Results 79
5 Portfolio allocation with predictable returns and five predictor variables 84
5.1 Introduction 84
5.2 Stock and bond predictability 84
5.3 Predictive variables 85
5.3.1 Vix index 86
5.3.2 Term spread 89
5.3.3 Credit spread 92
5.3.4 Risk-free asset 95
5.3 Predictability analysis model 95
5.4 Results 98
5.5 The role of the predictor variables 105
5.6 Other samples results 108
5.6.1. Sample 1990-2000 108
5.6.2. Sample 2002-2006 109
5.6.3 Sample 2007-2012 111
6 Portfolio allocation under loss aversion 114
6.1 Introduction 114
6.2 Critiques to the Expected Utility theory 114
6.3 Behavioral Finance 116
6.4 Prospect theory 118
6.5 Long horizon asset allocation 121
6.6 Results 123
Conclusions 127

Appendix A 130
Sample 1990 – 2000 130
Sample 2002 – 2006 132
Sample 2006 - 2012 134
Appendix B 136
Bibliography 153

Introduction

Portfolio choice problems are the leading edge of financial research. The portfolio theory underlying an investor's optimal portfolio choice, pioneered by Markowitz's

Mean-Variance Anlysis (1952), is by now well comprehended. The reborn interest in portfolio choice problems follows the relatively recent empirical evidence of time-varying return distributions (predictability and conditional heteroskedasticity). The purpose of this work is indeed to examine the effects of predictability for an investor trying to take portfolio allocation decisions. According to Samuelson (1969) and Merton (1969), when asset returns are i.i.d., an investor who rebalances his portfolio optimally and whose preferences are described by a power utility function, should choose the same asset allocation regardless of the investment horizon. However, considering the growing evidence of predictability in returns, the investor's horizon may no longer be unimportant. We therefore address this issue of portfolio choice from the perspective of horizon effects**:** "*Given the demonstration of predictability in asset returns, should a long horizon investor allocate his wealth differently from a short-horizon investor?"* (Barberis, 2000)

Our work draws on Nicholas Barberis' paper (2000) about long run predictability of asset returns. In his work he studies the effects of predictability for an investor making sensible portfolio choices. He analyzes portfolio choice in discrete time for an investor with power utility function over terminal wealth, employing two assets: a stock index and a risk-free asset. In order to examine how predictability affects portfolio choices he compares the allocation of an investor who does not recognize predictability, that is when asset returns are described by a i.i.d. model, to that of an investor who takes predictability into account. In particular he uses only one predictor variable in order to describe asset returns' dynamics, the dividend yield. He finds that predictability in asset returns leads to strong horizon effects, involving a much higher allocation to stocks for a long-horizon investor than for a short-horizon investor, this being because predictability makes stocks look less risky at long horizons.

In our work we try to understand if a risk averse investor, who decides today how to invest his wealth and does not change the allocation until the predetermined maturity, distributes his wealth differently for long horizons compared to short horizons. We firstly focus on studying the predictive power of only one variable, the dividend yield, for stock returns. Afterwards, we devote most of our work to examining in what way the optimal portfolio allocation changes when investors have the opportunity to choose how to allocate their wealth among three different assets, instead of the previous two: a stock index, a bond index, and a risk-free asset. We then investigate the predictability of excess stock and bond returns, availing ourselves of a set of five predictor variables gathered from the financial literature.

Particular attention is paid to estimation risk, which can be defined as the uncertainty about the true values of model parameters. We analyze estimation risk in order to take into account the uncertainty about the true predictive power of predictor variables, that sometimes could be weak. This approach constitutes therefore a middle ground between rejecting the null hypothesis of returns predictability, and analyzing the problem taking the parameters as fixed and known precisely.

In addition to what Barberis handled in his paper, we then devote our attention to introducing an alternative method to the Expected Utility approach, that is the Prospect Theory developed by Kahneman and Tversky (1979), whose goal is to capture people's attitudes to risky gambles as parsimoniously as possible. According to this theory a value function replaces the usual utility function, in particular the loss aversion utility function explains the investors' behavior of being risk averse for gains and risk seeking for losses. Moreover it describes the principle of loss aversion, according to which losses loom larger than corresponding gains. Our purpose is therefore to examine how the optimal portfolio allocation changes depending on whether the function employed to describe investors' preferences over wealth is a power utility function or a loss aversion function.

Regarding the application we evaluate a vector autoregressive model in order to explain the time-variation in asset returns throughout the predictor variables. Afterwards uncertainty about the model parameters is incorporated by the posterior distribution of the parameters given the data

The purpose of the first chapter is to explain in detail some concepts and ideas used throughout the work. After a brief description of financial markets returns over the last two centuries, we define the notions of asset return, excess return and risk-free rate.
The Expected Utility Theory and Mean-Variance Analysis are then illustrated. Finally we consider the case handled by Samuelson and Merton, when long-term

investors act myopically, choosing the same portfolio as short-term investors, and we specify the main approaches an investor can adopt.

In the second chapter attention is paid to the estimation risk, in other words we study the optimal portfolio allocation assuming that parameters are not known precisely. Our purpose is to understand how parameter uncertainty alone affects portfolio choice. According to a Bayesian approach, we define the posterior distribution of the model parameters given the data, and integrating over the uncertainty in the parameters captured by the posterior distribution, we construct predictive distribution for future returns, conditional only on observed data, and not on any fixed parameter value. The model implemented is then applied to a real dataset. Finally we illustrate the results obtained both assuming that excess returns have a normal distribution and adopting a resampling approach in order to understand if the assumption of normality attributed to assets returns affects the optimal portfolio allocation.

The third chapter focuses on how predictability affects portfolio choice. For the initial study of predictability of excess stock returns only one variable is taken into account, the dividend yield. A vector autoregressive model of the first order with some restrictions on its parameters is defined in order to examine how the evidence of predictability in asset returns affects optimal portfolio choice. The model is then applied to a real dataset and the results of the optimal portfolio allocation for different investment horizons are presented for a buy-and-hold investor who is risk-averse. Finally, the results obtained considering different initial values of the dividend yield are reported in order to understand the role of the predictor variable

We devote the fourth and fifth chapters to develop some extensions to the model implemented in chapters 2 and 3. We study the optimal portfolio allocation when investors have the opportunity to choose how to invest their wealth among three different assets: a stock index, a bond index, and the risk-free asset . The purpose of the fourth chapter is similar to the one of the second chapter, that is to understand how estimation risk alone affects portfolio choice. Some changes to the original model are therefore implemented in order to define an appropriate framework, which allows to us to examine the impact of parameter uncertainty when the investor can allocate his wealth among three different assets instead of two assets. The model we implemented is then applied to a real dataset, and the results of the optimal portfolio allocation for a buy-and-hold investor who is risk-averse are illustrated.

In the fifth chapter we focus on the study of predictability of excess stock and bond returns, and in order to do that, we avail ourselves of a set of five predictor variables. The model is similar in essence to the one we implement in the third chapter, a vector autoregressive model of the first order with some restrictions on its parameters. Applying it to a real dataset, we examine how the evidence of predictability affects portfolio choice when investors can choose to allocate their wealth between a stock index, a bond index and a risk free asset.

Finally, in the fifth chapter, after having related the main critiques to the Expected Utility Theory we bring up some experimental evidence that led to the emergence of Behavioral Finance. We then introduce the Prospect Theory, a behavioral economic theory that tries to describe investors' real-life choices, and we investigate how the optimal portfolio allocation changes when investors' preferences are described by a loss aversion utility function.

Chapter 1

Short run portfolio allocation

1.1 Introduction

The first one, is primarily a review chapter, whose purpose is to illustrate some ideas and concepts used throughout our work.
We firstly make a brief description of financial markets returns over the last two centuries. Afterwards we define the concept of asset return and illustrate some returns' appealing statistical properties. In this paragraph the meanings of risk-free rate and excess return are also explained.
The third paragraph is devoted to the Expected Utility Theory, which is used in order to describe economic agents' decisions under uncertainty.
In the fourth paragraph it is described the Mean-Variance Analysis, a portfolio choice theory whose main objective is to define the optimal portfolio allocation in the short-run; and its limitations are then given.
Finally we consider the case handled by Samuelson and Merton, when long-term investors act myopically, choosing the same portfolio as short-term investors.

1.1 Financial market returns from 1802

Risk and return are the fundamental blocks of finance and portfolio management. Once the risk and expected return of each asset are specified, modern financial theory can help investors define the best portfolios. But the risk and return on stocks and bonds are not physical constants. Despite the overwhelming quantity of historical data, one can never be certain that the underlying factors that generate asset prices have remained unchanged. One cannot, as in the physical sciences, run repeated controlled experiments, holding all other factors constant while estimating the value of the parameter in question.

However, one must start by analyzing the past in order to understand the future. In the next few paragraphs we carry a short analysis of past returns on stocks and bonds over the last two centuries. During this two-century period great changes have revolutionized the United States. The United States firstly made a transition from an agrarian to an industrialized economy and then became the main political and economic power in the world. Modern times led to the 1929 to 1932 stock collapse, the Great Depression, and the postwar expansion. The story is illustrated in Figure 1.1. It displays the real total return indexes for stocks, long and short-term bonds, gold, and commodities from 1802 through 2011. Since the focus of every long-term investor should be the growth of purchasing power that is, monetary wealth adjusted for the effect of inflation, the data in the graph are constructed by taking the dollar total returns and correcting them by the changes in the price level. Total return means that all returns, such as interest and dividends and capital gains, are automatically reinvested in the asset and allowed to accumulate over time

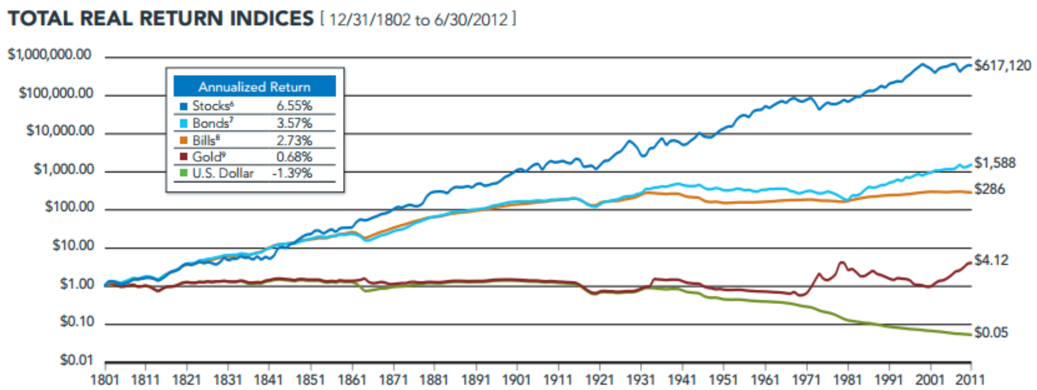

Figure 1.1: *Total real return indices, 1802 through June 2012*

It can be easily seen that the total real return on equities dominates all other assets and also shows remarkable long-term stability. Indeed, despite extraordinary changes in the economic, social, and political environment over the past two centuries, stocks have yielded about 6.6 percent per year after inflation. The wiggles on the stock return line represent the bull and bear markets that equities have suffered throughout history. The short-term fluctuations in the stock market, which appear so large to investors when they occur, are insignificant when compared to the upward movement of equity values over time. The long-term perspective radically changes one's view of the risk of stocks.

In contrast to the remarkable stability of stock returns, real returns on fixed-income assets have declined considerably over time. Until the twenties, the annual returns on bonds and bills, although less than those on equities, were significantly positive. But since those years, and especially since World War II, fixed-income assets have returned little after inflation.

Must however be said that in the real world investors consume most of the dividends and capital gains, so that the growth of the capital stock is not greater than the economy's rate of growth even though the total return on stocks is substantially higher. It is rare for anyone to accumulate wealth for long periods of time without consuming part of his or her return. The stock market has the power to turn a single dollar into millions by the perseverance of generations, but few will have the patience or desire to suffer the wait.

Although it might appear to be riskier to accumulate wealth in stocks rather than in bonds over long periods of time, precisely the opposite seems to be true: there is indeed evidence that the safest long-term investment for the preservation of purchasing power is a diversified portfolio of equities.

Indeed, according to the data Siegel(1994) availed himself of in his analysis , standard deviation, that is the measure of risk used in portfolio theory and asset allocation models, is higher for stock returns than for bond returns over short-term holding periods, however, once the holding period increases, stocks become less risky than bonds. The standard deviation of average stock returns falls nearly twice as fast as for fixed income assets as the holding period increases.

Theoretically the standard deviation of average annual returns is inversely proportional to the holding period if asset returns follow a random walk. But the historical data show that the random walk hypothesis can not be maintained for equities. Indeed the actual risk of stock declines far faster than the predicted rate under the random walk assumption. All that highlights one of the most relevant factors to be considered in making investment choices, that is the holding period.

Although the dominance of stocks over bonds is readily apparent in the long run, it is also important to note that in the short run, stocks outperform bonds or bills only about three out of every five years according to Siegel's research. The high probability that bonds and even bank accounts will outperform stocks in the short run is the primary reason why it is so hard for many investors to stay in stocks.

After a brief explanation of the main concepts and tools that will be used throughout our work, we will dedicate the next chapters to the exploration of the critical idea of how the holding period can affect the optimal allocation decision of an investor. We will firstly consider an investor who is allowed to choose how to invest his wealth only between a risk-free asset and a stock index, and afterwards we will add to the assets he can avail himself of a bond index.

1.2 Asset returns

When an empirical analysis in carried out, it is very important to use data whose type can supports the pursued objectives. Most financial studies involve returns instead of asset prices. There are at least two reason to contemplate returns rather than prices. Firstly, for the average investor, financial markets may be considered close to perfectly competitive, so that the size of the investment does not affect prices changes. Therefore, the return is a complete and scale-free summary of the investment opportunity. Secondly, returns have more attractive statistical properties than prices, such as stationarity and ergodicity

There are, however, several definitions of asset returns, we discuss some of them, that will be used throughout our work.

We denote by P_t the price of an asset at time t. We assume for the moment that the asset pays no dividends.

One-Period Simple Return

Holding the asset for one period from date $t-1$ to date t would result in a simple gross return :

$$1+R_t = \frac{P_t}{P_{t-1}}, \tag{1.1}$$

The corresponding one-period simple net return or simple return is:

$$R_t = \frac{P_t}{P_{t-1}} - 1 = \frac{P_t - P_{t-1}}{P_{t-1}}. \tag{1.2}$$

.

Continuously Compounded Return

The natural logarithm of the simple gross return of an asset is defined as the continuously compounded return or log return:

$$r_t = \log(1+R_t) = \log\frac{P_t}{P_{t-1}} = p_t - p_{t-1}, \tag{1.3}$$

where $p_t = \log P_t$.

Continuously compounded returns r_t enjoy some advantages over the simple net returns R_t. First statistical properties of log returns are more tractable, indeed it has not any lower bound and it is therefore compatible with the hypothesis of Normality. If r_t has normal distribution with mean μ mu and variance σ^2, the simple return has lognormal distribution with mean $E(1+R_t) = e^{\left(\mu+\frac{\sigma^2}{2}\right)}$ and variance $Var(1+R_t) = (e^{\sigma^2} - 1)e^{\left(2\mu+\sigma^2\right)}$. Secondly, when we consider multiperiod returns:

$$\begin{aligned} r_t(k) &= \log(1+R_t(k)) = \log((1+R_t)\cdot(1+R_{t-1})...(1+R_{t-k+1}) \\ &= \log(1+R_t) + \log(1+R_{t-1}) + ... + \log(1+R_{t-k+1}) \\ &= r_t + r_{t-1} + ... + r_{t-k+1}, \end{aligned} \tag{1.4}$$

Thus, the continuously compounded multiperiod return is simply the sum of the continuously compounded one-period returns involved. However, the simplification is more in the modeling of the statistical behavior of asset returns over time, indeed the previous assumption of normality hold true for multiperiod returns as well, since the sum of normally distributed variables is also normally distributed.

Dividend Payment

If an asset pays periodic dividends, the definitions of asset returns must be modified. Denote by D_t the asset's dividend payment between dates $t-1$ and t, and by P_t the asset's price at the end of period t . Thus, dividend is not included in P_t . Then the simple net return and continuously compounded return at time *t* may be defined as

$$R_t = \frac{P_t + D_t}{P_{t-1}} - 1 \quad \text{and} \quad r_t = \log(P_t + D_t) - \log(P_{t-1}). \tag{1.5}$$

Note that the continuously compounded return on a dividend-paying asset is a nonlinear function of log prices and log dividends. However, when the ratio of price to dividends is not too variable , this function can be approximated by a linear function of log prices and dividends.

Throughout our work we will use continuously compounded returns. Continuous compounding is usually preferred when the focus of interest is the temporal behavior of returns, since multiperiod returns can be computed overtly. Conversely, it is common to use simple returns when a cross-section of assets is being studied.

1.2.1 Portfolio returns

An investor's portfolio can be defined as his collection of investment assets where he allocates his wealth. Denote by R_{it} the simple return connected with the asset i , belonging to a portfolio counting N assets, and by ω_i its weight in the portfolio.

The simple return on a portfolio consisting of N assets is a weighted average of the simple net returns of the assets involved, where the weight on each asset is the percentage of the portfolio's value invested in that asset. If portfolio p places weight ω_{ip} on asset i, then the simple return on the portfolio at time t , R_{pt}, is related to the returns on individual assets R_{it}, by $R_{pt} = \sum_{i=1}^{N} \omega_{ip} R_{it}$ where $\sum_{i=1}^{N} \omega_{ip} = 1$.

Continuously compounded returns of a portfolio, unfortunately, do not have the above convenient property. Since the continuously compounded return on a portfolio is the logarithm of this linear combination, that is not equal to the linear combination of logarithms, in other words: $r_{pt} \neq \sum_{i=1}^{N} \omega_{ip} r_{it}$

Moreover, the sum of log-normal distributions is not defined as a log-normal. In empirical applications this problem is usually minor. When returns are measured over short intervals of time, and are therefore close to zero, the continuously compounded return on a portfolio is close to the weighted average of the continuously compounded returns on the individual assets: $r_{pt} \approx \sum_{i=1}^{N} \omega_{ip} r_{it}$.

1.2.2 Excess returns and risk-free asset

For the analysis that will be carried out later it is necessary to refer to a risk-free asset. The return on a risk-free asset may be defined as theoretical return of an investment with no risk of financial loss. The assumption is based on the evidence that in the market it is possible to find an asset that has a sure and well-known ex ante return. In practice, these assets are usually short-term government bonds of absolutely reliable countries, money market funds, or bank deposit. Formally, the risk-free random variable has constant expected value and a variance equal to zero. But it may appear risky since its returns can fluctuate over time and its variance move usually away from zero. Nevertheless their variability is minimal compared to the one of the risky assets and therefore can be well approximated to zero.

Since the risk free return can be obtained with no risk, it is implied that any additional risk taken by an investor should be rewarded with an higher return than the risk-free one. We measure the reward as the difference between the expected return on the risky asset and the risk-free rate. This difference is defined as the risk premium on common stocks.

It is often convenient to handle an asset's excess return, in place of the asset's return. Excess return is defined as the difference between the asset's return and the return on some reference asset, where the reference asset is usually assumed to be the risk-free one. In the next equation, z_{it} contains the simple excess return on the risky asset i relative to the risk-free asset.

$$z_{it} = r_{it} - rf \tag{1.6}$$

where r_f specifies the risk-free return.

An investor could choose to invest a portion of his wealth in the risk-free asset, as well as in the N risky assets. If you specify with ω_0 the portfolio's share of wealth invested in the risk-free asset, the portfolio return will then be:

$$r_p = \omega_0 r_f + \boldsymbol{\omega}'\mathbf{r}, \quad \text{where} \quad \omega_0 + \mathbf{i}'\boldsymbol{\omega} = 1, \tag{1.7}$$

alternatively

$$r_p = r_f + \boldsymbol{\omega}'(\mathbf{r} - r_f\mathbf{i}), \tag{1.8}$$

where $(\mathbf{r} - r_f\mathbf{i}) = \mathbf{z}$, vector of excess returns.

Subtracting r_f to both members of the expression above, we can obtain the portfolio excess return formula as function of risky assets' excess return.

$$z_p = \boldsymbol{\omega}\mathbf{z} \tag{1.9}$$

Here the weight vector does not sum to 1, since $\boldsymbol{\omega}$ only represents the proportion invested in risky assets.

Since the risk-free random variable is assumed to have a constant mean and a variance equal to zero, the riskiness of risky assets is often measured by the standard deviation of excess returns. However, due to r_f fluctuation over time, excess returns sample variances and covariances are not equal to returns'. Nevertheless the fluctuations of the risk-free assets are negligible compared with the uncertainty of stock market returns, the difference between the two variances will thus be small. Most of the time this condition is observed and the difference between the empirical variances of $\mathbf{r}$ and $\mathbf{z}$ is not significant.

The majority of economic models is based on hypothesis, not always verified, that return and excess return are independent realization from the same multivariate normal distribution.

1.3 Expected Utility Theory

Uncertainty plays a remarkable role in the investors' processes of taking decisions. Since the future is unknown, investors make their choice within an uncertain overall framework, where every action carries different consequences depending on the state of nature that it will come true. Each state of nature has its own probability of success, and therefore they have a specific probability distribution.

Economic agents' decisions under uncertainty can be represented as the choice of a particular prospect within a set of alternatives. In the case where individuals do not bother about the risk related to the choice of an uncertain prospect, their decisions are driven solely by the expected value criterion , which takes into account only the sizes of the payouts and the probabilities of occurrence. The alternative with the highest expected value will then be chosen. However, most people are not indifferent to the risk. Intuitively, one would rank each prospect as more attractive when its expected return is higher, and lower attractive when its risk is higher. But when risk increases along with return, the most attractive portfolio is not easy to be found anymore. How can investors quantify the rate at which they are willing to trade off return against risk? In situations involving uncertainty (risk), individuals act as if they choose on the basis of expected utility, the utility of expected wealth, rather than expected value.

Economists use Expected Utility Theory in order to explain decisions taken under uncertainty. This perspective, which focuses on man as a rational and predictable being in his actions, was developed in 1947 by Neumann and Morgenstern and has been widely accepted and applied as a model of economic behavior. According to this theoretical model, individuals, who are required to choose between several options, do not evaluate financial quantities depending on their amount, but on the satisfaction they subjectively confer on them. Investors can assign a welfare, or utility, score to competing investment portfolios based on the expected return and risk of those portfolios. The utility score may be viewed as a means of ranking portfolios, resulted from a criterion of personal choice, therefore it will depend on preferences of investors in a specific moment or situation. Higher utility values are assigned to portfolios with more attractive risk-return profile.

This theory allows us to study individual preferences, which are represented by a utility function u, which is defined barring a monotonic increasing transformation. This function has two properties: it must respect the preferences order of the individual and it must be increasing, that is it must have positive marginal utility of wealth, since it is reasonable to confer more utility to greater payoffs.

Given a function $u(x)$ where x corresponds to the wealth in $t+1$ and assuming $u'(x) > 0$ (rational investor), the expected utility of wealth result from

$$E[u(x)] = \sum_{i=1}^{S} p_i u(x_i) \tag{1.10}$$

where $\sum_{i=1}^{S} p_i = 1$ and S are the states of nature.

Investors choice criteria among several risky alternatives are always based on the expected utility result. Rational individuals choose the option that maximize their utility, on the basis of the expected utility rather than expected value of the outcomes. Therefore the preferred alternative depends on which subjective expected utility is higher. Different people may take different decisions because they may have different utility functions or different beliefs about the probabilities of varied outcomes.

Asked to choose between two prospects, a risk-free one, with sure return *R*, and a risky one with expected return equal to *R,* investors always compare $u(E[x])$ and $E[u(x)]$.

A risk averse individual would prefer to receive a certain return R rather than having an uncertain prospect whose expected value corresponds to R. He is therefore willing to give up a part of income in exchange for a sure outcome, since he considers uncertainty as a negative element. Financial analysts generally assume investors are risk averse in the sense that, if the risk premium were zero, people would not be willing to invest any money in stocks. A risk-averse investor penalizes the expected rate of return of a risky portfolio by certain percentage to account for the risk involved. The greater the risk, the larger the penalty. In theory, then, there must always be a positive risk premium on stocks in order to induce risk-averse investors to hold the existing supply of stocks instead of placing all their money in risk-free assets.

In contrast to risk-averse investors, risk-neutral investors judge risky prospectus solely by their expected rates of return. The level of risk is irrelevant to the risk-neutral investor, meaning that there is no penalty for risk.

A risk lover is willing to engage in fair games and gambles; this investor adjusts the expected return upward to take into account the pleasure of confronting the prospect's risk. Risk lovers will always take a fair game because their upward adjustment of utility for risk gives the fair game a higher utility than the risk-free investment.

The concept of risk aversion is useful to estimate risk effects in individuals' satisfaction level and in their preferences.

Risk attitude is directly related to the curvature of the utility function:

- A risk averse individual has concave utility function. Moreover the concavity shows diminishing marginal wealth utility.
- A risk neutral individual has linear utility function.
- A risk lover individual has convex utility function.

The degree of risk aversion can therefore be measured by the curvature of the utility function. Since the risk attitudes are unchanged under affine transformations of u, the first derivative , u' , is not an adequate measure of the risk aversion of a utility function. Instead, it needs to be normalized. This leads to the definition of the Arrow–Pratt measure risk aversion.

The Arrow–Pratt measure of absolute risk aversion is:

$$R_A(x) = -\frac{u''(x)}{u(x)}, \qquad (1.11)$$

Where u' and u'' are the first and second derivatives of the utility function and x is the generic outcome. The reasons behind the choice of this coefficient is intuitive: a function is concave if its second derivative is nonpositive. It is a local measure of risk, it depends in general on x, and its unit is the inverse of the outcome x one. This coefficient define the absolute amount an investor is willing to pay in order to avoid a risky situation. It is commonly assumed that absolute risk aversion decreases with wealth.

The Arrow–Pratt measure of relative risk aversion is:

$$R_R(x) = -x\frac{u''(x)}{u(x)} = x R_A . \qquad (1.12)$$

It has the advantage over the coefficient of absolute risk aversion to be independent of the monetary unit for wealth. It defines the share of wealth an investor is willing to pay in order to avoid a risky situation. Long term economic behavior shows that relative risk aversion is almost independent from wealth.

When investors are risk averse, and therefore the utility function is concave, the indicators are positive and the degree of risk aversion increases as their value raises.

1.4 Mean-Variance Analysis

History shows us that, in the short run, long-term bonds have been riskier investments than investments in Treasury bills, and that stock investments have been riskier still. On the other hand, the riskier investments have offered higher average returns. Investors, of course, do not make all-or-nothing choices from these investment classes. They can and do construct their portfolios using securities from all asset classes. Portfolio selection, that is the definition of the optimal allocation obtained maximizing expected utility, is indeed one of the most relevant issue an investor must deal with. The process of building an investment portfolio usually begins by deciding how much money to allocate to broad classes of assets, such as stocks, bonds, real estate, commodities, and so on. The choice among these broad asset classes is referred as asset allocation. Then, the portfolio's construction continues with the capital allocation between the risk-free asset and the risky portfolio. However, to define portfolio's shares that minimize risk and maximize return is the final purpose of this process.

Portfolio choice theory was originally developed by Markowitz (1952). In his Mean-Variance Analysis model he showed how investors should pick assets if they care only about mean and variance of portfolio returns over a single period. The main objective of this approach it to define the optimal portfolio and to track the efficient frontier that gather all the risk-return efficient opportunities. The system consists of two parts. In the first one, where investor's expectations and his risk aversion do not come into play, the risk-return combinations available from the set of risky assets are identified and the optimal portfolio of risky assets is selected. Secondly the investor chooses his appropriate optimal portfolio, combination of risk–free asset and optimal risky portfolio, maximizing his own satisfaction. In this last step, the introduction of individuals' preferences makes it possible to compare the efficient portfolios, and to take the final decision among them. The Expected Utility theory, that fully quantify the investor's position, represents the connecting element between these two parts.

The principal idea behind the frontier set of risky portfolios is that, for any risk level, investors are interested only in that portfolio with the highest expected return, or alternatively for any given level of expected return they prefer the portfolio which has minimum variance.

The efficient frontier can be obtained in two ways:

- Minimizing the portfolio's variance for all the possible values of expected return;
- Maximizing investor's expected return changing the portfolio's variance.

These two methods return the same efficient frontier when there is a square utility function or when returns have an elliptical distribution, as the case of a multivariate normal distribution.

The investor maximizes an objective function, that depends on the mean and variance of returns, in order to define a set of efficient portfolios, that constitute the efficient frontier. It is important to highlight that the set of efficient portfolios does not depend on the investor's expectation or on his risk aversion level.

When the first step is completed, the investor has a list of efficient portfolios, that is the efficient frontier of risky assets . Thus, he proceeds to step two and introduces the risk-free asset. The efficient frontier is now given as a straight line tangent to the efficient frontier of risky assets, and it is defined as Capital Market Line. The set of admissible portfolios is specified, now the investor must choose his optimum according to his own preferences and level of risk aversion. The optimum portfolio is therefore defined as the tangency point between the efficient frontier and the indifference curves derived from his utility function.

What the investor does in order to solve the risk-return trade-off, is to maximize his utility function defined over wealth in $t+1$. The wealth at the end of the period depends on the allocation decisions. And since the assets where he can invest are risky, his wealth will also have risky returns, whose we can compute the expected value and variance. Then the maximization problem is:

$$\max_{\omega} E[u(W_{t+1})] \tag{1.13}$$

subject to $W_{t+1} = (1+R_{t+1})W_t$, and where ω is a portfolio's share invested in the risky asset, or :

$$\max_{\omega} E[u(W_t(1+R_t))] = \max_{\omega} u(CE) \tag{1.14}$$

where the certainty equivalent is the guaranteed amount of money that an individual would view as equally desirable as a risky asset, $u(CE) = E[u(x)]$.
Similar results are available if we assume instead that the investor maximizes an objective function that is a liner combination of mean and variance with a positive weight on mean and a negative weight on variance.

$$\max_{\omega}\left(E_t(R_{p,t+1}) - \frac{R_A(W_{t+1})}{2}\sigma^2_{p,t} \right) \tag{1.15}$$

Where ω is the portfolio's share invested in the risky asset, $W_{t+1} = (1+R_{t+1})W_t$ and $\sigma^2_{p,t}$ is the portfolio's variance at time t.

The result of Markowitz analysis are shown in the mean-standard deviation diagram of Figure 1.2. The vertical axis shows expected return, and the horizontal axis shows risk as measured by standard deviation. Stocks offer a high expected return and a high standard deviation, bonds a lower expected return and lower standard deviation. The risk-free asset has a lower mean again, but is riskless over one period, so it is plotted on the vertical zero-risk axis. Investors can achieve any efficient combination of risk and return along the curve, that it is the efficient frontier, by changing the proportion of stock and bonds. Moving up the curve they increase the proportion in stocks and correspondingly reduce the proportion in bonds. As stock are added to the all-bond portfolio, expected returns increase and risk decreases, a very desirable combination for investors. But after the minimum risk point is reached, increasing stocks will increase the return of the portfolio only with extra risk. The slope of any point on the efficient frontier indicates the risk-return trade-off for that allocation. When the risk-free asset is added to a portfolio of risky assets, the efficient frontier becomes the straight line that passes through the risk-free point and is tangent to the curved line . This straight line, the Capital Market Line, offers the highest expected return for any given standard deviation. All investors who care only about mean and standard deviation will hold the same portfolio of risky assets. Conservative investors will combine this portfolio with a risk-free asset to achieve a point on the mean-variance efficient frontier that is low down and to the left; moderate investors will reduce their holdings in the risk-free asset, moving up and to the right; aggressive investors may even borrow to leverage their holdings of the tangency portfolio, reaching a point on the straight line that is even riskier than the tangency portfolio. But none of these investors should alter the relative proportions of risky assets in the tangency portfolio.

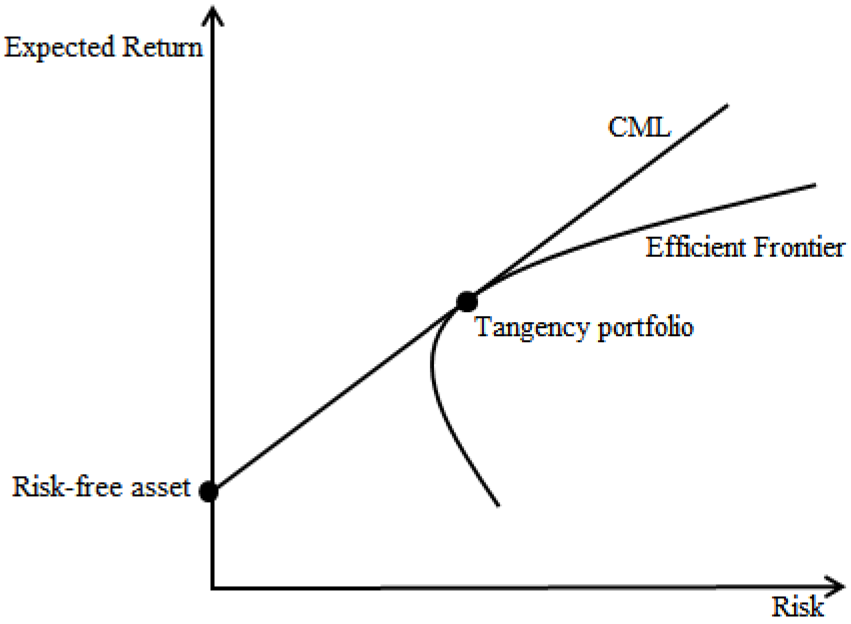

Figure 1.2: *Mean-standard deviation diagram*

1.4.1 The form of the utility function

As we mentioned before, models of portfolio choice require assumptions about the form of the utility function and about the distribution of asset returns. There are three alternative sets of assumptions that generate consistent result with those of the mean-variance analysis.

Investors have quadratic utility defined over wealth. That is, $U(W_{t+1}) = aW_{t+1} - bW_{t+1}^2$. Under this assumption maximizing expected utility is equivalent to maximizing a linear combination of mean and variance. No distributional assumptions on asset returns are required. Quadratic utility implies that absolute risk aversion and relative risk aversion are increasing in wealth.

Investors have exponential utility, $U(W_{t+1}) = -\exp(-\theta W_{t+1})$, and returns are normally distributed. Exponential utility implies that absolute risk aversion is a constant θ, while relative risk aversion increases in wealth.

Investors have power utility, $U(W_{t+1}) = (W_{t+1}^{A} - 1)/(1-A)$, and asset returns are lognormally distributed. Power utility implies that absolute risk aversion is declining in wealth, while relative risk aversion is a constant A . As A approaches one the limit is log utility: $U(W_{t+1}) = \log(W_{t+1})$

The power utility function seems to be the most suitable choice to explain investors' preference. Indeed, since absolute risk aversion should decline, or at the very least should not increase with wealth, the quadratic utility can be excluded, and the power utility can be preferred to the exponential utility. The power-utility property of constant relative risk aversion is attractive, and is required to explain the stability of financial variables. The choice between exponential and power utility also implies distributional assumptions on returns. Power utility function produces simple results if returns are lognormal. The assumption of lognormal returns, unlike the one of normal returns, can hold at every time horizon since products of lognormal random variables are themselves lognormal.The assumption of lognormal returns has another limit, however. It does not carry over straightforwardly from individual assets to portfolios. Anyway this difficulty can be avoided by considering short time intervals. Indeed, as the time interval shrinks, the non-lognormality of the portfolio return diminishes. Therefore, in the portfolio choice analysis that we are carrying out later, we use a power utility function to describe the investor's preferences.

1.4.2 Limitations of the Mean-Variance Model

The striking conclusion of Markowitz's analysis is that all investors who care only about mean and standard deviation must hold the same portfolio of risky assets and none of these investors should alter the relative proportions of risky assets in the tangency portfolio. But financial planners have traditionally resisted the simple investment advice embodied in Markowitz's Mean-Variance theory. One common pattern in financial advice is that conservative investors are typically encouraged to hold more bonds, relative to stocks, than aggressive investors, contrary to the constant bond-stock ratio suggested by the mean-variance model.

One possible explanation for this pattern of advice is that aggressive investors are unable to borrow at the riskless interest rate, and they thus cannot reach the upper right portion of the straight line in Figure 1.2. In this situation, aggressive investors should move along the curved line, increasing their allocation to stocks and reducing their allocation to bonds. The fact is that this explanation only applies once the constraint on borrowing starts to commit the investor, that is, once cash holdings have been reduced to zero; but the bond-stock ratio often changes even when cash holdings are positive.

Markowitz's mean-variance approach can be applied only when investor's preferences are described by a quadratic utility function, of the mean-variance kind, or when the distribution of risky returns is elliptical. Although these hypothesis allow to obtain explicit solutions, they are strong assumptions, that do not describe the reality: the quadratic utility function is not enough flexible and for some specific combination it may violate the non satiety assumption (for high values of wealth you can have a reduction in utility), the normal distribution of returns can be used when markets are not excessively volatile; increasing the frequency of observations from annual to monthly or weekly the returns' distribution usually deviates from a normal one. Therefore, for long time horizons and for violation of one of these two assumption the approximation included in the mean-variance approach is not sufficiently accurate. An additional possibility is the hypothesis that investor's preferences violate the axioms of Expected Utility theory, as in the Prospect Theory of Kahneman and Tversky (1979).

Moreover, so far we have assumed that the investor has a short investment horizon and cares only about the distribution of wealth at the end of the next period. In reality, investors are more interested in maintaining a certain standard of living through long-term investment. If individuals with long horizon invest repeatedly in the efficient uniperdiodal portfolio, they achieve an efficient strategy when:

- they have constant relative risk aversion and own only financial wealth;
- asset returns are i.i.d.;
- there is no uncertainty in the estimated parameters;
- there are no transaction costs.

Most of these assumptions are not realistic, therefore we can not consider Mean-variance analysis an appropriate model for long-term investment. Merton (1969) found that in a multiperiod context portfolio choice can be significantly different.

1.5 The holding period

Beyond single agent's preferences, a lot of other factors affects optimal portfolio choice. For instance, an individual with a long investment horizon may consider risk differently from a short-horizon investor. Thus, the optimal portfolios of long-horizon investors do not need necessarily to have the same composition of those of short-horizon investors. Given these important results, it might seem puzzling that the holding period has almost never been mentioned before.

In order to understand the optimal portfolio allocation when several holding periods are taken into account, it is essential to specify the behaviors an investor can adopt.

- *Buy-and-hold*

  An agent with investment horizon of t years chooses the portfolio allocation at the beginning of the first year and does not touch his portfolio again until the t years are over. The buy-and-hold strategy is a passive and static investment strategy: once the portfolio is created, it is not handled in any way.
- *myopic rebalancing*

  The investor chooses some arbitrary intervals to rebalance the portfolio, for example every year. He then chooses an allocation at the beginning of the first year, knowing that he will always choose the initial allocation at the beginning of every year. This strategy is called myopic because the individual does not use any of the new information he has once a year is passed to allocate the portfolio in an optimal way for the subsequent years. Moreover , it is similar to the buy-and-hold strategy since over the years always the same allocation is chosen, as if the investor would not intervene until the end of the investment horizon.

- *Optimal rebalancing*

  The investor chooses today the allocation of his portfolio, knowing that at regular intervals he may reallocate the portfolio using all the new information available up to that moment. This is the most sophisticated technique to manage a portfolio in a dynamic and uncertain context as the financial market.

This paper presents the results for a buy-and-hold investor who faces the problem of portfolio choice in several investment horizons.

1.5.1 Long-run portfolio choice

Illustrating the classic Mean-Variance Analysis we assumed that the investor has a short investment horizon and cares only about the distribution of wealth at the end of the next period. However, most of the time, investors are more interested in maintaining a certain standard of living through long-term investment.
Financial economists recognized the need for a long-term portfolio choice theory in the 70's . They started to develop empirical models of portfolio choice for long term investors, building on the fundamental insights of Samuelson and Merton; important contributions came from Rubinstein, Stigliz and Breeden.

Below we try to explain those special cases in which long-term investors should take the same decisions as short-term investors. In these special examples the investment horizon is irrelevant; portfolio choice is therefore said myopic.

Classic results of Samuelson (1969) and Merton (1969, 1971) show two sets of conditions under which the long-term agent acts myopically, choosing the same portfolio as a short-term agent.

Firstly, portfolio choice will be myopic, if the investor has power utility and returns are i.i.d**.** As we already stated, power utility implies the presence of constant relative risk aversion. With constant relative risk aversion, portfolio choice does not depend on wealth, and hence does not depend on past returns. Moreover if returns are i.i.d, no new information emerges between one period and the next so there is no reason for portfolio choice to change over time in a random way. The investor with power utility function, who rebalances over time his portfolio, will choose the same allocation of short period regardless of the investment horizon. The choice of myopic portfolio is therefore optimal if investors do not have labor income and if investment opportunities are constant over time.

The second condition for myopic portfolio choice is that investor has log utility. In this case portfolio choice will be myopic even if asset return are not i.i.d.. Hence also if investment opportunities vary over time, with this utility function the horizon becomes irrelevant. The argument here is simple. Indeed, if the log utility investor chooses a portfolio that maximizes the expected log return, K-period log return is just the sum of 1-period log returns. Since the portfolio can be chosen freely each period, the sum is maximized by maximizing each of its elements separately, that is, by choosing each period the portfolio that is optimal for a 1-period log utility investor.

Nonetheless a typical pattern in financial advice is the tendency for financial planners to encourage young investors, with a long horizon, to invest mainly in stocks compared to older investors who have a shorter horizon. In this work we will explore the conditions under which a long investment horizon indeed justifies a different allocation, therefore contrasting with Samuelson and Merton conclusion.

We devote the next chapters to studying the optimal portfolio decision when Samuelson and Merton’s assumptions are infringed, in particular in the third chapter we allow for predictability in asset return rather than consider an i.i.d. context, still employing a power utility function.

Chapter 2

Portfolio allocation with parameter uncertainty

2.1 Introduction

In this chapter we deal with the optimal portfolio allocation assuming that parameters are not known precisely. Our purpose is to understand how parameter uncertainty alone affects portfolio choice.

We devote the third paragraph to a brief description of the data set used throughout our work, and to some preliminary analysis of the data in order to examine their features.

We then present the model that handles portfolio choice under several investment horizons and under the case where the investor either ignores or accounts for parameter uncertainty.

In the fifth paragraph the results obtained by implementing the model to the data set are reported and explained.

Finally we adopt the resampling approach in order to simulate data from the real and unknown generating process. We therefore understand if the assumption of normality attributed to assets returns, that has a critical role in the construction of the model, affects the portfolio optimal allocation.

2.2 Parameter uncertainty

Theoretical models often assume that an investor who makes an optimal financial decision knows the true parameters of the model, but the true parameter are rarely if ever known to the decision maker. In reality, model parameters need to be estimated and, hence, the model's usefulness depends in part on how good the estimates are. This gives rise to estimation risk in virtually all financial models. Estimation risk is defined as the investor's uncertainty about the true values of model parameters. The parameter uncertainty increases the perceived risk in the economy and necessarily influences portfolio decisions, it is therefore the primary source of deviation from reasonably satisfactory and consistent solutions. At present, estimation risk is commonly minimized based on statistical criteria such as minimum variance and asymptotic efficiency. The reasons why this type of risk exists may be attributed to two specific factors: sampling error, when inputs are estimated , and non-stationarity of the time series.

A first example of parameter uncertainty arises from the classic portfolio choice problem. Markowitz's work shows that the optimal portfolio for an investor who cares only about mean and standard deviation is a combination of tangency portfolio and the risk-free asset. Despite its limitation as a single-period model already mentioned before, the mean-variance framework is one of the most important benchmark models used in practice today. However the framework requires knowledge of both the mean and covariance matrix of the asset returns, which in practice are unknown and have to be estimated from the data. The standard approach, ignoring estimation risk, simply treats the estimates as the true parameters and plugs them into the optimal portfolio formula derived under the mean-variance framework. Even though we assume that investors know these parameters with certainty, we can not be sure that the estimated values coincide effectively with the true values of the parameters. The investor would face two problems at the same time: a portfolio allocation problem and an inferential problem.

The concept of parameter uncertainty was first investigated by Bawa, Brown and Klein (1979) who explore the issue in the context of i.i.d. returns. Whereas Kandel and Stambaugh (1996) were the first to explore the problem of parameter uncertainty in the context of portfolio allocation with predictable returns. They show that for a short-horizon investor, the optimal allocation can be sensitive to the current value of predictor variables, even though regression evidence for such predictability may be weak. In our analysis we focus on a wider range of horizons, from one month to 10 years, rather than the one-month horizon of Kandel and Stambaugh.

The studies on estimation risk typically focuses on the subjective distribution perceived by investors. Since investors do not know the true distribution, they must estimate the parameters using whatever information is available, which can be formally modeled using Bayesian analysis. The subjective distribution combines investors' prior beliefs with the information contained in observed data. Indeed, rather than constructing the distribution of future returns conditional on fixed parameter estimates, they can integrate over the uncertainty in the parameters captured by the posterior distribution. This allows them to construct what is known in Bayesian analysis as the predictive distribution for future returns, conditional only on observed data, and not on any fixed parameter values. This distribution represents investors' best guess about future returns, and is therefore relevant for investment decisions.

Our first set of results relates to the case where parameter uncertainty is ignored, that is, the investor allocates his portfolio taking the parameters as fixed at their estimated values; then we consider the case where the investor takes into account uncertainty about model parameters. By comparing the solution in the case where we condition on fixed parameters, and where we use the predictive distribution conditional only on observed data, we see the effect of parameter uncertainty on the portfolio allocation problem.

2.3 Data set

To illustrate our approach, we use monthly U.S. financial data for the period January 1990-November 2012, the sample consisting therefore of 275 monthly data. We begin our analysis including only one risky asset that, combined with the risk-free one, constitute the investor optimal portfolio choice.

In our study, the risky asset is the S&P 500 Index, and the risk-free asset is a short-term debt instrument.

The S&P 500 is the most widely accepted barometer of the market. This value weighted index was firstly compiled in 1957 when it included 500 of the largest industrial, rail, and utility firms that traded on the New York Stock Exchange. It soon became the standard against which the performance of institutions and money managers investing in large U.S. stocks was compared. It now includes 500 large-cap stocks, which together represent about 75% of the total U.S. equities market. The

S&P 500 thus provide a convenient way to examine the behavior of stock returns. Returns on the index were computed assuming continuous compounding, from the monthly total return time series downloaded from Datastream.

The risk-free asset used in the analysis is the 3-month Treasury Bill, downloaded from FRED (Federal Reserve Economic Data) a database of the of the Federal Reserve Bank of St. Luis. The available data are annualized, therefore we divided the annualized rates by 12 in order to get the monthly rates of return.

2.3.1 Preliminary analysis

Equity index's total return time series is non-stationary and it has frequent changes in mean, as it is displayed in Figure 2.1.

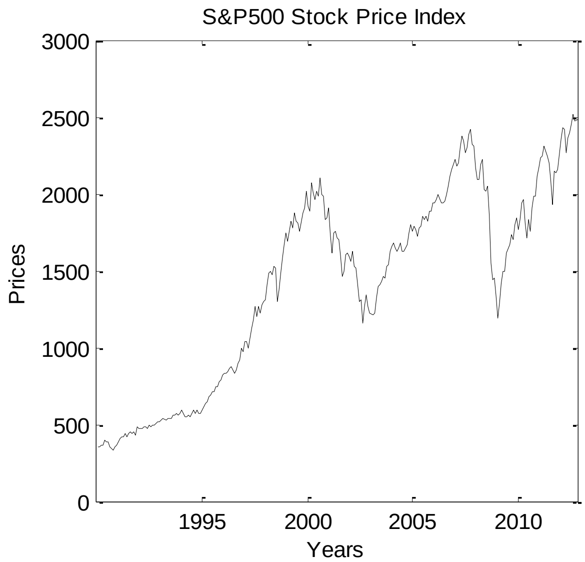

Figure 2.1: *S&P 500 Stock Price Index over the period 1990-2012.*

Returns instead exhibit more attractive properties, that is the reason why we use returns in place of prices series throughout our work. Continuous compounded returns are computed according to equation (1.3), starting from the total return series of the stock index. We now make a brief analysis of these returns properties. The returns considered here are stationary, and the autocorrelation function confirms that.

Moreover analyzing the empirical autocorrelation function we can see that returns are uncorrelated. They have a positive mean of 0.0070, that is significant since the t-statistic (obtained dividing the returns' mean by its standard error) , is equal to 2.6945 , which is greater than the critical value 1.96.

S&P500 Logarithmic Returns			
Mean	0.007082	St. Error	0.002628
Minimum	-0.183863	Variance	0.001900
Maximum	0.108277	St. Dev	0.043587
1° Quartile	-0.017341	Skewnees	-0.773007
3° Quartile	0.035467	Kurtosis Excess	1.579464

Table 2.1: *Main descriptive statistics of S&P 500 Continuously Compounded Returns over the period 1990-2012.*

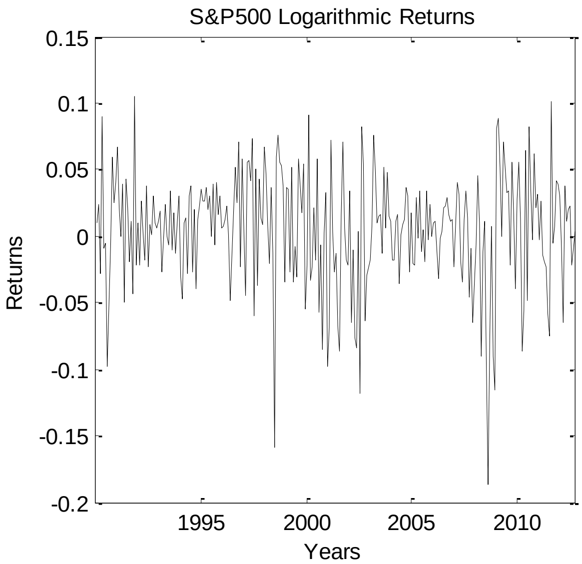

Figure 2.2: *S&P 500 Continuously Compounded Returns over the period 1990-2012*

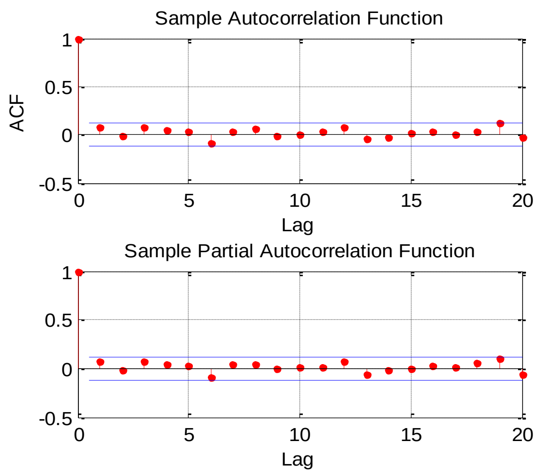

Figure 2.3: *Empirical correlogram of S&P 500 returns.*

When returns are calculated assuming continuous compounding they are hypothesized to have a Normal distribution. This hypothesis hold true for multiperiod returns as well, since they are simply the sum of the continuously compounded one-period returns involved. The assumptions of normality, attributed to the assets' returns, has a fundamental role in the construction of the model, however there are empirical reasons to believe that it does not represent an adequate description of the returns' generator process. We now test for the normality of our sample.

There are several test statistics that can be used in order to verify the normality of the returns series. The simplest ones are based on the properties of the indexes of skewness and kurtosis. Indeed under normality assumption $\hat{S}(x)$ and $\hat{K}(x)-3$ are distributed asymptotically as normal with zero mean and variance $6/T$ and $24/T$, respectively. These asymptotic properties can be used to test the normality of asset returns. Given our asset series, the skewness and excess of kurtosis of returns can be verified throughout the use of marginal tests respectively based on S and K. Jarque and Bera (1987) combine the two tests and use the test statistic

$$JB=\left(\frac{\hat{S}}{\sqrt{6/T}}\right)^2+\left(\frac{\hat{K}-3}{\sqrt{24/T}}\right)^2,$$

which is asymptotically distributed as a chi-squared random variable with 2 degrees of freedom, to test for the normality of asset return series.

Another statistic used in order to test for the hypothesis of normality, when the mean and variance are not specified, is the Lilliefors one. Initially the empirical mean and variance are estimated from the available data, then the maximum discrepancy between the empirical distribution function and the cumulative distribution function of the normal distribution, with the estimated mean and variance, is found . Finally the obtained statistic value is compared with the critical values of the Lilliefors distribution in order to assess whether the maximum discrepancy is large enough to be statistically significant , thus requiring rejection of the null hypothesis.

Normality Test	
Jarque-Bera	0.001
Lilliefors	0.023

Table 2.2: *Normality tests' P-values for the returns series.*

Our results reject the null hypothesis of normal returns for a significance level of 0.05. A confirmation of what has been said, the normal probability plot in Figure 2.5 shows a departure of sample quantiles from the theoretical ones of the normal distribution, in particular on the left queue. Moreover, the empirical density function of the returns series in Figure 2.4, has a particularly high peak around its mean and exhibits a skewness on the left side and leptokurotsis, sign that extreme returns are more likely to happen compared to a normal distribution.

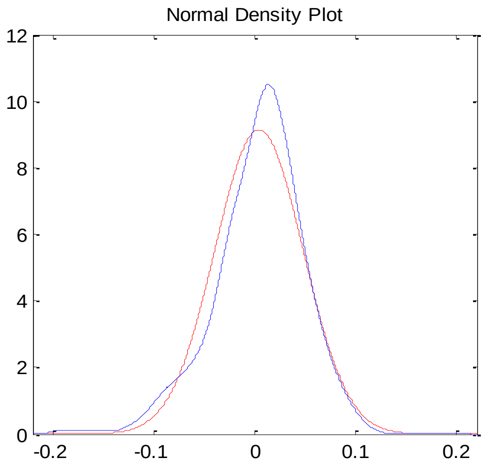

Figure 2.4: *Empirical density function of the S&P 500 returns series and normal probability density function evaluated by using the sample mean and standard deviation.*

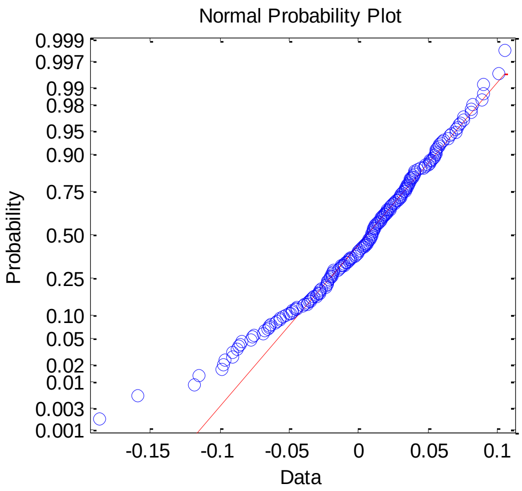

Figure 2.5: *Normal probability plot of S&P 500 returns series.*

We now consider the short-term interest rate series, the U.S. Treasury Bill with a maturity of three months. Looking at the autocorrelation function in Figure 2.6 we hypothesize a non stationary series. When we implement the Dickey-Fuller test without constant, since it is not significant, we obtain a value of the t-statistic equal to -2.14, which is smaller, in absolute value, than the critical value -2.58, and confirm the presence of unit root at a significance level of 1%. However the monthly interest rate is a very small number and has a lower variance compared to the stocks returns' one. Therefore interest rates can be considered almost constant and they can be set equal to their sample mean r_f . The effect of the approximation can be considered irrelevant to the analysis. For these reasons, the interest rate is treated as risk-free and it is used to build the equity index excess returns. The excess returns, obtained as difference between stock returns and r_f, retain all the properties that characterize the equity index. Only some descriptive statistics on position indexes, such as average, quartiles and extremes change .

3-Month treasury Bill			
Mean	0.002735	St. Error	6.527e-006
Minimum	8.33e-006	Variance	3.199e-006
Maximum	0.006562	St. Dev	0.001789
1° Quartile	0.000991	Skewnees	-0.128481
3° Quartile	0.004158	Kurtosis Excess	-1.113750

Table 2.3: *Main descriptive statistics of 3-Month Treasury Bill over the period 1990-2012.*

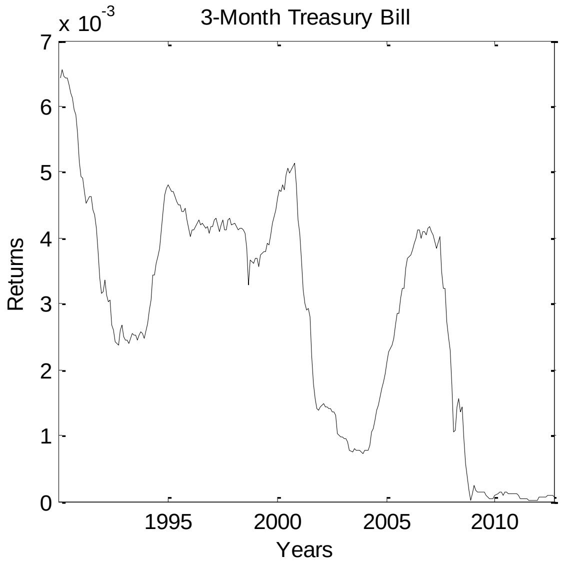

Figure 2.6: *3-Month Treasury Bill over the period 1990-2012.*

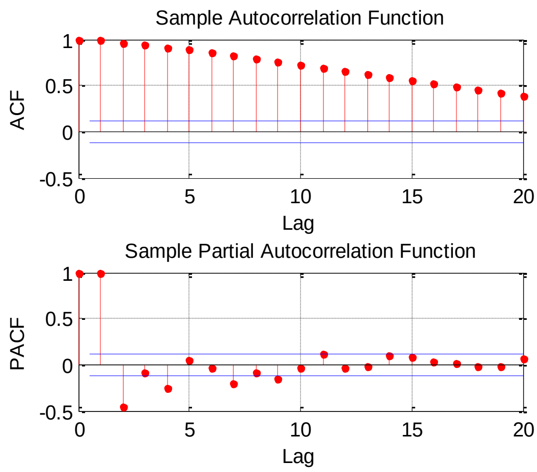

Figure 2.7: *Empirical correlogram of 3-Month Treasury Bill*

If we look at the Figure 2.8 , we can see that the empirical density function of the interest rate differs from normal probability density evaluated by using the sample mean and standard deviation. Furthermore, the normal probability plot exhibits a strong departure of the empirical queues from the theoretical ones. This is confirmed by the normality tests we implemented, which lead to reject the hypothesis of normality of the risk-free asset.

Normality Test	
Jarque-Bera	0.005
Lilliefors	0.01

Table 2.4: *Normality tests' P-values for 3-Month Treasury Bill series.*

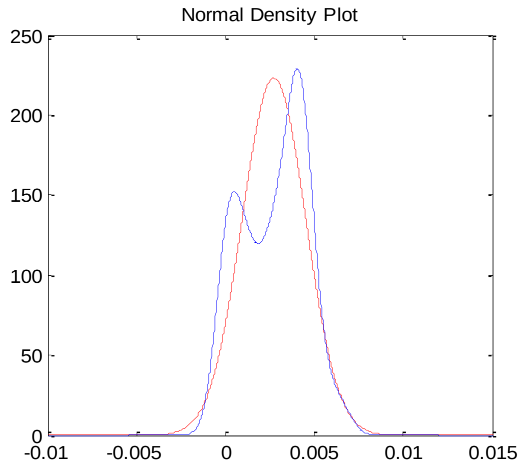

Figure 2.8: *Empirical density function of the 3 Month Treasury Bill series and normal probability density function evaluated by using the sample mean and standard deviation.*

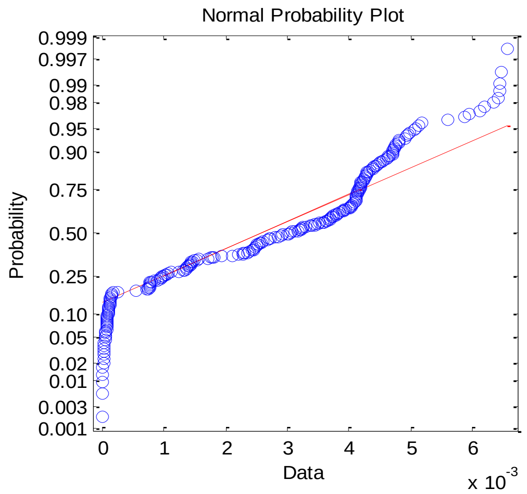

Figure 2.9: *Normal probability plot of the 3 Month Treasury Bill series.*

2.4 Long horizon portfolio allocation

This section is dedicated to the presentation of the model developed by Barberies (2000), that deals with the portfolio choice under several investment horizons and under the case where the investor either ignores or accounts for parameter uncertainty and returns predictability. We start out our analysis by considering the case where no predictor variables are included in the model, and hence where asset returns are i.i.d., and look at how parameter uncertainty alone affects portfolio allocation. There are two assets: Treasury bill and equity index, in this case the value-weighted index S&P 500. For simplicity, we suppose that the continuously compounded monthly return on Treasury Bills is a constant r_f . The excess return on the risky asset is obtained as difference between the stock return and r_f, and it is continuously compounded.

As we have just said, we model excess returns on the stock index assuming that they are i.i.d., so that

$$r_t = \mu + \varepsilon_t, \tag{2.1}$$

where r_t is the continuously compounded excess return on the equity index over month t, and where $\varepsilon_t \sim \text{i.i.d.} N(0,\sigma^2)$.

Suppose we are at time T and want to write down the portfolio problem for a buy-and-hold investor with a horizon of $\hat{T}$ months. If the investor has no chance to buy or sell assets between time T and horizon $T+\hat{T}$, he will only be interested in the distribution of wealth at the end of the investment period, that is $u(W_{T+\hat{T}})$. The most used utility function for portfolio allocation problem is the power utility function, that has absolute risk aversion declining in wealth, while relative risk aversion is constant. The investor's preferences over terminal wealth are then described by a constant relative risk-aversion power utility function of the form:

$$u(W) = \frac{W^{1-A}}{1-A} \tag{2.2}$$

where A is the coefficient of relative risk aversion

If initial wealth $W_T = 1$ and ω is the allocation to the stock index, then end-of-horizon wealth is given by

$$W_{T+\hat{T}} = (1-\omega)\exp(r_f \hat{T}) + \omega \exp(r_f \hat{T} + r_{T+1} + ... + r_{T+\hat{T}}) \tag{2.3}$$

If we write the cumulative excess stock return over $\hat{T}$ periods as

$$R_{T+\hat{T}} = r_{T+1} + r_{T+2} + ... + r_{T+\hat{T}}, \tag{2.4}$$

the buy-and-hold investor's problem is to solve

$$\max_{\omega} E_T \left(\frac{\left\{(1-\omega)\exp(r_f \hat{T}) + \omega \exp(r_f \hat{T} + R_{T+\hat{T}})\right\}^{1-A}}{1-A} \right) \tag{2.5}$$

E_t denotes the fact that the investor calculates the expectation conditional on his information set at time T, adopting the distribution of cumulative excess returns $R_{T+\hat{T}}$. We have therefore to define which distribution the investor should use in calculating this expectation. Indeed, the distribution may be different depending on whether the investor accounts for parameter uncertainty or not. The effect of parameter uncertainty is then revealed by comparing the optimal portfolio allocation obtained in these two cases.

Ignoring parameter uncertainty

Once the parameters $\theta = (\mu, \sigma^2)$ have been estimate, a distribution for future stock excess returns conditional on a set of parameter values and on the data observed by the investor up until the start of his investment horizon is generated, which we write as $p(R_{T+\hat{T}} \mid \hat{\mu}, \hat{\sigma}^2, r)$. Since $R_{T+\hat{T}}$ is the sum of $\hat{T}$ normally distributed random variables with mean μ and variance σ^2, the sum $R_{T+\hat{T}}$ is normally distributed conditional on μ and σ^2 with mean $\hat{T}\mu$ and variance $\hat{T}\sigma^2$.

The investor then solves

$$\max_{\omega} \int \upsilon(W_{T+\hat{T}})\, p(R_{T+\hat{T}} \mid r, \hat{\theta})\, dR_{T+\hat{T}}. \tag{2.6}$$

The shortcoming with this approach is that it ignores the fact that theta $\theta = (\mu, \sigma^2)$ is not known precisely. There may be substantial uncertainty about the regression mean of μ and σ^2 .

Incorporating parameter uncertainty

A natural way to take the uncertainty in the estimations into account is to use Bayesian concept of posterior distribution $p(\theta | r)$, which summarizes the uncertainty about the parameters given the data observed so far. To construct the posterior distribution $p(\mu, \sigma^2 | r)$ a prior is required. A potential choice could be the uninformative prior

$$p(\mu, \sigma^2) \propto \frac{1}{\sigma^2}. \tag{2.7}$$

But we could also have used a more informative prior, which for instance puts zero weight on negative values of μ , reflecting the consideration of Merton (1980) that expected market risk premium should be positive.

The resulting posterior distribution derived by Zellner (1971) consists of the marginal distribution Inverse Gamma

$$\sigma^2 | r \sim IG\left(\frac{T-1}{2}, \frac{1}{2}\sum_{i=1}^{T}(r_t - \bar{r})^2\right) \tag{2.8}$$

and of the conditional Normal distribution

$$\mu | \sigma^2, r \sim N\left(\bar{r}, \frac{\sigma^2}{T}\right), \tag{2.9}$$

Indeed, to sample form the posterior $p(\mu, \sigma^2 | r)$, we firstly sample from the marginal $p(\sigma^2 | r)$, an Inverse Gamma distribution, and then, given the σ^2 drawn, from the conditional $p(\mu | \sigma^2, r)$, a Normal distribution.

Integrating over this distribution, we obtain the predictive distribution for long-horizon returns. This distribution is conditioned only on the sample observed, and not on any fixed theta θ:

$$p(R_{T+\hat{T}} \mid r) = \int p(R_{T+\hat{T}} \mid r, \theta)\, p(\theta \mid r)\, d\theta. \tag{2.10}$$

The investor then solve

$$\max_{\omega} \int \upsilon(W_{T+\hat{T}})\, p(R_{T+\hat{T}} \mid r)\, dR_{T+\hat{T}}. \tag{2.11}$$

It could be helpful to rewrite the problem as

$$\begin{aligned} &\max_{\omega} \int \upsilon(W_{T+\hat{T}})\, p(R_{T+\hat{T}}, \theta \mid r)\, dR_{T+\hat{T}}\, d\theta \\ &= \max_{\omega} \int \upsilon(W_{T+\hat{T}})\, p(R_{T+\hat{T}} \mid r, \theta)\, p(\theta \mid r)\, dR_{T+\hat{T}}\, d\theta. \end{aligned} \tag{2.12}$$

The integral can therefore be evaluated by sampling from the joint distribution $p(R_{T+\hat{T}}, \theta \mid r)$, and then averaging $\upsilon(W_{T+\hat{T}})$ over those draws. We sample from the joint distribution by first sampling from the posterior $p(\theta \mid r)$ and then from the conditional $p(R_{T+\hat{T}} \mid r, \theta)$, $N(\hat{T}\mu, \hat{T}\sigma^2)$.

In order to solve the maximization problems (2.6) and (2.11) we calculate the integrals for several values of the proportion invested in the equity index, that is $\omega = 0, 0.01, 0.02, ..., 0.98, 0.99$, and report the ω that maximizes expected utility. We therefore restrict the allocation to the interval $0 \le \omega \le 1$ precluding short selling and buying on margin**.** In section 2.5, we present the optimal allocation ω which maximize expected utility for a variety of risk aversion levels A and investment horizons ranging from 1 month to 10 years, and for each of the two cases where the investor either ignores or account for parameter uncertainty.

The integrals themselves are evaluated numerically by simulation. For instance, if we are trying to evaluate

$$\int g(y)\, p(y)\, dy,$$

where $p(y)$ is a probability density function. We can approximate the integral by

$$\frac{1}{I}\sum_{i=1}^{I} g\,(y^{(i)}),$$

where $y^{(1)},...,y^{(I)}$ are independent draws from the probability density $p(y)$.

Thus we approximate the integral for the calculation of the expected utility by taking a sample $R_{T+\hat{T}}^{(i)}$ from one of the two possible distributions, and then computing

$$\frac{1}{I}\sum_{i=1}^{I}\left(\frac{\left\{(1-\omega)\exp(r_f\hat{T}) + \omega\exp(r_f\hat{T} + R_{T+\hat{T}}^{(i)})\right\}^{1-A}}{1-A}\right).$$

We chose to avail ourselves of the interactive environment of numerical computation and programming MATLAB, in order to implement the model described before. The employed commands are listed in Appendix B.

2.4.1 Sampling process

As we have just mentioned in equation (2.10) , there are two steps to sampling from the predictive distribution for long-horizon returns $p(R_{T+\hat{T}}\,|\,r)$. Firstly, we generate a large sample from the posterior distribution for the parameters $p(\mu,\sigma^2\,|\,r)$. We sample form the marginal $p(\sigma^2\,|\,r)$, an Inverse Gamma distribution, and then, given the $\hat{\sigma}^2$ drawn, from the conditional $p(\mu\,|\,\hat{\sigma}^2,r)$, a Normal distribution. To ensure a high degree of accuracy we fix the sample size I = 200000 throughout, and we repeat this 200000 times in order to give an accurate representation of the posterior distribution. The second step in sampling from the predictive distribution is to sample from the distribution of returns conditional on fix parameter values and past data $p(R_{T+\hat{T}}\,|\,\hat{\mu},\hat{\sigma}^2,r)$. The sum $R_{T+\hat{T}} = r_{T+1} + r_{T+2} + ... + r_{T+\hat{T}}$ is Normally distributed conditional on $\hat{\mu}$ and $\hat{\sigma}^2$ with mean $\hat{T}\hat{\mu}$ and variance $\hat{T}\hat{\sigma}^2$. Therefore for each of the pairs of $\hat{\mu}$ and $\hat{\sigma}^2$ drawn from the posterior $p(\mu,\sigma^2\,|\,r)$, we sample one point form the Normal distribution with mean $\hat{T}\hat{\mu}$ and variance $\hat{T}\hat{\sigma}^2$. This gives a sample of size 200000 from the predictive distribution $p(R_{T+\hat{T}} \mid r)$ which we can use to compute the optimal allocation when taking estimation risk into account.

When parameter uncertainty is ignored, the investor samples instead from the distribution of future returns conditional on fixed parameters and past data $p(R_{T+\hat{T}} \mid \hat{\mu}, \hat{\sigma}^2, r)$. We assume that the investor takes the posterior mean of μ and σ^2 as the fixed values of parameters, and then draws 200000 times forma a Normal distribution with mean $\hat{T}\hat{\mu}$ and variance $\hat{T}\hat{\sigma}^2$.

2.5 Results

The framework we have just introduced allows us to understand how parameter uncertainty affects portfolio choice. This section presents the results of our analysis. The objective is to show how the portfolio allocation changes as the investment horizon of a buy-and-hold investor increases, and how the optimal allocation ω changes depending on whether parameter uncertainty is taken into account or ignored in the model.

We simply compare the solution to problem (2.6) which ignores parameter uncertainty, with the solution to problem (2.11) which takes uncertainty into account. The result are based on the model $r_t = \mu + \varepsilon_t$, where r_t is the continuously compounded excess stock index return in month t and $\varepsilon_t \sim \text{i.i.d.} N(0, \sigma^2)$.

Table 2.5 gives the mean and standard deviation (in parentheses) of the posterior distribution $p(\mu, \sigma^2 \mid r)$ for each parameter μ and σ^2.

1990-2012	
μ	σ^2
0.0044	0.0019
(0.0026)	(0.0002)

Table 2.5: *Mean and standard deviation (in parenthesis) of each parameter's posterior distribution.*

For an investor using the entire sample from 1990 to 2012, the posterior distribution for the mean monthly excess return μ has mean 0.0044 and standard deviation 0.0026. This seems to be an important source of parameter uncertainty for the investor. The posterior distribution for the variance σ^2 is more compact and is centered around 0.0019.

Ignoring parameter uncertainty

When the investor does not take into account parameters uncertainty , he solves the maximization problem (2.6), employing a distribution for future excess returns conditional on the parameter values and on the observed data of this form $p(R_{T+\hat{T}} \mid \hat{\mu}, \hat{\sigma}^2, r)$, which is normally distributed with mean $\hat{T}\hat{\mu}$ and variance $\hat{T}\hat{\sigma}^2$. In this case $\hat{\mu}$ and $\hat{\sigma}^2$ are the means of each parameter's posterior distribution shown in Table 2.5.

Figure 2.10 shows the optimal portfolio allocation for a buy-and-hold investor, whose preferences over terminal wealth are described by a constant relative risk-aversion power utility function. The optimal percentage ω allocated to the stock index, is plotted against the investment horizon that range from 1 month to 10 years. The graph on the left side is based on a relative risk-aversion level of $A = 5$, the one on the right is for $A = 10$.

The line, that shows the percentage ω allocated to the stock index on varying holding period, is completely horizontal in both the graphs. An investor ignoring the uncertainty about the mean and variance of asset allocation returns would therefore allocate the same amount to stocks, regardless of his investment horizon. This is similar to Samuelson's result where he showed that with power utility function and i.i.d. returns, the optimal allocation is independent of the horizon. However, it is important to note that he proves this for an investor who optimally rebalances his portfolio at regular intervals, rather than for an investor who follows a buy-and-hold strategy.

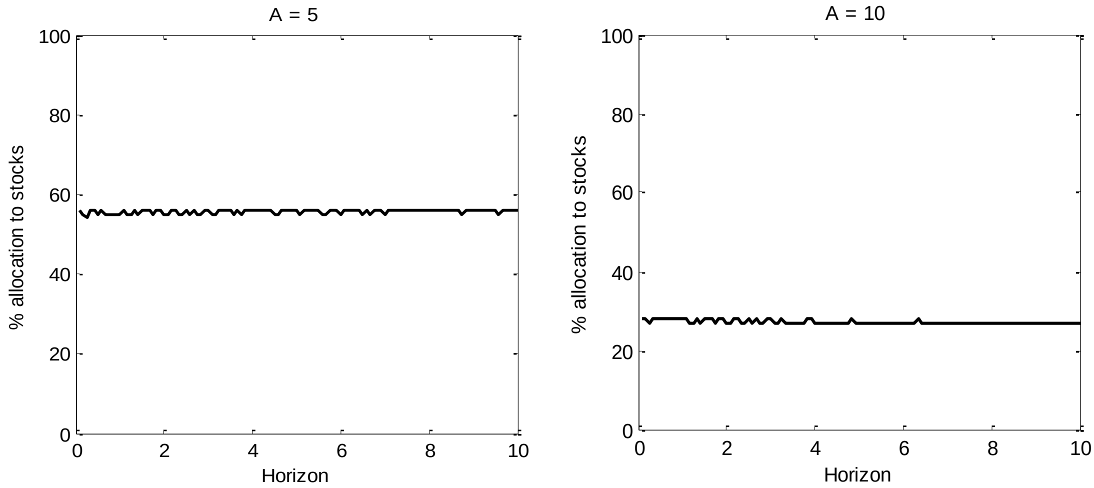

Figure 2.10: *Optimal allocation to stocks for a buy-and-hold investor with power utility function. The percentage invested in stocks is plotted against the investment horizon in years. The graph on the left side corresponds to a level of risk-aversion of 5, the graph on the right to a level of risk-aversion of 10.*

When the investor ignores parameter uncertainty, he uses a Normal distribution with mean $\hat{T}\hat{\mu}$ and variance $\hat{T}\hat{\sigma}^2$ in order to forecast log cumulative returns. We indeed assume that continuously compounded excess stock returns are homoscedastic, uncorrelated**.** Both the mean and the variance grow linearly with the investor's horizon $\hat{T}$. A natural consequence of this is that the investor chooses the same stock allocation, regardless of the holding period.

When $A = 5$, the optimal percentage ω that the investor allocates to the stock index is 56% , whereas for an investor with $A = 10$ the percentage allocated to the stock index falls to 28%. As the level of relative risk-aversion increases, the allocation to the stock index falls, indeed a conservative investor prefers a portfolio where the risk-free asset constitutes the main proportion.

Ignoring parameter uncertainty

In this section we try to show how the allocation differs when parameter uncertainty is explicitly incorporated into the investor's decision making framework. When he takes into account parameter uncertainty, he solves the maximization problem (2.11), throughout the application of the predictive distribution $p(R_{T+\hat{T}} \mid r)$ conditional only on past data.

Figure 2.11 shows that in this context, the stock allocation falls as the horizon increases. Therefore we note that parameter uncertainty can introduce horizon effect even in the context of i.i.d. model returns. Accounting for estimation risk, the investor's distribution for long-horizon returns incorporates an extra degree of uncertainty, involving an increase in its variance. Moreover, this extra uncertainty makes the variance of the distribution for cumulative returns increase faster than linearly with the horizon $\hat{T}$. This makes stocks appear riskier to long-horizon investors, who therefore reduce the amount they allocate to equities in favor of risk-free asset.

The explanation why variances increase faster than linearly with the horizon is because, in the presence of parameter uncertainty, returns are no longer i.i.d. form perspective of the investor, but rather positively serially correlated. An important source of uncertainty in the parameters surrounds the mean of the stock return. Returns are positively serially correlated in the sense that, if the stock return is high over the first month, then it will probably be high over the second month because it is likely that the state of world is one with a high realization of the uncertain stock mean parameter μ.

The magnitude of the effects included by parameter uncertainty are meaningful. An investor using the full data set, with $A = 5$, at an investment horizon of one month allocates to the stock index 56%, the same portion he would have invested ignoring parameter uncertainty . On the other hand, after ten years the percentage allocated to the stock index falls to 41% , a difference of more than 10 percent . When the level of risk-aversion grows to 10 , the difference in allocation at a 10-year horizon becomes 8%, that is, the investor passes from an allocation to stocks equal to 28% to an allocation of 20%.

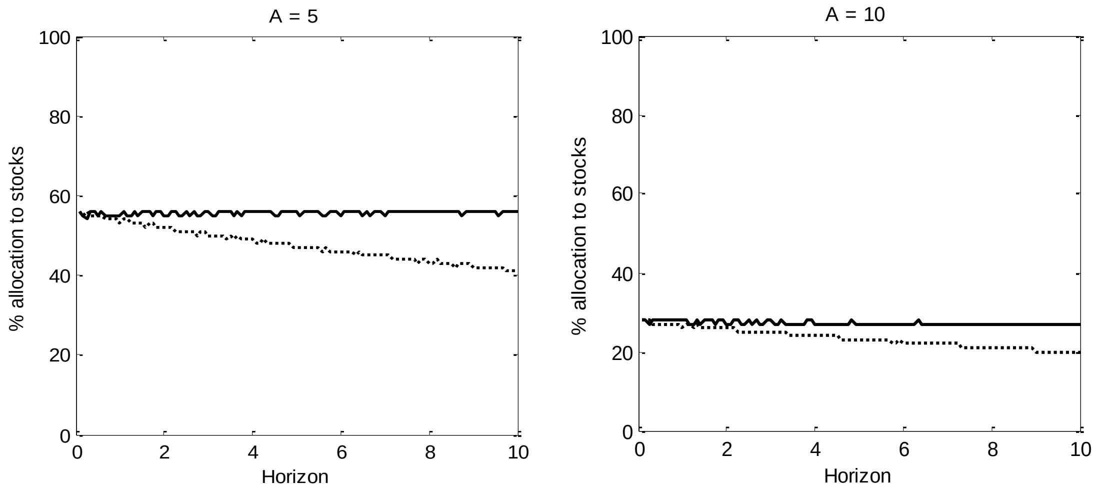

Figure 2.11: *Optimal allocation to stocks for a buy-and-hold investor with power utility function. The percentage invested in stocks is plotted against the investment horizon in years. The solid line refers to the case where the investor ignores parameter uncertainty, the dot line to the case where he accounts for it. The graph on the left side corresponds to a level of risk-aversion of 5, the graph on the right to a level of risk-aversion of 10.*

2.6 Resampling

Thus far we hypothesized that returns, calculated assuming continuous compounding, had a Normal distribution, or more precisely that the model for excess return over month t, $r_t = \mu + \varepsilon_t$, whit $\varepsilon_t \sim \text{i.i.d.}N(0,\sigma^2)$, held true. This assumption of normality, attributed to the assets returns, has a critical role in the construction of the model; however, as we can note from the preliminary analysis of S&P 500 stock price index, there are empirical reasons to believe that it does not represent an adequate description of the returns' generator process. After testing for normality we ended up rejecting the null hypothesis of normal returns for a significance level of 0.05. Furthermore the distribution of future excess returns is unknown since we do not know the future realizations of the stock index. In this section we find a way to obtain the empirical distribution of future excess returns using past data to simulate future returns from the available sample. This technique is defined as resampling. It only allows the assumption that all sample data have the same probability to occurring, no additional hypothesis is made. This method can be implemented by constructing a number of resamples of the observed dataset of excess returns ,of equal size to the observed dataset, each of which is obtained by random sampling with replacement from the original dataset. This process is repeated thousands of times in order to generate a probability distribution anchored to the true but unknown distribution of returns. The cumulative excess returns over $\hat{T}$ periods are simply the sum of $\hat{T}$ samples generated using a resempling method.

When we want the sampling method to take into account parameters uncertainty, we firstly sample from the standardized returns

$$u_t = \frac{r_t - \mu}{\sigma},$$

so that the mean is equal to 0 and the variance is equal to 1. Every drawn value is then multiplied by a value $\hat{\sigma}$ obtained from the posterior distribution of σ^2, an added to a value $\hat{\mu}$ obtained from the posterior distribution of μ .

Comparing the optimal allocations obtained assuming normally distributed cumulative excess returns to the allocations obtained resampling the excess returns, we get a measure of the sensitivity of the results to departures form the normality hypothesis.

2.6.1 Results

Figure 2.12 shows the optimal portfolio allocation for a buy-and-hold investor, whose preferences over terminal wealth are described by a power utility function. The optimal percentage ω allocated to the stock index, is plotted as a function of the investment horizon, that range from 1 month to 10 years. The graphs on the left side refer to the analysis that assumes normally distributed excess returns based on the model $r_t = \mu + \varepsilon_t$, where $\varepsilon_t \sim \text{i.i.d.}N(0,\sigma^2)$. The graphs on the right side allude to the case where excess returns are generated by resampling.

The graphs in figure 2.12 exhibit a strong similarity. Whether the investors take into account the estimation risk or ignore it, the optimal allocations obtained under the hypothesis of normality are essentially the same as the ones obtained by resampling

Although we rejected the null hypothesis of normality for the distribution of excess returns, the optimal allocation does not appear to be affected by this assumption. In the next analysis we will therefore keep on hypothesizing a normal distribution of excess returns .

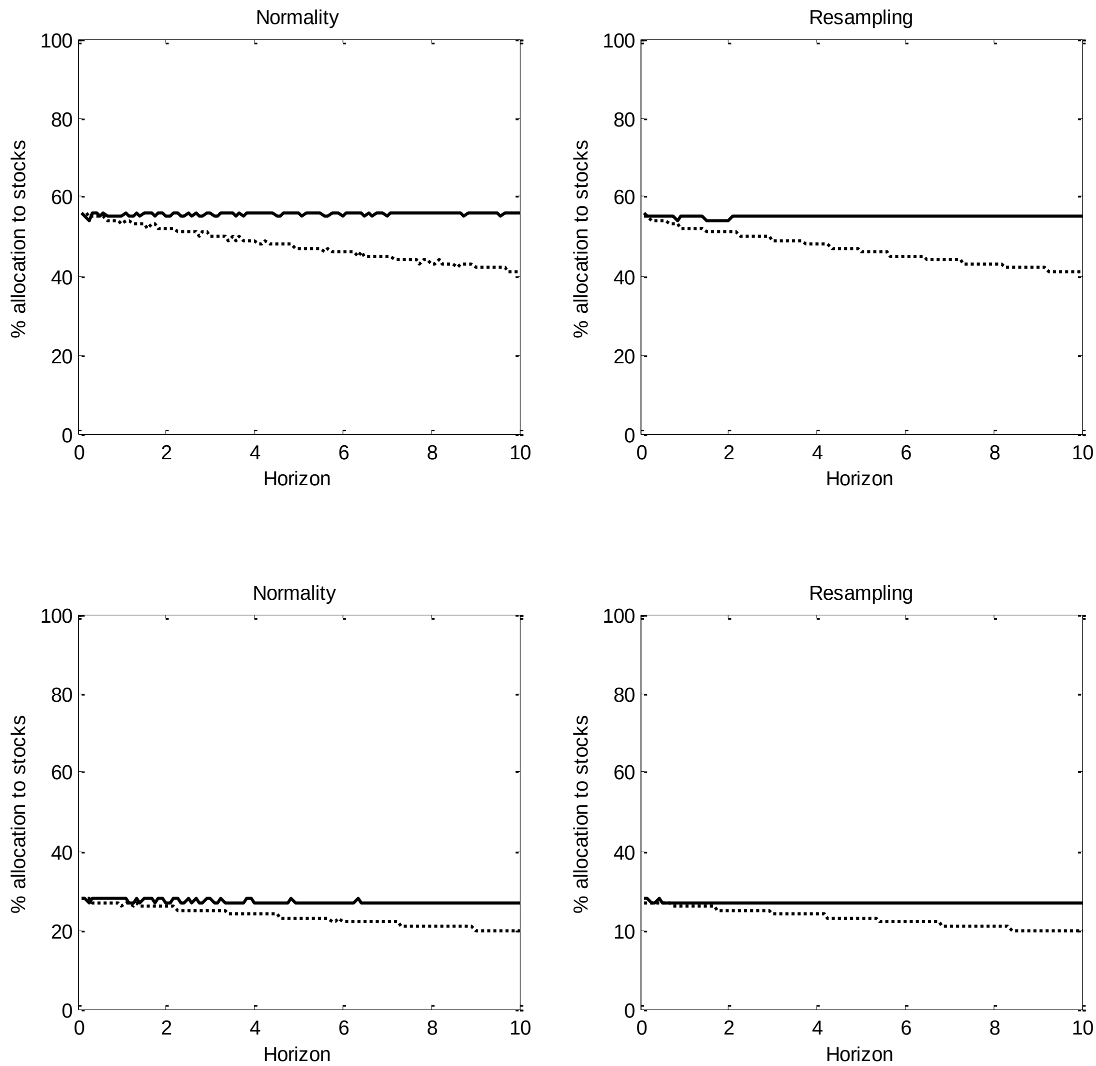

Figure 2.12: *Optimal allocation to stocks for a buy-and-hold investor with power utility function. The percentage invested in stocks is plotted against the investment horizon in years. The two upper graphs correspond to the case where the investor's level of risk aversion equals 5, the graphs below to the case where his level of risk aversion equals 10.*

Chapter 3

Portfolio allocation with predictable returns

3.1 Introduction

This chapter focuses on how predictability affects portfolio choice. An important aspect of this analysis is that in constructing optimal portfolios, we account for the fact that the true extent of predictability in returns is highly uncertain.

For the study of the predictability of excess stock returns only one variable is considered, the dividend yield, which is introduced and analyzed in the third paragraph.

A VAR model is then defined in order to examine how the evidence of predictability in asset returns affects optimal portfolio choice.

In the sixth paragraph we describe the procedure to incorporate parameter uncertainty in the portfolio allocation problem and then explain the sampling process, a critical step in computing optimal allocations.

In the seventh paragraph the results of the optimal portfolio allocation for a buy-and-hold investor who is risk-averse are presented. To see whether predictability in returns has any effect on portfolio choice we compare the allocation of an investor who recognizes predictability to that of an investor who is blind to it.

Finally, the results obtained considering different initial values of the dividend yield are reported in order to understand the role of the predictor variable.

3.2 Returns predictability

Economists have long been concerned by the nature of variations in the stock market. By the early 1970's a consensus emerged among financial economists suggesting that stock prices could be well approximated by random walk model, and that changes in stock returns were basically unforecastable. Samuelson (1965) showed that in an informationally efficient market, price changes must be unpredictable. However, random walk model had been around for many years; having been originally discovered by Louis Bachelier back in 1900. The main idea behind the random walk theory is that investors react instantaneously to any informational advantages they have, eliminating therefore profit opportunities. Thus, prices must fully reflect the information available in the market and no profit can be made from information based trading.

However, recently there has been an emergence of counter arguments. One branch of the literature asserts that expected returns contain a time-varying component that implies predictability of future returns. Recent researches (Keim and Stambaug (1984), Campbell (1984), Fama and French (1989)) have drawn attention to the ability of some economic variable, to partially predict stock and bond returns and interest rates.

Typical predictor variables employed in this kind of researches are financial ratios, such as the dividend-price ratio, the earnings-price ratio, and the book-to-market ratio, which have a quantity that represent the market in the denominator; but also measures of equity risk such as squared returns, or interest rates measures which capture the level or slope of the interest rates' term structure and finally financial and economic variables such as the inflation rate. Depending on their nature these variables can capture variation throughout time of expected excess returns or variation in the variance and also in the covariance matrix.

3.3 Predictor variable: dividend yield

Given actual historical data on asset returns and a predictor variable, we try to understand the magnitude of these effects by computing optimal asset allocation for an investor who adopts a static buy-and-hold strategy, and whose preference over terminal wealth are described by a constant relative risk-aversion power utility function.

We firstly develop the model using only one predictor variable in order to describe return's dynamics. In this section we use dividend/price , henceforth called dividend yield, to forecast returns on the value-weighted index S&P 500, for return horizons ranging from one month to ten years. Dividend yield occupies a salient role in much of the empirical literature on the predictability of stock returns. As discussed by Keim and Stambaugh (1986) , given that asset's current price is inversely related to the discount rate applied to expected future cash flows, variables that are inversely related to price levels, such as the dividend yield, are suitable candidates ex ante as predictors for returns. At high frequency, dividends are smooth relative to stock prices, so the dividend yield displays a strong inverse association with the level of equity prices and thereby arises as a plausible predictor variable.

There is however evidence that dividend yields forecast stock return in Rozeff (1984), Shiller(1984), Flood, Hodrick, and Kaplan (1986), Campbell and Shiller (1988), and Fama and French (1988b).

We downloaded the S&P 500 monthly dividend yield time series from Datastream database for the period Jenuary 1990-November 2012.

3.3.1 Preliminary Analysis

As we can see in figure 3.1 the dividend yield time series is non-stationary and has frequent changes in mean, this is further confirmed by the autocorrelogram in figure 3.2. Furthermore when we implement the Dickey-Fuller test without constant, since it is not significant, we obtain a value of the t-statistic equal to -1.32, which is smaller, in absolute value, than the critical value -1.95 and confirm the presence of unit root at a significance level of 0.05. In developing a VAR model only stationary variables should be taken into account, unfortunately the dividend yield does not exhibit this property. Stambaugh (1999) asserts that using highly persistent variables in a VAR model can lead to small biases in the coefficients' estimate if the sample size is not large enough. He finds that the bias has an opposite sign to the correlation between innovations in excess returns and dividend yield. He also notes that the bias disappear as this correlation approaches zero. The non stationarity can clearly have some effects on the values of the estimated coefficients, but for now no correction is made. Estimated values will be treated as given and known by investors, or alternatively, the uncertainty in the VAR model parameters will be taken into account so as to not give too much credit to the particular estimated value.

Dividend Yield			
Mean	0.021151	St. Error	0.000390
Minimum	0.010800	Variance	0.000042
Maximum	0.040300	St. Dev	0.006457
1° Quartile	0.017000	Skewnees	0.655035
3° Quartile	0.026600	Kurtosis Excess	-0.163063

Table 3.1: *Main descriptive statistics of S&P 500 dividend yield over the period 1990-2012.*

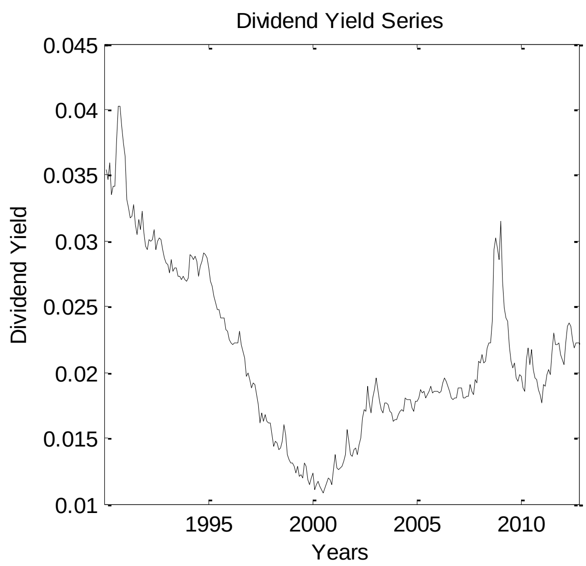

Figure 3.1: *S&P 500 dividend yield over the period 1990-2012.*

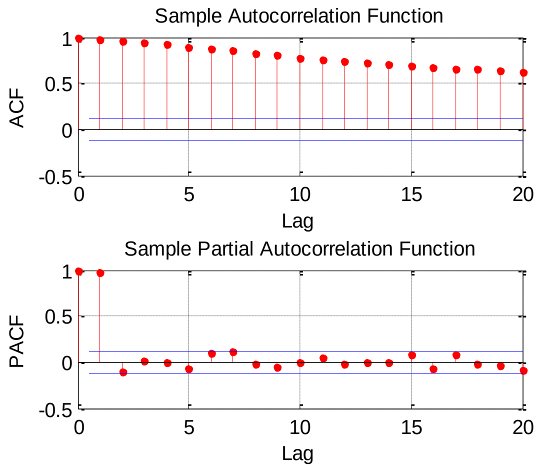

Figure 3.2: *Empirical correlogram of the S&P 500 dividend yield series.*

The normality tests implemented easily reject the null hypothesis of normality for a significance level of 0.05. We can indeed observe that the empirical density function of the dividend yield differs from the normal probability density evaluated by using the sample mean and standard deviation, in particular it has two significant peaks. The probability plot of Figure 3.4 further confirms this result.

Normality Test	
Jarque-Bera	0.002
Lilliefors	0.001

Table 3.2: *Normality tests' P-values for the dividend yield series..*

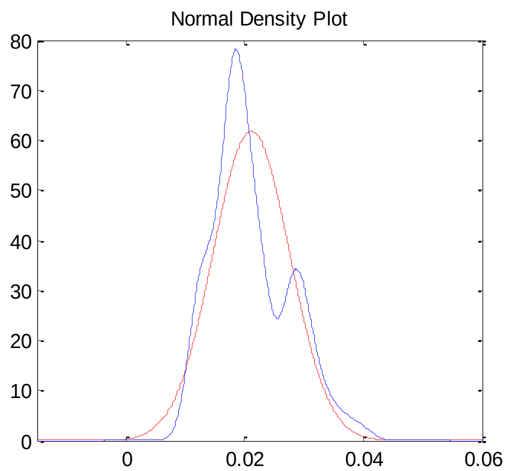

Figure 3.3 : *Empirical density function of the S&P 500 dividend yield series and normal probability density function evaluated by using the sample mean and standard deviation.*

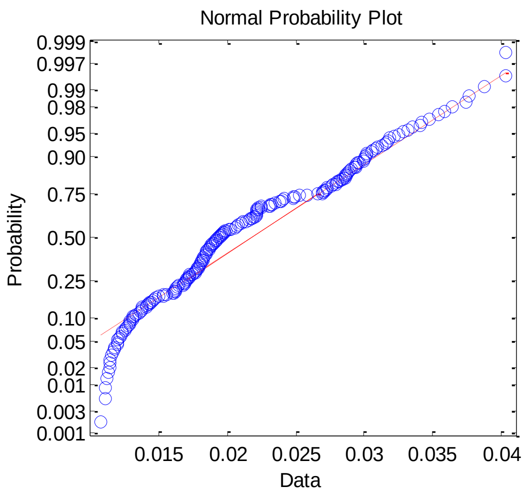

Figure 3.4: *Normal probability plot of S&P 500 dividend yield series.*

3.4 Long horizon predictability and parameter uncertainty

In light of the growing evidence that returns are predictable, the investor's horizon may be highly relevant. It has been known since Samuelson and Merton that variation in expected returns over time can potentially introduce horizon effects. Time-variation in returns can therefore invalidate the assumptions under which a long-term investor acts myopically, choosing the same portfolio as a short-term investor

The extent to which the holding period does play a role serves as an interesting and convenient way of thinking about how predictability affects portfolio choice. Moreover, the results may shed light on the common but controversial advice that investors with long horizons should allocate more heavily on stocks.

An important aspect of our analysis is that in constructing optimal portfolios, we account for the fact that the true extent of predictability in returns is highly uncertain. This is of particular concern in this context because the evidence of time variation in expected returns is sometimes weak. A typical example is the following. Denote by rt the continuously compounded return on the value-weighted index S&P500 in month *t*, and by dy_{t-1} be the portfolio's dividend yield in month $t-1$. An OLS regression of the returns on the lagged dividend yield, using monthly returns from January 1990 to November 2012, gives

$$\underset{}{r_t} = \underset{(0.0097)}{-0.0072} + \underset{(0.4602)}{0.6171}\ dy_{t-1} + \varepsilon_t, \tag{3.1}$$

where standard errors are in parentheses and the R^2 is 0.0029 . The coefficient on the dividend yield is not quite significant, and the R^2 is very low. Some investors might react to the weakness of this evidence by discarding the notion that returns are predictable; others might instead ignore the substantial uncertainty regarding the true predictive power of the dividend yield and analyze the portfolio problem assuming that parameters are known precisely. However, the optimal stock-versus-cash allocation of the investor can depend importantly on the current value of a predictive variable, such as the dividend yield, even though a null hypothesis of no predictability might not be rejected at conventional significance levels. The approach we choose in our work could be considered as a middle ground: we explicitly account for the uncertainty about the parameters, also known as estimation risk, when constructing optimal portfolios.

How is parameter uncertainty incorporated? It is natural to take a Bayesian approach here. The uncertainty about the parameters of the predictive variables is summarized by the posterior distribution of parameters given the data. Rather than constructing the distribution of future returns conditional on fixed parameter estimates, we integrate over the uncertainty in the parameters captured by the posterior distribution.

It may be important that the investor take into account uncertainty about the model parameters such as the coefficient on the predictor variable in equation (3.1). The standard errors in equation (3.1) indicate that the true forecasting ability of the dividend yield may be much weaker than that implied by the raw parameter estimate. The investor's portfolio decisions can be improved by adopting a framework that recognizes this.

Our framework assumes a risk-averse investor with initially vague beliefs about the distribution of stock returns. The investor uses the above regression evidence to update those beliefs, and these revised beliefs are then used by the investor to compute the optimal asset allocation. We find that the asset allocation chosen by the investor depends importantly on the level of the current dividend yield.

3.5 Predictability analysis model

To examine how the evidence of predictability in asset returns affects optimal portfolio choice we analyze a vectorial autoregressive process, VAR. Barberies (2000) develops this model that it is suitable to describe the dynamic behavior of stocks returns. The model is similar in structure to the one implemented by Kandel and Stambaugh (1991), Cambpbell (1991), and by Hodrick (1992).

The investor uses a VAR model to forecast returns, where the state vector in the VAR can include asset returns and predictors variables. This is a convenient framework for examining how predictability affects portfolio choice: by changing the number of predictor variables in the state vector, we can compare the optimal allocation of an investor who takes return predictability into account to that of an investor who is blind to it. In the calculations presented in this section, the vector z_t contains only two components: the excess stock index return r_t, and a single predictor variable, the dividend yield $x_{1,t}$, which captures an important component of the variation in expected returns. Hereinafter in this work we will take into account other variables. The model takes this form

$$z_t = a + Bx_{t-1} + \varepsilon_t, \tag{3.2}$$

with $z_t' = (r_t, x_t')$, $x_t = (x_{1,t}, ..., x_{n,t})'$, since the number of predictor variable n is equal to 1 $x_t = x_{1,t}$, and $\varepsilon_t \sim \text{i.i.d.} N(0, \Sigma)$.

The first component of z_t, namely r_t, is the continuously compounded excess return over month t. The remaining components of z_t, which together make up the vector or explanatory variables x_t, consist of variables useful for predicting returns, such as the dividend yield. The first equation in the system specifies expected stock returns as a function of the predictor variables. The other equations specify the stochastic evolution of the predictor variables. Considering the dividend yield as the sole predictor variable the model takes this form:

$$
\begin{aligned}
r_{t+1} &= a_1 + b_1 x_{1,t} + \varepsilon_{1,t+1}\,, \\
x_{1,t+1} &= a_2 + b_2 x_{1,t} + \varepsilon_{2,t+1}\,,
\end{aligned}
\tag{3.3}
$$

where

$$
\begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} \sim N\left(0, \begin{pmatrix} \sigma_1^2 & \sigma_{12}^2 \\ \sigma_{12}^2 & \sigma_2^2 \end{pmatrix}\right).
$$

The variance-covariance matrix of contemporaneous innovations Σ is invertible and not necessarily diagonal; thus we allow the shocks to be cross-sectionally correlated, but assume that they are homoscedastic and independently distributed over time. The hypothesis of homoscedasticity is of course restrictive. It rules out the possibility that the predictor variables predict change in risk; they can affect portfolio choice only by predicting changes in expected return. However, even though the assumption of homoscedasticity is not entirely realistic, empirical evidence suggests that changes in risk is a short-lived phenomenon that does not affect the long-term portfolio choice(Chacko e Viceira,1999).

The model we handle is not exactly a first order VAR, since all the variables here evaluated should also depend on the lagged value of r_t . Basically we analyze a VAR(1) model with some restrictions on its parameters , indeed we can write:

$$
z_t = a + B_0 z_{t-1} + \varepsilon_t\,, \tag{3.4}
$$

Where B_0 is a square matrix and its first column contains only zeros so that z_t does not depend on r_{t-1} .

$$
B_0 = \begin{bmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} & B \end{bmatrix}.
$$

3.6 Long horizon portfolio allocation

In this section we introduce the impact of predictability as well as of parameter uncertainty, that we already analyzed in chapter 2. We implement the VAR model illustrated before in order to explore how the evidence or predictability in returns affects optimal portfolio choice. Our pursue is to study the portfolio allocation problem for a buy-and-hold investor with an investment horizon of $\hat{T}$ months.

We now rewrite the model in a more convenient way:

$$\begin{pmatrix} z_2' \\ \vdots \\ z_T' \end{pmatrix} = \begin{pmatrix} 1 & x_1' \\ 1 & \vdots \\ 1 & x_{T-1}' \end{pmatrix} \begin{pmatrix} a' \\ B' \end{pmatrix} + \begin{pmatrix} \varepsilon_2' \\ \vdots \\ \varepsilon_T' \end{pmatrix}, \tag{3.5}$$

or

$$Z = XC + E, \tag{3.6}$$

where Z is a $(T-1, n+1)$ matrix with the vectors $z_2',...,z_T'$ as rows; X is a $(T-1, n+1)$ matrix with the vectors $(1\ x_1'),...,(1\ x_{T-1}')$ as rows, and E is a $(T-1, n+1)$ matrix with vectors $\varepsilon_2',...,\varepsilon_T'$ as rows. Instead C is a $(n+1)(n+1)$ matrix . Since in this section we study the predictive effect of one variable only, n equals 1, and matrix C takes this form:

$$\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}$$

where the first row contains the intercepts and the second one contains the coefficients of x_{t-1} .

Now, we want to write down the problem faced at time T by a buy-and-hold investor with a horizon of $\hat{T}$ months , given by the next equation. Since he has no chance to buy or sell assets between time T and horizon $T+\hat{T}$, he is interested only in the distribution of wealth at the end of the holding period. The investor problem therefore concerns the maximization of his expected utility defined over final wealth.

$$\max_{\omega} E_T \left(\frac{\left\{ (1-\omega)\exp(r_f \hat{T}) + \omega \exp(r_f \hat{T} + R_{T+\hat{T}}) \right\}^{1-A}}{1-A} \right). \tag{3.7}$$

E_t denotes the fact that the investor calculates the expectation conditional on his information set at time T. We have therefore to define which distribution the investor should use in calculating this expectation. If we consider the case when the investor recognizes predictability, there are two possible distributions he can use when computing the expectation in equation (3.7) depending on whether he accounts for parameter uncertainty or not.

Ignoring parameter uncertainty

We evaluate the model $z_t = a + Bx_{t-1} + \varepsilon_t$. When the uncertainty in the model parameters is ignored the investor uses the distribution of future returns conditional on both past data and fixed parameters values $\theta = (a, B, \Sigma)$, $p(R_{T+\hat{T}} \mid \theta, z)$. Once the parameters estimates have been obtained from the posterior distribution, it is generated a distribution for future stock excess returns conditional on a set of parameter values and on the data observed by the investor up until the start of his investment horizon, which we write as $p(R_{T+\hat{T}} \mid z, \hat{\theta})$, where $z = (z_1, ..., z_T)'$ is the data observed by the investor until the start of his investment horizon .
The investor then solves:

$$\max_{\omega} \int \upsilon(W_{T+\hat{T}})\, p(R_{T+\hat{T}} \mid z, \hat{\theta})\, dR_{T+\hat{T}}. \tag{3.8}$$

In order to define the cumulative excess returns conditional distribution $p(R_{T+\hat{T}} \mid z, \hat{\theta})$ we can write the model as $z_t = a + B_0\, z_{t-1} + \varepsilon_t$, therefore

$$
\begin{aligned}
z_{T+1} &= a + B_0 z_T + \varepsilon_{T+1} \\
z_{T+2} &= a + + B_0 a + B_0^2 z_T + \varepsilon_{T+2} + B_0 \varepsilon_{T+1} \\
&\vdots \\
z_{T+\hat{T}} &= a + + B_0 a + B_0^2 a + \ldots + B_0^{T-1} a \\
&\quad + B_0^{\hat{T}} z_T \\
&\quad + \varepsilon_{T+\hat{T}} + B_0 \varepsilon_{T+\hat{T}-1} + B_0^2 \varepsilon_{T+\hat{T}-2} + \ldots + B_0^{\hat{T}-2} \varepsilon_{T+2} + B_0^2 \varepsilon_{T+1}.
\end{aligned} \tag{3.9}
$$

Conditional on a, B and Σ the sum $Z_{T+\hat{T}} = z_{T+1} + z_{T+2} + \ldots + z_{T+\hat{T}}$ is Noramlly distributed with mean and variance given by:

$$
\begin{aligned}
\mu_{sum} &= \hat{T} a + (\hat{T}-1) B_0 a + (\hat{T}-2) B_0^2 a \\
&\quad + \ldots \\
&\quad + B_0^{\hat{T}-1} a + (B_0 + B_0^2 + \ldots + B_0^{\hat{T}}) z_T ,
\end{aligned} \tag{3.10}
$$

$$
\begin{aligned}
\Sigma_{sum} &= \Sigma \\
&\quad + (I + B_0) \Sigma (I + B_0)' \\
&\quad + (I + B_0 + B_0^2) \Sigma (I + B_0 + B_0^2)' \\
&\quad \vdots \\
&\quad + (I + B_0 + \ldots + B_0^{\hat{T}-1}) \Sigma (I + B_0 + \ldots + B_0^{\hat{T}-1})'.
\end{aligned} \tag{3.11}
$$

In this case, we assume that the distributions for future returns are $N(\mu_{sum}, \Sigma_{sum})$, where μ_{sum} and Σ_{sum} are constructed using the posterior means of a, B and Σ as fixed values.

Incorporating parameter uncertainty

In contrast, when we take parameter uncertainty into account we refer to a Bayesian approach. Zellner (1971) discusses the Bayesian analysis of a multivariate regression model in the traditional case with exogenous regressors. The form of the likelihood function is the same in the cases of endogenous regressors , so long as we condition on the first observation in the sample, z_1 . Therefore we can take advantages of his analysis for our dynamic regression framework with endogenous regressors.

Throughout a posterior distribution $p(\theta | z)$ we summarize the uncertainty about the parameters $\theta = (a, B, \Sigma)$ given the observed data.

To construct the posterior distribution $p(a, B, \Sigma | z)$ we consider, as in the previous section, an uninformative prior as

$$p(C, \Sigma) \propto |\Sigma|^{-(n-2)/2}.$$

the posterior $p(C, \Sigma^{-1} | z)$ is then given by

$$\Sigma^{-1} | z \sim Wishart(T-n-2, S^{-1})$$

$$vec(C) | \Sigma, z \sim N(vec(\hat{C}), \Sigma \otimes (X'X)^{-1})$$

where $S = (Z - X\hat{C})'(Z - X\hat{C})$ with $\hat{C} = (X'X)^{-1} X'Z$.

Integrating over this distribution, we obtain the predictive distribution for long-horizons returns. This distribution is conditioned only on the sample observed, and not on any fixed a, B and Σ.

$$p(R_{T+\hat{T}} | z) = \int (R_{T+\hat{T}} | z, \theta)\, p(\theta | z)\, d\theta. \tag{3.12}$$

The problem the investor has to solve is then

$$\max_{\omega} \int \upsilon(W_{T+\hat{T}})\, p(R_{T+\hat{T}} | z)\, dR_{T+\hat{T}}. \tag{3.13}$$

Or alternatively

$$\begin{aligned} &\max_{\omega} \int \upsilon(W_{T+\hat{T}})\, p(R_{T+\hat{T}}, \theta | z)\, dR_{T+\hat{T}}\, d\theta \\ &= \max_{\omega} \int \upsilon(W_{T+\hat{T}})\, p(R_{T+\hat{T}} | z, \theta)\, p(\theta | z)\, dR_{T+\hat{T}}\, d\theta. \end{aligned} \tag{3.14}$$

Excess returns distribution conditional on a set of parameter values and on the observed data is given by

$$Z_{T+\hat{T}} | C, \Sigma, z \sim N(\mu_{sum}, \Sigma_{sum}). \tag{3.15}$$

where μ_{sum} and Σ_{sum} are computed using the estimated parameters of the posterior distribution.

In order to solve the maximization problem (3.8) and (3.13) we calculate the integrals for several values of the proportions invested in the equity index, that is $\omega = 0, 0.01, 0.02, ..., 0.98, 0.99$ and report the ω that maximizes expected utility. We calculate the optimal allocation ω which maximizes expected utility for a variety of risk aversion levels A and investment horizons ranging from 1 month to 10 years, and for each of the two cases where the investor either ignores or account for parameter uncertainty.

We chose to avail ourselves of the interactive environment of numerical computation MATLAB in order to implement the model described before. The employed commands are listed in Appendix B.

3.6.1 Sampling process

The next few paragraphs explain how we sample from the predictive distribution, an important step in computing these optimal allocations.

The procedure for sampling is similar to that in the second chapter. Firstly, we generate a sample of size I=200000 from the posterior distribution for the parameters $p(a, B, \Sigma \mid z)$. We sample from the posterior distribution by first drawing from the marginal $p(\Sigma^{-1} \mid z)$, Wishart, and then given the $\hat{\Sigma}$ drawn, from the conditional $p(vec(C) \mid \hat{\Sigma}, z)$, a Normal distribution. We therefore generate a sample of size 200000 from the posterior distribution for C and Σ. Repeating this 200000 gives an accurate representation of the posterior distribution. Secondly, for each of the 20000 realizations of the parameters $(\hat{C}, \hat{\Sigma})$ in the sample from the posterior $p(a, B, \Sigma \mid z)$, we sample once from the distribution of returns conditional on both pasta data and the parameters $p(Z_{T+\hat{T}} \mid \hat{C}, \hat{\Sigma}, z)$, a Normal distribution of cumulative returns conditional on past data and on parameters $\hat{C}$ and $\hat{\Sigma}$. This gives us a sample of size 200000 form the predictive distribution for returns, conditional only on past returns, with the parameter uncertainty integrated out.

In contrast when parameter uncertainty is ignored we assume that the distributions for future returns are constructed using the posterior means of $\hat{a}$, $\hat{B}$ and $\hat{\Sigma}$ as the fixed values of the parameters, and then drawing 200000 times from the Normal distribution with mean and variance given by equations (3.10) and (3.11) above.

3.7 Results

To see whether predictability in returns has any effect on portfolio choice of a buy-and hold investor, our strategy is to compare the allocation of an investor who recognizes predictability to that of an investor who is blind to it. The VAR model provides a convenient way of making this comparison because by simply altering the number of predictor variables included in the vector x_t, it simulates investors with different information sets.

In this section we compute the optimal allocations ω which maximize the quantity in expression (3.7) for a variety of risk aversion levels A and investment horizons $\hat{T}$, and for different cases where the investor either ignores or accounts for parameter uncertainty.

The results are based on the model $z_t = a + Bx_{t-1} + \varepsilon_t$, where $z_t = (r_t, x_t)'$ includes continuously compounded monthly excess stock returns r_t and the dividend yield $x_{1,t}$, and where $\varepsilon_t \sim \text{i.i.d.}N(0,\Sigma)$.

Table 3.3 presents the mean and standard deviation (in parentheses) of the posterior distribution $p(C,\Sigma \mid z)$ for each parameter a, B and Σ.

The predictive power of the dividend yield is summarized in the first row of the B matrix. We note that the posterior distribution for that coefficient has mean 0.5420 and standard deviation 0.4117, which appears to be an important source of parameter uncertainty for the investor. Moreover the second row of the B matrix confirms us the high persistency of the dividend yield, that we already mentioned before. The variance matrix shows the strong negative correlation between innovations in stock returns and the dividend yield, estimated here at -0.7940; this correlation has an important influence on the distribution of long-horizon returns. Indeed if the dividend yield falls unexpectedly, since $\sigma_{12}^2 < 0$, it is likely to be accompanied by a contemporaneous positive shock to stock returns. However, since the dividend yield has fallen, stock returns are forecasted to be lower in the future, since $b_1 > 0$. This rise, followed by a fall in returns generate a component of negative serial correlation in returns which slows the evolution of the variance of cumulative returns as the horizon grows.

1990-2012	
a	B
-0.0071 (0.0091	0.5420 (0.4117)
0.0004 (0.0002	0.9805 (0.0088)
Σ	
0.0019 (0.0002	-3.28e-05 (3.23e-06)
	8.82e-07 (7.66e-08)

Table 3.3: *Mean and standard deviation (in parenthesis) of each parameter's posterior distribution.*

The aim of this section is to understand how predictability in asset returns and parameter uncertainty affects portfolio choice. To do this, we compute optimal allocation using four different choices for the distribution of future returns. These distributions differ in whether they take into account predictability and estimation risk. In the second chapter we explored the issue of parameter uncertainty in the context of i.i.d. returns. Here we want to see whether predictability in returns has any effect on portfolio choice throughout the implementation of a VAR model. In any case the investor may account for parameter uncertainty in the model, and thus use a predictive distribution of the form $p(R_{T+\hat{T}} \mid z)$, or he may ignore parameter uncertainty in the model; in this case we assume that the distribution for future returns are constructed using the posterior means of a , B and Σ, given in Table 3.3, as the fixed values of the parameters.

Ignoring parameter uncertainty

When the investor ignores parameters uncertainty , he solves the maximization problem (3.8), employing a distribution for future excess returns conditional on the estimated parameter values and on the observed data of this form $p(Z_{T+\hat{T}} \mid \hat{a}, \hat{B}, \hat{\Sigma})$, which is normally distributed with mean $\hat{\mu}_{sum}$ and variance $\hat{\Sigma}_{sum}$. The investor's distribution for future returns of course depends on the value of the dividend yield at the beginning of the investment horizon, $x_{1,T}$. If the value of the yield is low, this forecasts low returns, lowering the mean of the distribution for future returns and reducing the allocation to the stock index. In our set of result we set the initial value of the dividend yield to its mean in the sample, namely $x_{1,T} = 2.12\%$, in order not to consider the impact of the initial value in the portfolio choices, and investigate how the optimal allocation changes with the investor's horizon for this fixed initial value of the predictor.

Figure 3.5 shows the optimal portfolio allocation for a buy-and-hold investor, whose preferences over terminal wealth are described by a power utility function. The optimal percentage ω allocated to the stock index is plotted as a function of the investment horizon that range from 1 moth to 10 years. The graph on the left is based on a relative risk-aversion level of 5 , the one on the right side is for $A = 10$. The two lines on each graph correspond to the two possible distributions the investor could use once the fact he ignores parameter uncertainty is assumed. The black line represents the case where the investor ignores predictability, that is when he assumes $r_t = \mu + \varepsilon_t$, with $\varepsilon_t \sim \text{i.i.d.}N(0, \sigma^2)$; on the other hand the green line represents the case where the investor uses a VAR model which allows for predictability in returns.

We can note that the green line, that represents the optimal percentage ω allocated to the stock index when the investor takes into account predcitability, rises dramatically as the investment horizon increases. Hence, when we ignore parameter uncertainty about the model parameters, the optimal allocation to equities for a long-horizon investor is much higher than for a short-horizon investor. When we acknowledge that returns may be predictable rather than i.i.d., the mean and variance of cumulative returns may not grow linearly with the investor horizon $\hat{T}$ anymore, as in the case when asset returns are modeled as i.i.d. In the context of predictability in returns the variance of cumulative stock returns may grow slower than linearly with the investor's horizon, lowering the perceived long-run risk of stocks and hence leading to higher allocations to stocks in the optimal portfolio.

This point can be verify mathematically, performing the matrixial calculation of Σ_{sum} described in equation (3.11) . For instance, the conditional variances of one- and two-period cumulative stock returns are

$$Var_T(r_{T+1}) = \sigma_1^2 , \tag{3.16}$$

$$Var_T(r_{T+1} + r_{T+2}) = 2\sigma_1^2 + b_1^2\sigma_2^2 + 2b_1\sigma_{12}^2 . \tag{3.17}$$

If we plug in the parameter values estimated from the data, the posterior means in table 3.3 , we find that $b_1^2\sigma_2^2 + 2b_1\sigma_{12} < 0$,which implies that the conditional variance of two-period returns is less than twice the conditional variance of one-period returns. When we take into account the predictive power of the dividend yield, conditional variances greow more slowly than linearly with the investor's horizon, making stocks look relatively less risky at longer horizon and increasing their optimal weight in the investor's portfolio.

The insight behind this result can partially be explained by the effect of the negative correlation between innovations in stock returns and the dividend yield, that has already been described above. However the results obtained here should not be considered as being specific to the particular way we have modeled returns, nor to the specific parameter values estimated from the data. There is a strong economic intuition behind the concept that time variation in expected returns induces mean-reversion in realized returns. The essence of this concept is the assumption that both a stock's high and low returns are temporary and stock's returns will tend to move to the average over time. Or even, when there is a positive shock to expected returns, it is very reasonable that realized returns should suffer a contemporaneous negative shock since the discount rate for discounting future cash flows has suddenly increased. This negative shock to current realized returns, followed by higher forecasted returns, are the provenience of mean-reversion, which in turn makes stocks more appealing in the long run.

In his study, Barberies underlines the fact that horizon effects can be present even without negative serial correlation in returns. He asserts that the predictability in returns may be sufficient to make stocks more attractive at long horizons, without being strong enough to induce mean-reversion in returns.

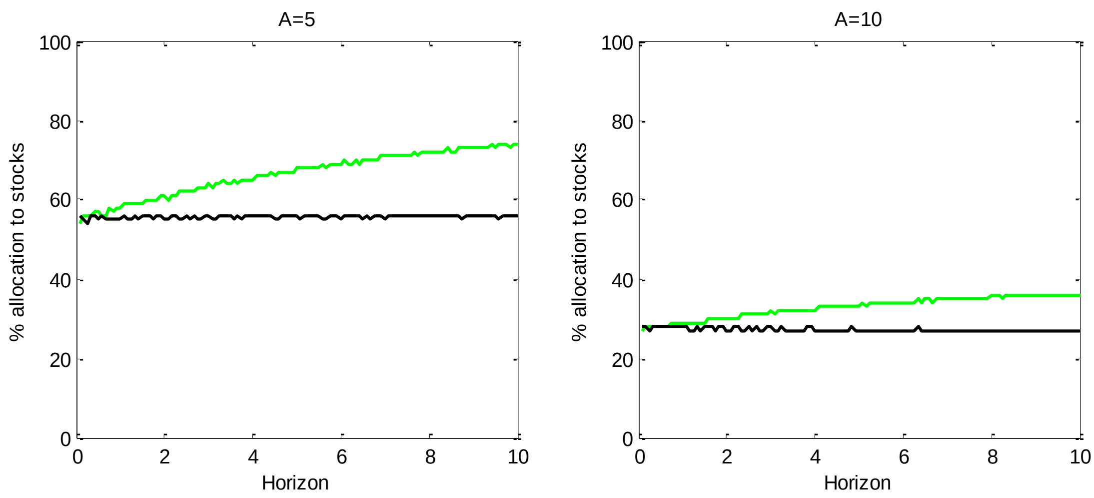

Figure 3.5: *Optimal allocation to stocks for a buy-and-hold investor with power utility function. The percentage invested in stocks is plotted against the investment horizon in years. The green line refers to the cases where the investor accounts for predictability, the black line to the cases where he ignores it. The graph on the left side corresponds to a level of risk-aversion of 5, the graph on the right to a level of risk-aversion of 10.*

Incorporating parameter uncertainty

In this section we try to show how the allocation differs when parameter uncertainty is explicitly incorporated into the investor's decision making framework. Our strategy for understanding the effect of parameter uncertainty is to compare the allocation of an investor who uses the predictive distribution to forecast returns with the allocation of an investor who uses instead the distribution of returns conditional on fixed parameters $\hat{a}$, $\hat{B}$ and $\hat{\Sigma}$.

Figure 3.6 shows the optimal portfolio allocation for a buy-and-hold investor, whose preference over terminal wealth are described by a power utility function. The optimal percentage ω allocated to the stock index is plotted against the investment horizon that range from 1 moth to 10 years. The four lines in the graphs correspond to the four possibilities for the distribution of future returns, depending on whether the investor allows for predictability and parameter uncertainty. The dotted lines correspond to cases where investor accounts for parameter uncertainty, the solid ones to cases where he ignores it. The green lines refers to the case where the investor accounts for predictability, whereas the black one to the case where the investor is blind to it.

Figure 3.6 shows that when we account for predictability and parameter uncertainty together, there is still horizon effect, in other words, the optimal allocation changes as the investment horizon increases. However the long-horizon allocation is again higher than the short-horizon allocation, but not nearly as much higher as when we ignore estimation risk. We can deduce that incorporating parameter uncertainty can considerably reduce the size of the horizon effect. Moreover in this case the optimal allocation to equities is not monotonic anymore, we can indeed observe that it first rises with the investment horizon , and then it starts falling as the investment horizon grows. In any case it always remains under the optimal allocation of an investor who assumes that asset returns are modeled as i.i.d. , and above the allocation of this investor when he takes parameter uncertainty into account. But we need to bear in mind that the posterior distribution for b_1 has a meaningful standard deviation of 0.4117.

This effect firstly arises from the investor’s uncertainty about the mean stock return. Exactly in the same way of chapter 2 , incorporating the uncertainty about the mean makes conditional variances grow faster as the horizon increases, making stocks look more risky and inducing a lower allocation to stocks compared to the case where estimation risk is ignored. Moreover the true predictive power of the dividend yield is uncertain to the investor; therefore it is also uncertain whether the dividend yield really does slow the evolution of conditional variances, and hence whether stocks’ riskiness diminish with the horizon. The investor acknowledge both that the predictive power may be weaker than the point estimate suggests, in which case he would be more cautious to allocate more to stocks at long horizons, and that it may be stronger, in which case he would be enthusiastic to allocate more to stocks at longer horizons. These effects go on opposite directions. On net, the investor invests less at long horizon because he is risk-averse. Other two effects go on opposite direction, accounting for predictability makes stocks look less risky at long horizons; whether incorporating the estimation risk makes them look more risky, this therefore lead, which is the case, to stock allocations that are not monotonic as a function of the investment horizon.

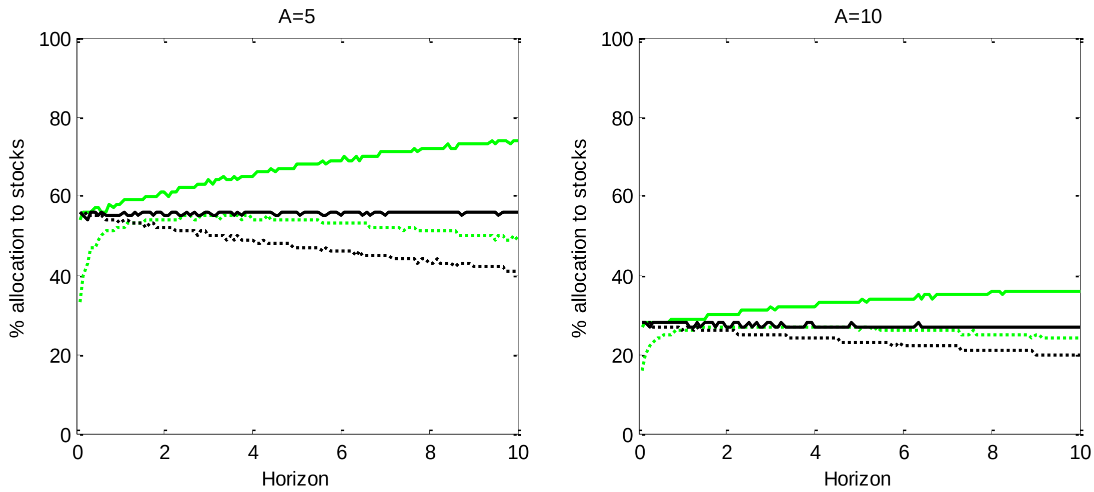

Figure 3.6: *Optimal allocation to stocks for a buy-and-hold investor with power utility function. The percentage invested in stocks is plotted against the investment horizon in years. The green lines refer to the cases where the investor accounts for predictability, the black lines to the cases where he ignores it. The solid lines refers to the case where the investor ignores parameter uncertainty, the dot line to the cases where he accounts for it. The graph on the left side corresponds to a level of risk-aversion of 5, the graph on the right to a level of risk-aversion of 10.*

Up to this point we have examined just one consequence of including the dividend yield as a predictor variable in the VAR. Taking into account the predictive power of the dividend yield reduces the variance of predicted long-horizon cumulative returns, , lowering the perceived long-run risk of stocks and hence leading to higher allocations to stocks for long horizon investor. Conditioning on the dividend yield affects not only the conditional variance but also the mean of the cumulative excess returns. Indeed, when the dividend yield is low relative to its historical mean, an investor forecasts lower than average stock returns and hence reduces his allocation to stocks. This effect has not been taken into account so far because the initial value of the dividend yield has been kept fixed at its sample mean.

In this section we examine the results on the optimal portfolio allocation for different initial values of the dividend yield $x_{1,T}$. Figure 3.7 presents the optimal allocations, estimated running a regression over the period 1990 to 2012 . The graphs on the left side refer to the case where the investor ignores parameter uncertainty, the one on the right to the case where he accounts for it. Each graph exhibits the optimal stock allocation as a function of the investor's horizon for five different initial values of the predictor variable: the historical mean of the dividend yield in our sample, the first and third quartile and the 37.5% and 67.5% percentiles.

Both graphs on the left side show that for all the initial values of the dividend yield considered, the allocation to stocks rises with the investor's horizon. The result we obtained earlier in this section continues therefore to hold. Moreover, for any fixed horizon, the optimal allocation to stocks is higher for higher values of the predictor variable. Since the dividend yield affects the mean of the distribution for future returns, the investor expects higher future returns when the dividend yield is high. Besides, we can notice that the optimal stock allocation of an investor with 10-year horizon is just as sensitive to the initial value of the dividend yield $x_{1,T}$ as the optimal allocation of a one-year horizon investor. So, the various allocation do not converge to a specific value in the long run.

The two graphs on the right illustrate the optimal allocation to stocks when parameter uncertainty is incorporated. The results are extremely different from the previous one, when parameter uncertainty is ignored. At low value of the dividend yield, the stock allocation is generally increasing in the investment horizon, whereas that allocation is generally decreasing in the horizon at higher dividend yield. In other words, the allocation of an investor with a 10-year horizon is less sensitive to the initial value of the predictor variable than the allocation of a one-year horizon investor, and much less sensitive than the allocation of a 10-year horizon investor who ignores parameter uncertainty. The allocation lines show therefore sign of converging. It is reasonable to think that the degree of predictability of returns in more distant future months is less than in nearby months, the effect of the initial value of the dividend yield on future expected returns therefore diminishes as the investment horizon grows.

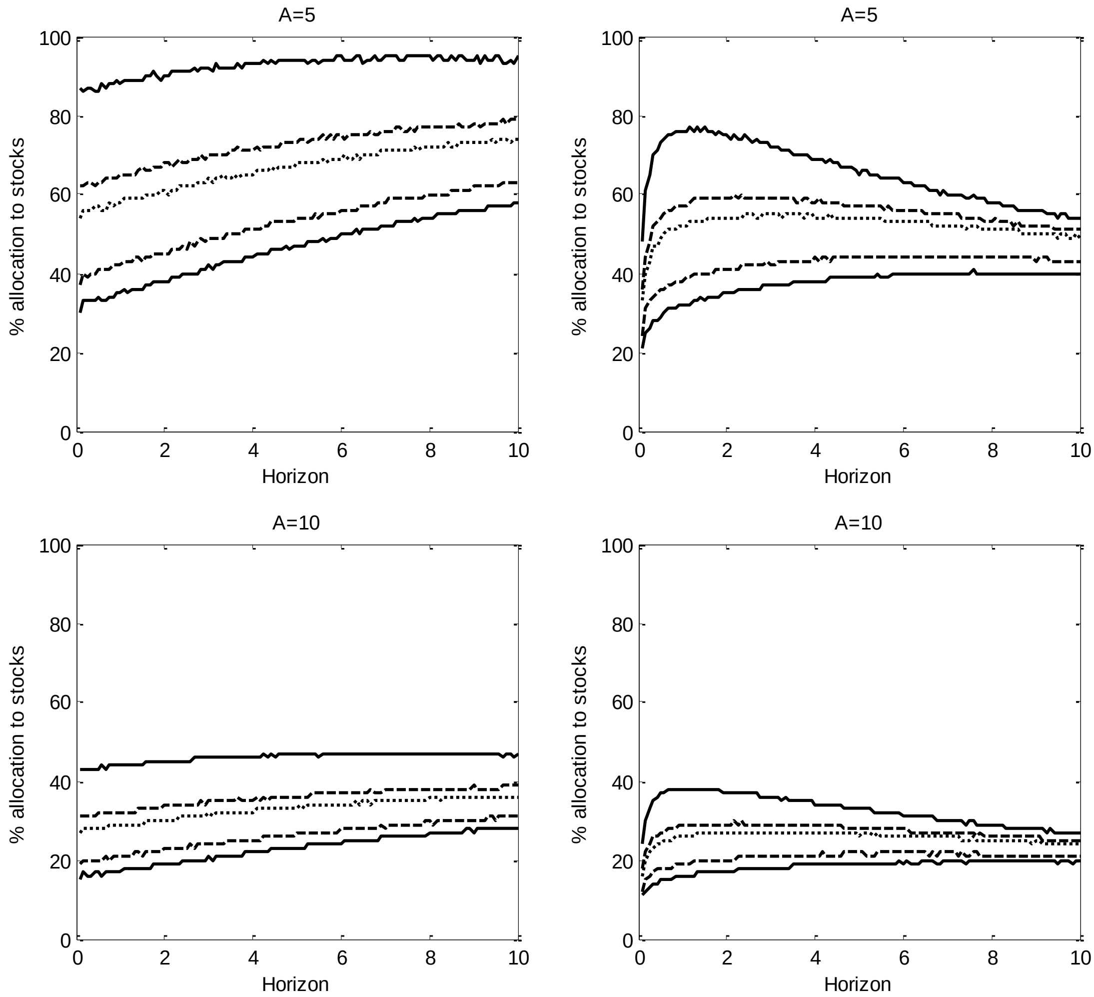

Figure 3.6: *Optimal allocation to stocks for a buy-and-hold investor with power utility function. The percentage invested in stocks is plotted against the investment horizon in years. The two graphs on the left ignore parameter uncertainty, the ones on the right account for it. The five lines within each graph correspond to different initial value of the dividend yield:* $x_t = 1.17\%$ *(solid line),* $x_t = 1.83\%$ *(dashed line),* $x_t = 2.12\%$ *(dotted line),* $x_t = 2.23\%$ *(dashed line),* $x_t = 2.66\%$ *(solid line).*

Chapter 4

Portfolio allocation with parameter uncertainty: two risky assets

4.1 Introduction

In this chapter we introduce some extensions to the model implemented thus far. We indeed want to study the optimal portfolio allocation when investors can choose how to allocate their wealth among three different assets: a stock index, a bond index, and the risk-free asset . Our purpose is similar to the one of the second chapter, that is to understand how parameter uncertainty alone affects portfolio choice.

The third paragraph is dedicated to the description of the new dependent variable, the bond index, and to some preliminary analysis.

Some changes to the framework presented in the second chapter are then implemented, and an adequate model, that deals with portfolio choice under the case where the investor either ignores or accounts for parameter uncertainty, is defined. We then explain the sampling process needed to implement this model.

In the sixth paragraph the results of the optimal portfolio allocation for a buy-and-hold investor who is risk-averse are presented.

4.2 An extra risky asset: the bond index

In Modern portfolio theory it is described how an investor may alter risk and return of a portfolio by changing the mix of assets. In particular, according to the Mean-Variance Analysis, the investor chooses his appropriate optimal portfolio, combination of risk–free asset and optimal risky portfolio, maximizing his own satisfaction. In the previous chapters we assumed that the stock index was the only risky asset available to the investor, therefore simplifying his decision process.

Indeed he was only asked to choose the proportion to be allocated to the stock index and the one to be allocated to the risk-free asset. From now on, we devote our attention to studying the optimal portfolio allocation for a buy-and-hold investor who is allowed to allocate his wealth among two risky assets, the stock index and the bond index, and a risk-free one. Adding another risky asset, the bond index, the investor can achieve any combination of risk and return along the efficient frontier by changing the proportion of stocks and bonds.

Bonds are the most important financial assets competing with stocks, they promise fixed monetary payments over time. In contrast to equity, the cash flows from bonds have a maximum monetary value set by the terms of the contract and except in the case of default, bond returns do not vary with the profitability of the firm. That said, an investor could consider it advantageous to allocate all his wealth in these debt instruments, however, we already said in the first chapter that although it might appear safer to accumulate wealth in bonds rather than in stocks over long periods of time, precisely the opposite seems to be true. As Siegel asserts, standard deviation is higher for stock returns than for bond returns over short-term holding period, but once the holding period increases, bonds become riskier than stocks. He finds that the probability that stocks outperform fixed income assets increases dramatically with the holding period, although in the short run bonds and even bank accounts outperform stocks with a high probability. Even though over long periods returns on bonds fall short of that on stocks, bonds may still serve to diversify a portfolio and lower the overall risk.

In the next two chapter we intend to explore the issue of portfolio allocation among three assets, the stock index, the bond index and the risk-free asset. In particular our purpose is to throw light on the commonly held view that investors with long horizon should allocate more heavily on stocks. We desire to investigate the question in a broader context, compared to the one of the previous chapters, where the investor is now allowed to choose how to invest his wealth between two risky assets; and we want therefore to observe how the addition of a risky asset affects optimal portfolio choice.

The long term debt instrument we employ in order to carry out our analysis is the 20 years U.S. Treasury bond index downloaded from Datastream data set. We compute the bond index returns starting from the monthly total returns time series and assuming continuous compounding.

4.2.1 Preliminary analysis

Bond index total return time series is non-stationary and a trend in mean is easily identifiable in Figure 4.1. . The logarithmic returns calculated starting from this series are stationary, and the autocorrelation function in Figure *** is a confirmation of that. They have a positive mean of 0.0074, and it is significantly different from zero, since the t-test, obtained from the ratio between returns' mean and the corresponding standard error, is equal to 3.82.

20 U.S. Year Treasury Bond Logarithmic Returns			
Mean	0.007418	St. Error	0.001941
Minimum	-0.141319	Variance	0.001032
Maximum	0.136941	St. Dev	0.032133
1° Quartile	-0.010567	Skewnees	0.055695
3° Quartile	0.026208	Kurtosis Excess	3.182042

Table 4.1: *Main descriptive statistics of the 20 Year U.S. Year Treasury Bond Logarithmic Returns over the period 1990-2012.*

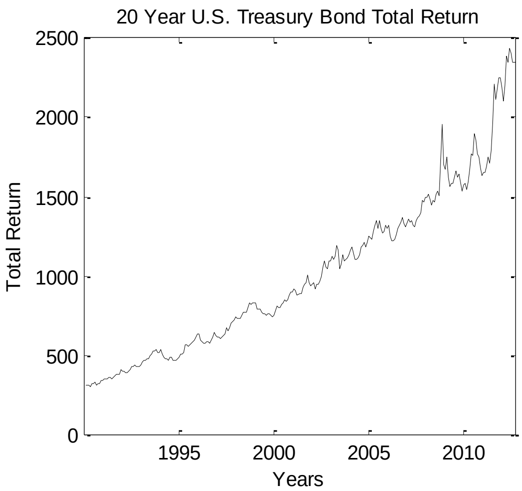

Figure 4.1: *20 Year U.S. Treasury Bond Total Return series over the period 1990-2012.*

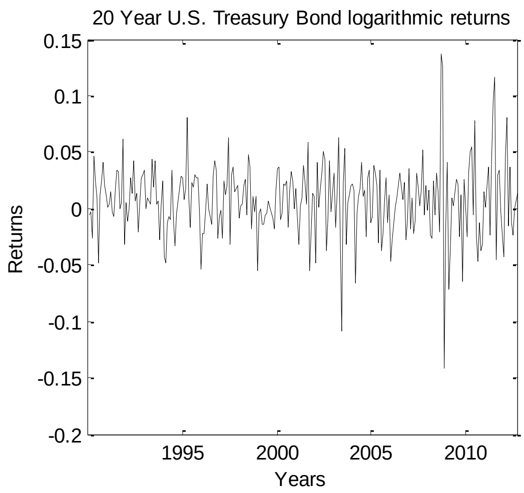

Figure 4.2: *20 Year U.S. Treasury Bond returns series over the period 1990-2012*

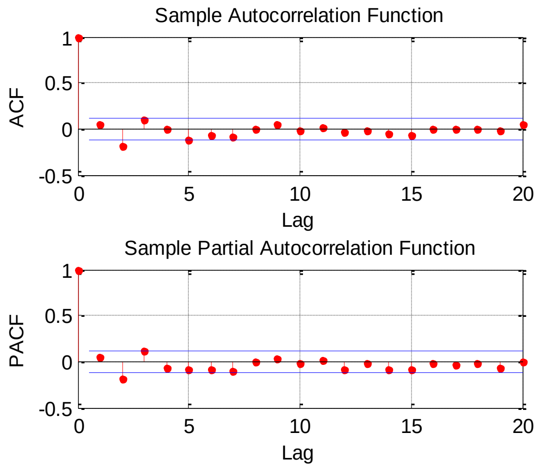

Figure 4.3: *Empirical correlogram of 20 Year U.S. Treasury Bond return series*

As we have already mentioned in the second chapter there are empirical reasons to believe that assumption of i.i.d. normal returns, that is behind several models, does not represent an appropriate description of the returns' generator process. Taking a look at Figure 4.4, we can see how the empirical density function of the returns series moves away from the normal probability density function evaluated by using the sample mean and standard deviation, we can moreover recognize a skewness on the left side. This is furthermore confirmed by the normality tests implemented, that reject the null hypothesis of normality. If we look then at the normal probability plot in Figure 4.5 we notice a strong departure of the empirical queues from the theoretical ones.

Normality Test	
Jarque-Bera	0.001
Lilliefors	0.0236

Table 4.2: *Normality tests' P-values for the bond returns series.*

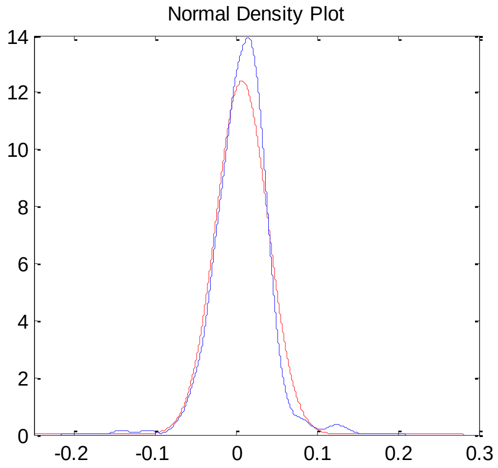

Figure 4.4: *Normal density plot of the 20 Year U.S Treasury Bond returns series.*

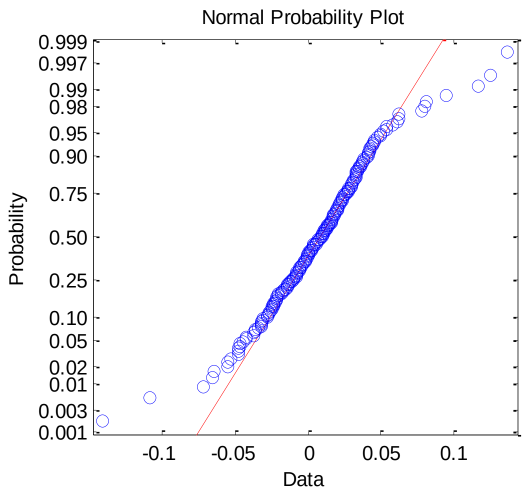

Figure 4.5: *Normal probability plot of the 20 Year U.S Treasury Bond returns series.*

4.3 Model with two risky assets

In this section we introduce some extensions to the model developed by Barberies (2000), that deals with the portfolio choice under several investment horizons.
Some changes to the initial model should be made in order to evaluate how portfolio choice changes when the investor can choose how to allocate his wealth no longer between two alternatives but rather among three different assets: the risk-free asset, an equity index and a bond index. We firstly consider the case where no predictor variables are included in the model, and afterword we focus on a more generic model.

As we did when only one risky asset was available, we begin our analysis by considering the context where no predictor variables are included in the model, and hence where asset returns are i.i.d., and look at how parameter uncertainty alone affects portfolio allocation. There are three assets: Treasury bill, an equity index, in this case the value-weighted index S&P 500, and a bond index , the 20-Year U.S. Treasury Bond. As before, for the sake of simplicity, we suppose that the continuously compounded monthly return on Treasury Bills is a constant r_f, and that the excess returns on the risky assets, obtained as difference between the returns and r_f, are continuously compounded. We therefore assume a normal distribution for the excess returns.

We therefore model excess returns on the stock and bond indexes assuming that they i.i.d., so that

$$\begin{aligned} r_{1,t} &= a_{r_1} + \varepsilon_{r_1,t} \\ r_{2,t} &= a_{r_2} + \varepsilon_{r_2,t}, \end{aligned} \tag{4.1}$$

Where $r_{1,t}$ is the continuously compounded excess return on the equity index over month t , $r_{2,t}$ is the continuously compounded excess return on the bond index over month t and where $\varepsilon_t \sim \text{i.i.d.}N(0,\Sigma)$. The variance matrix of contemporaneous innovations is invertible and unexpected excess returns realizations are allowed to covariate among them. Moreover, as we did in the third chapter we assume that Σ does not vary over time.

In matrix notation the model becomes:

$$r_t \;=\; a \;+\; \varepsilon_t\,, \tag{4.2}$$

with $r_t' = (r_{1,t}, r_{2,t})$, $a' = (a_{r_1}, a_{r_2})$ and $\varepsilon_t \sim \text{i.i.d.} N(0,\Sigma)$

and, if we consider the entire time series, takes this form:

$$\begin{pmatrix} r'_2 \\ \vdots \\ r'_T \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \left(a'\right) + \begin{pmatrix} \varepsilon'_2 \\ \vdots \\ \varepsilon'_T \end{pmatrix}, \tag{4.3}$$

or

$$R \;=\; IM \;+\; E\,, \tag{4.4}$$

where R is a $(T-1,2)$ matrix with the vectors $r_2',...,r_T'$ as rows; I is a $(T-1,1)$ vector of ones, M is a $(1,2)$ matrix containing the means of the process, and E is a $(T-1,2)$ matrix with vectors $\varepsilon_2',...,\varepsilon_T'$ as rows.

Although now we are not focused on studying the predictability dynamics of assets returns, we can rewrite the model in a different way, that henceforth will turn out to be useful . We therefore consider a model of the form:

$$r_t \;=\; a \;+\; B_0 r_{t-1} \;+\; \varepsilon_t\,, \tag{4.5}$$

with $r_t' = (r_{1,t}, r_{2,t})$, $a' = (a_{r_1}, a_{r_2})$, $\varepsilon_t \sim \text{i.i.d.} N(0,\Sigma)$ and B_0 equal to

$$B_0 \;=\; \begin{bmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} & B \end{bmatrix}. \tag{4.6}$$

where matrix B does not exist since we are not taking predictability into account , and therefore B_0 is a two columns matrix of zeros.

This is a convenient framework. In fact the state vector in the model could include not only assets returns, but predictor variables as well. All that is needed to move from the i.i.d. context to the one of predictability, is to add some predictor variables in the state vector and to change the composition of B matrix, and consequently of B_0. By changing the number of predictor variables in the state vector, we can indeed compare the optimal allocation of an investor who takes return predictability into account to that of an investor who is blind to it.

4.4 Long horizon portfolio allocation

We devote this section to analyzing the portfolio choice under several investment horizons and under the case where the investor either ignores or accounts for parameter uncertainty. To do this we employ the model described in the previous section where assets returns are assumed to be i.i.d.

The purpose of our analysis is to determine the optimal portfolio allocation for a buy-and-hold individual with a horizon of $\hat{T}$ months. Since the investor has no chance to buy or sell securities between time horizon T and $T+\hat{T}$, he is only interested in the distribution of wealth at the end of the investment period, that is $u(W_{T+\hat{T}})$. We employ the same utility function used in the previous chapters, that is the power utility function, which has a constant coefficient of relative risk aversion. The investor's preferences on final wealth in $T+\hat{T}$ are then described by a power utility function of the form

$$u(W) = \frac{W^{1-A}}{1-A} \tag{4.7}$$

where A is the coefficient of relative risk aversion.

Since the investor can now allocates his wealth among three different assets, instead of only two alternatives, the end-of-horizon wealth is thus given by

$$\begin{aligned} W_{T+\hat{T}} &= (1-\alpha-\beta)\exp(r_f\hat{T}) \\ &\quad + \alpha\exp(r_f\hat{T} + r_{1,T+1} + \ldots + r_{1,T+\hat{T}}) \\ &\quad + \beta\exp(r_f\hat{T} + r_{2,T+1} + \ldots + r_{2,T+\hat{T}}). \end{aligned} \tag{4.8}$$

Assumed initial wealth $W_T = 1$, α being the allocation to the stock index and β the allocation to the bond index.

If we write the cumulative excess stock return over $\hat{T}$ periods as

$$R_{1,T+\hat{T}} = r_{1,T+1} + r_{1,T+2} + \ldots + r_{1,T+\hat{T}}, \qquad (4.9)$$

and the cumulative bond excess return over $\hat{T}$ periods as

$$R_{2,T+\hat{T}} = r_{2,T+1} + r_{2,T+2} + \ldots + r_{2,T+\hat{T}}, \qquad (4.10)$$

the buy-and-hold investor's problem is to solve

$$\max_{\alpha,\beta} E_T \left(\frac{\left\{ W_{T+\hat{T}} = (1-\alpha-\beta)\exp(r_f\hat{T}) + \alpha\exp(r_f\hat{T} + R_{1,T+\hat{T}}) + \beta\exp(r_f\hat{T} + R_{2,T+\hat{T}}) \right\}^{1-A}}{1-A} \right).$$

(4.11)

The investor calculates the expectation conditional on his information set at time T. We have therefore to define once again which distribution he should use in calculating this expectation, depending on whether he accounts for parameter uncertainty or not. The effect of parameter uncertainty can then be studied by comparing the optimal portfolio allocation obtained in these two cases.

Ignoring parameter uncertainty

We evaluate the model $r_t = a + \varepsilon_t$. When the uncertainty in the model parameters is ignored the investor uses the distribution of future returns conditional on both past data and fixed parameters values $\hat{\theta} = (\hat{a}, \hat{\Sigma})$. Once the parameters estimates have been obtained from the posterior distribution, it is generated a distribution for future stock and bond excess returns conditional on a set of parameter values and on the data observed by the investor up until the start of his investment horizon, which we write as $p(R_{T+\hat{T}} \mid r, \hat{\theta})$, where we denote by $R_{T+\hat{T}} = (R_{1,T+\hat{T}}, R_{2,T+\hat{T}})'$ the cumulative excess returns of stocks and bonds and by $r = (r_1, \ldots, r_T)'$ the data observed by the investor until the start of his investment horizon.

The investor then solves:

$$\max_{\alpha,\beta} \int \upsilon(W_{T+\hat{T}})\, p(R_{T+\hat{T}} \mid r, \hat{\theta})\, dR_{T+\hat{T}} \,. \tag{4.12}$$

Since the model employed here is $r_t = a + \varepsilon_t$, we have that $R_{T+\hat{T}} = r_{T+1} + r_{T+2} + ... + r_{T+\hat{T}}$ is the sum of $\hat{T}$ bivariate normal random variables with mean a and variance Σ, the sum $R_{T+\hat{T}}$ is therefore normally distributed conditional on a and Σ with theoretical mean $\hat{T}\,a$ and variance $\hat{T}\,\Sigma$.

Alternatively, if we write the model as $r_t = a + B_0 r_{t-1} + \varepsilon_t$, with B_0 void matrix, we have that

$$\begin{aligned}
r_{T+1} &= a + B_0 r_T + \varepsilon_{T+1} \\
r_{T+2} &= a + + B_0 a + B_0^2 r_T + \varepsilon_{T+2} + B_0 \varepsilon_{T+1} \\
&\vdots \\
r_{T+\hat{T}} &= a + + B_0 a + B_0^2 a + ... + B_0^{T-1} a \\
&\quad + B_0^{\hat{T}} r_T \\
&\quad + \varepsilon_{T+\hat{T}} + B_0 \varepsilon_{T+\hat{T}-1} + B_0^2 \varepsilon_{T+\hat{T}-2} + ... + B_0^{\hat{T}-2} \varepsilon_{T+2} + B_0^2 \varepsilon_{T+1} \,.
\end{aligned} \tag{4.13}$$

The sum $R_{T+\hat{T}} = r_{T+1} + r_{T+2} + ... + r_{T+\hat{T}}$ conditional on a, B_0 and Σ is Normally distributed with mean and variance given by:

$$\mu_{sum} = \hat{T}a + (\hat{T}-1)\,B_0 a + (\hat{T}-2)\,B_0^2 a + ... + B_0^{\hat{T}-1} a + (B_0 + B_0^2 + ... + B_0^{\hat{T}})\, r_T \,, \tag{4.14}$$

$$\begin{aligned}
\Sigma_{sum} &= \Sigma \\
&\quad + (I + B_0)\Sigma(I + B_0)' \\
&\quad + (I + B_0 + B_0^2)\Sigma(I + B_0 + B_0^2)' \\
&\quad \vdots \\
&\quad + (I + B_0 + ... + B_0^{\hat{T}-1})\Sigma(I + B_0 + ... + B_0^{\hat{T}-1})'.
\end{aligned} \tag{4.15}$$

Assuming that B_0 is a void matrix, these ones become:

$$\mu_{sum} = \hat{T}a \tag{4.16}$$

$$\Sigma_{sum} = \hat{T}\Sigma \qquad (4.17)$$

that is exactly the same result obtained before.

Incorporating parameter uncertainty

Differently, when we take parameter uncertainty into account we refer to Zellner's Bayesian approach (1971). Throughout a posterior distribution $p(\theta | z)$ we summarize the uncertainty about the parameters $\theta = (a, \Sigma)$ given the observed data. To construct the posterior distribution $p(a, \Sigma | z)$ we consider, as we did in the third chapter, an uninformative prior of the form

$$p(\mathrm{M}, \Sigma) \propto |\Sigma|^{-1/2}.$$

The resulting posterior distribution consists of the marginal distribution

$$\Sigma^{-1} | z \sim Wishart(T-2, S^{-1})$$

and of the conditional Normal distribution

$$vec(M) | \Sigma, r \sim N(vec(\hat{M}), \Sigma)$$

where $S = (R - I\hat{M})'(R - I\hat{M})$ with $\hat{M} = I'R$.

Integrating over this distribution, we obtain the so-called predictive distribution for long-horizons returns, as we did when we considered a single risky asset. This distribution is conditioned only on the observed sample, and not on any fixed a and Σ .

$$p(R_{T+\hat{T}} | r) = \int (R_{T+\hat{T}} | r, \theta)\, p(\theta | r)\, dR_{T+\hat{T}}\, d\theta. \qquad (4.18)$$

The problem the investor has to solve is then

$$\max_{\alpha,\beta} \int \upsilon(W_{T+\hat{T}})\, p(R_{T+\hat{T}} | z)\, dR_{T+\hat{T}}. \qquad (4.19)$$

Or alternatively

$$\max_{\alpha,\beta} \int \upsilon(W_{T+\hat{T}})\, p(R_{T+\hat{T}}, \theta \mid r)\, dR_{T+\hat{T}}\, d\theta$$
$$= \max_{\alpha,\beta} \int \upsilon(W_{T+\hat{T}})\, p(R_{T+\hat{T}} \mid r, \theta)\, p(\theta \mid z)\, dR_{T+\hat{T}}\, d\theta . \tag{4.20}$$

As the decomposition in equation (4.20) shows, we sample from the joint distribution by first sampling from the posterior $p(\theta \mid r)$ and then from the conditional $p(R_{T+\hat{T}} \mid r, a, \Sigma)$, a $N(\hat{T}\, a, \hat{T}\, \Sigma)$.

The problem of expected utility maximization is solved calculating the integrals (4.12) and (4.19) for several combinations of α and β, the proportion invested in the equity index and the one invested in the bond index respectively. In other words we compute the integrals for all the available combinations of $\alpha = 0,\ 0.01,\ 0.02,\ ...,\ 0.98,\ 0.99$ and $\beta = 0,\ 0.01,\ 0.02,\ ...,\ 0.98,\ 0.99$ subject to $0 \le \alpha + \beta \le 1$ and report α and β that maximize expected utility. We therefore restrict the allocation to the interval $0 \le \alpha + \beta \le 1$ precluding short selling and buying on margin. For each of the two cases where the investor either ignores or account for parameter uncertainty, we calculate the optimal proportions α and β, which maximize expected utility for a variety of risk aversion levels A and investment horizons ranging from 1 month to 10 years.
The integrals themselves are evaluated numerically by simulation, generating 200000 values from the distributions defined earlier.

We chose to avail ourselves of the interactive environment of numerical computation MATLAB in order to implement the model described before. The employed commands are listed in Appendix B.

4.4.1 Sampling process

The procedure for sampling from the predictive distribution is similar to that in chapter 2 and 3. First, we generate a sample of size I=200000 from the posterior distribution for the parameters $p(a, \Sigma \mid r)$. We sample from the posterior distribution by first drawing from the marginal $p(\Sigma^{-1} \mid r)$, Wishart, and then given the $\hat{\Sigma}$ drawn, from the conditional $p(vec(M)|\hat{\Sigma},r)$, a Normal distribution. We therefore generate a sample of size 200000 from the posterior distribution for $\hat{M}$ and $\hat{\Sigma}$. Repeating this 200000 gives an accurate representation of the posterior distribution.

Secondly, for each of the 20000 realizations of the parameters $(\hat{M},\hat{\Sigma})$ in the sample from the posterior $p(a,\Sigma|r)$, we sample once from the distribution of returns conditional on both past data and the parameters $p(R_{T+\hat{T}}|\hat{M},\hat{\Sigma},r)$, a Normal distribution. This gives us a sample of size 200000 from the predictive distribution for returns, conditional only on past returns, with the parameter uncertainty integrated out.

In contrast, when parameter uncertainty is ignored we assume that the distributions for future returns are constructed using the posterior means of $\hat{a}$ and $\hat{\Sigma}$ as the fixed values of the parameters, and then drawing 200000 times from the Normal distribution with mean $\hat{T}\,\hat{a}$ and variance $\hat{T}\,\hat{\Sigma}$.

4.5 Results

In this section we illustrate the results obtained from our analysis. To see how parameter uncertainty affects portfolio choice, our strategy is to compare the allocation of an investor who takes into account estimation risk to that of an investor who ignores it. In the next paragraphs we present the optimal combinations of α and β which maximize the quantity in expression (4.11) for a variety of risk aversion levels A and investment horizons $\hat{T}$, and for different cases where the investor either ignores or accounts for parameter uncertainty,

The result are based on the model $r_t = a + \varepsilon_t$, where $r_t' = (r_{1,t}, r_{2,t})$ are the continuously compounded excess returns of the stock and bond index in month t, $a' = (a_{r_1}, a_{r_2})$ and $\varepsilon_t \sim \text{i.i.d.}N(0,\Sigma)$.

Table 4.3 gives the mean and standard deviation (in parentheses) of the posterior distribution $p(a,\Sigma|r)$ for each parameter a and Σ.

1990-2012	
a	
0.0043 (0.0027)	
0.0047 (0.0020)	
Σ	
0.0019 (0.0002)	-0.0002 (0.0001)
	0.0010 (0.0001)

Table 4.3: *Mean and standard deviation (in parenthesis) of each parameter's posterior distribution.*

For an investor using the entire sample from 1990 to 2012, the posterior distribution for the mean monthly excess stock return a_1 has mean 0.0043 and standard deviation 0.0027. The posterior distribution for the mean monthly excess bond return a_2 has instead mean 0.0047 and standard deviation0.0020. In both cases the standard deviations seem to be an important source of parameter uncertainty for the investor. The variance matrix shows the negative correlation between innovations in stock returns and bond returns, estimated here at -0.1074; this is a sign that bond can serve to diversify the portfolio and lower the risk

Ignoring parameter uncertainty

When the investor does not take into account parameters uncertainty , he solves the maximization problem (4.12), employing a distribution for future excess returns conditional on the parameter values and on the observed data of this form $p(R_{T+\hat{T}} \mid \hat{a}, \hat{\Sigma}, r)$, which is normally distributed with mean $\hat{T}\,\hat{a}$ and variance $\hat{T}\,\hat{\Sigma}$. In this case $\hat{a}$ and $\hat{\Sigma}$ are the means of each parameter's posterior distribution shown in Table 4.3

Figure 4.6 shows the optimal portfolio allocation for a buy-and-hold investor, whose preference over terminal wealth are described by a constant relative risk-aversion power utility function. The optimal combinations of α, proportion allocated to the stock index, and β, proportion allocated to the bond index, are plotted against the investment horizon that range from 1 month to 10 years. The graph on the left side is based on a relative risk-aversion level of $A=5$, the one on the right is for $A=10$.

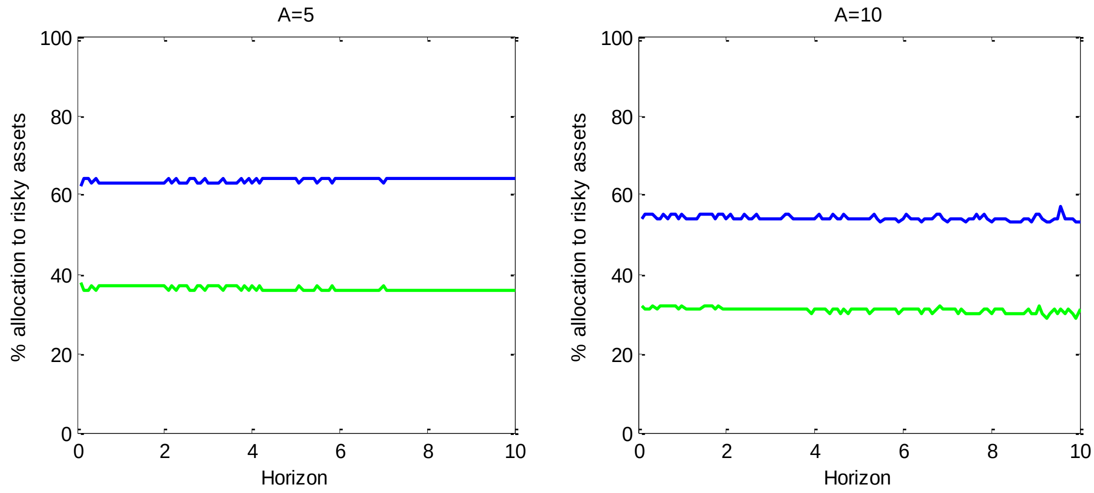

Figure 4.6: *Optimal allocation to risky assets for a buy-and-hold investor with power utility function. The percentage invested in risky assets is plotted against the investment horizon in years. The green line corresponds to the percentage invested in stocks, the blue line to the percentage invested in bonds. The graph on the left side corresponds to a level of risk-aversion of 5, the graph on the right to a level of risk-aversion of 10.*

In each graph there are two lines, a green one representing the percentage α allocated to the stock index , and a blue one representing the percentage β allocated to the bond index. Both these lines are completely horizontal in each of the two graphs. An investor ignoring the uncertainty about the mean of each parameter's posterior distribution would therefore allocate the same amount to stocks and bonds, regardless of the investment horizon. Independently from the time horizon then, the percentage allocated to the bond index is always greater than the one allocated to the stock index, whether the risk-aversion level A is equal to 5 or to 10.

This result is consistent with the context we are analyzing. Indeed, when the investor ignores parameter uncertainty, he uses a Normal distribution with mean $\hat{T}\hat{a}$ and variance $\hat{T}\hat{\Sigma}$ in order to forecast log cumulative returns; and both the mean and the variance grow linearly with the investor's horizon $\hat{T}$. A natural consequence of this is therefore that the investor chooses the same stock allocation, regardless of the holding period.

For an investor using the full data set, and when $A=5$, the optimal combination of risky assets is when α equals 34% and β equals almost 63%, whereas for an investor with $A=10$ the optimal proportion of stocks and bonds is when α equals 31.5% and β equals 54.5% . We notice that the percentage allocated to the risky assets is almost 100% when the level of risk-aversion is 5, and it falls to 86% when the level of risk aversion increases to 10. Therefore the proportion allocated to risky assets diminishes as a function of the risk-aversion level, sign that conservative investors prefer to portion their wealth between risky and risk-free assets, instead of invest all their money in risky assets. However, it is important to underline that the proportion of risky assets invested in stocks and bonds is not especially sensitive to the investor's level of risk-aversion.

Incorporating parameter uncertainty

In this section we try to show how the allocation to stocks bonds and risk-free asset differs when parameter uncertainty is explicitly incorporated into the investor's decision making framework. When he takes into account parameter uncertainty, he solves the maximization problem (4.19), throughout the application of the predictive distribution $p(R_{T+\hat{T}} \mid r)$ conditional only on past data.

Figure 4.7 shows that when $A=10$, the allocation to risky assets falls as a function of the investment horizon, on the other hand, when $A=5$, there is no considerable reduction of the allocation to risky assets as the horizon increases. Therefore we note that, in the context of i.i.d. model, the appearance of horizon effect due to parameter uncertainty, strongly depends on the investor's level of risk-aversion.

When the investor accounts for estimation risk, his distribution for long-horizon returns incorporates an extra degree of uncertainty, involving an increase in its variance. As we explained in the second chapter, this extra uncertainty makes the variance of the distribution for cumulative returns increase faster than linearly with the horizon $\hat{T}$. This makes stocks and bonds appear riskier to long-horizon investors. We therefore presume that an investor with a risk-aversion level of 5 is not affected as much as a more conservative investor by this increase in the variance. Indeed, if a conservative investor reduces the amount allocated to equities and bonds in favor of the risk-free asset, an aggressive one does not alter his allocation to stocks and bonds.

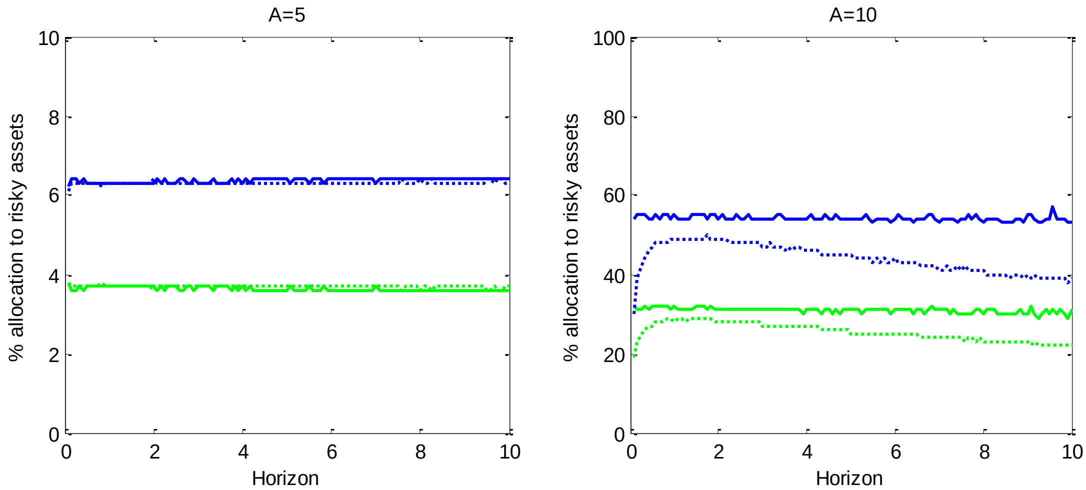

Figure 4.7: *Optimal allocation to risky assets for a buy-and-hold investor with power utility function. The percentage invested in risky assets is plotted against the investment horizon in years. The green line corresponds to the percentage invested in stocks, the blue line to the percentage invested in bonds. The solid lines refers to the cases where the investor ignores parameter uncertainty, the dotted line to the cases where he accounts for it. The graph on the left side corresponds to a level of risk-aversion of 5, the graph on the right to a level of risk-aversion of 10.*

An investor, whose level of risk aversion is equal to 5, reduces his allocation in risky assets only by 1% during a period of 10 years. On the other hand, an investor with risk-aversion level of 10 , after ten years diminishes his allocation to stocks and bonds respectively by 9% and 14% , reducing the amount allocated to risky assets by 23%.

Chapter 5

Portfolio allocation with predictable returns and five predictor variables

5.1 Introduction

This chapter focuses on how predictability affects optimal portfolio allocation, when the investor is allowed to allocate his wealth between two risky assets, the stock index and the bond index, and a risk-free asset.
For the study of the predictability of excess stock and bond returns we consider a set of five predictor variables, that are introduced and analyzed in the third paragraph.
A VAR model is then introduced to investigate how the evidence of predictability in asset returns affects optimal portfolio choice. The framework is similar to the one we implemented in the third chapter, the only difference now is that we want to study the predictive effects on stocks and bonds allocation of five predictor variables.
In the last paragraph we implement the same strategy used in chapter 3. We compare the allocation of an investor who recognizes predictability to that of an investor who is blind to it, to see whether predictability in returns has any effect on portfolio choice, and report the results obtained in our analysis.

5.2 Stock and bond predictability

In the previous chapter, we added to our model another risky asset, the bond index, therefore introducing some extensions to the framework drawn on Barberis' article. We analyzed how the portfolio choice problem changes when an individual has the opportunity to invest his wealth among three different assets instead of the usual two alternatives. We devoted the chapter to studying the portfolio decision in the context of i.i.d. returns, where no predictor variable was included in the model. However , expected returns on long term bonds can vary through time for at least two reasons:

variation in default premium, variation in term or maturity premium. In this chapter we intend to study how predictability in stock and bond returns affect optimal portfolio choice, in order to carry out this analysis we incorporate a set of five predictor variables to the previous model.

Until the 80s, in literature, most of the evidence of ex ante variables that predict excess returns was confined especially to specific types of assets. There have been steps in that direction, however. Campbell (1984) finds that, in the 1959-1979 period , several measures constructed from interest rates on U.S. Government securities predict risk premiums of Treasury bills, 20-year Government bonds, and the value-weighted portfolio of New York Stock Exchange (NYSE) common stock. In the same year Keim and Stambaugh find that several ex ante observable variables based on asset price levels predict ex post excess returns on common stocks of NYSE firms of various sizes, long-term bonds of various default-risks, and U.S Government bonds of various maturities. In 1989 Fama and French find that expected excess returns on corporate bonds and stocks move together, and that dividend yields, commonly used to forecast stock returns, also forecast bond returns. According to them, predictable variation in stock returns is, in turn tracked by variables commonly used to measure default and term premiums in bond returns.

In order to carry out our analysis we chose a set five predictor variables that are among the most used in recent financial studies. Before introducing the model we implemented, we devote the next section to a brief review of the variables we avail ourselves of.

5.3 Predictive variables

We decided to incorporate in our model a heterogeneous set of variables. The first variable is the dividend yield; we already used it in the third chapter in order to study its predictive effect in the portfolio choice, and it has a long tradition among practitioners and academics. The second one is the VIX index, which captures the stock market volatility. Then we considered the term spread and the credit spread, that mainly refer to the bond market. Finally, the risk-free rate, which is often used in financial literature to forecast returns both of equities and bonds.

5.3.1 Vix index

The Vix index, or better the Chicago Board Options Exchange Market Volatility Index, is a measure of the implied volatility of S&P 500 index options. It represents one indicator of the market's expectation of stock market risk over the next 30 day period. The monthly volatility index that we downloaded from *Yahoo! Finance* is annualized, we therefore divided it by the square root of 12 in order to convert it to a monthly measure of volatility.

Preliminary analysis

The tendency of the VIX index is displayed in figure 5.1 . Its mean is 5.895%
If we look at the autocorrelation function in figure 5.2. we can recognize many significant lags, although they appear to die out rapidly. When we implement the Dickey-Fuller test with constant, since it is significant, we obtain a value of the t-statistic equal to -4.959, which is greater, in absolute value, than the critical value -3.42, and reject the null hypothesis of presence of unit root at a significance level of 0.05, the VIX time series can therefore be considered stationary.

Vix			
Mean	0.589548	St. Error	0.001357
Minimum	0.030080	Variance	5.048e-04
Maximum	0.172888	St. Dev	0.022468
1° Quartile	0.042002	Skewness	1.584198
3° Quartile	0.069917	Kurtosis Excess	3.935083

Table 5.1: *Main descriptive statistics of the VIX series over the period 1990-2012.*

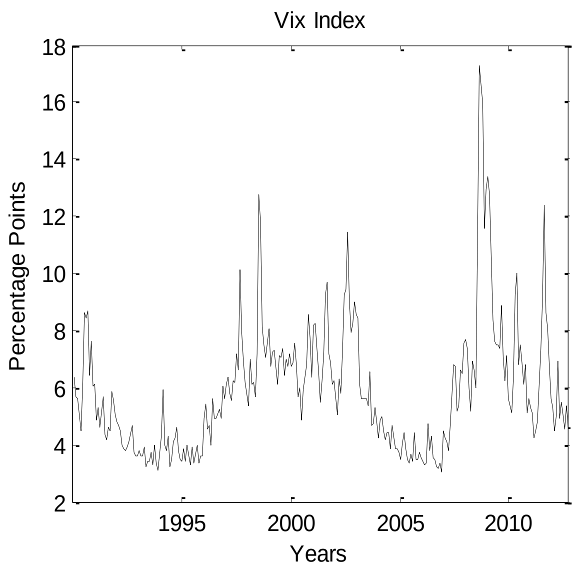

Figure 5.1: *VIX series over the period 1990-2012.*

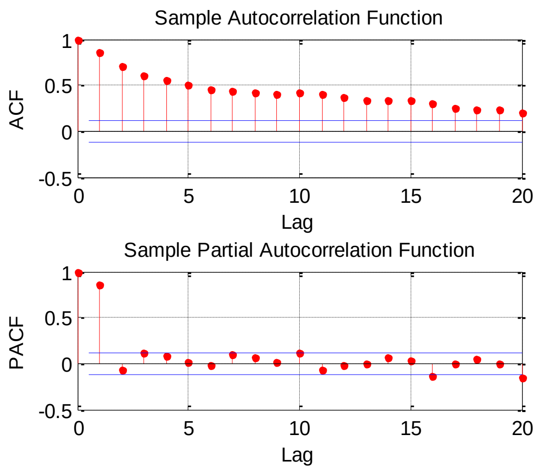

Figure 5.2: *Empirical correlogram of the VIX series.*

The VIX series has positive skewness of 1.584 and an excess of kurtosis of 3.935. We can indeed recognize that the empirical density function of the series moves away from the normal probability density function evaluated by using the sample mean and standard deviation. Moreover, looking at the Normal probability plot we can see a departure of sample quantiles from theoretical ones of the normal distribution. The normality test implemented easily rejects the null hypothesis of normality.

Normality Test	
Jarque-Bera	<0.001
Lilliefors	<0.001

Table 5.2: *Normality tests' P-values for the VIX series.*

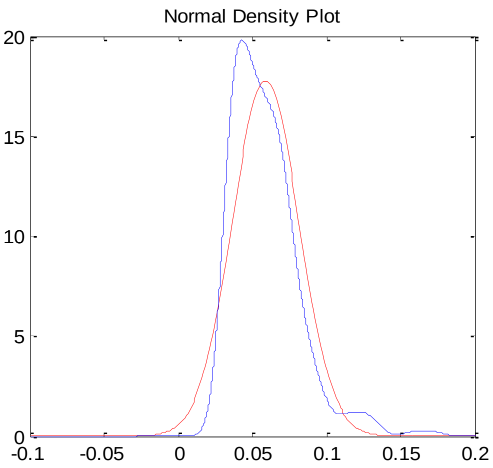

Figure 5.3: *Empirical density function of the VIX series and normal probability density function evaluated by using the sample mean and standard deviation.*

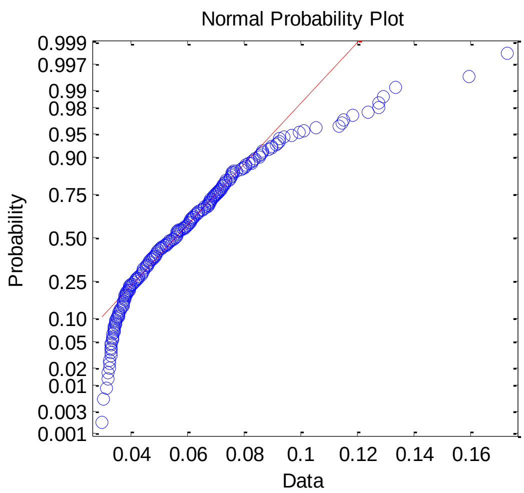

Figure 5.4: *Normal probability plot of the VIX series.*

5.3.2 Term spread

We define the term spread as the difference between the yield to maturity on long-term bonds and the yield to maturity on short term bonds. In this work we obtain the term spread as difference between the yield on the 10-year U.S. Treasury bond and the 3-month U.S. Treasury bill rate. We downloaded both of them from FRED (Federal Reserve Economic Data) . The available data are annualized, we therefore divided the annualized rates by 12 in order to get the monthly rates of return.

Preliminary analysis

The term spread series has a positive mean, this is natural since bonds with long maturities are usually characterized by a higher yield than the short maturity ones. Looking at the autocorrelation function in Figure 5.6 we see that the series has a strong persistency. When we implement the Dickey-Fuller test without constant, since it is not significant, we obtain a value of the t-statistic equal to -1.14, which is smaller, in absolute value, than the critical value -1.95, and accept the null hypothesis of presence of unit root at a significance level of 0.05. We can conclude that the term spread time series is not stationary.

Term Spread			
Mean	0.001569	St. Error	5.830e-005
Minimum	-4.396e-004	Variance	9.314e-007
Maximum	0.003119	St. Dev	9.651e-004
1° Quartile	0.000739	Skewnees	-0.161072
3° Quartile	0.002399	Kurtosis Excess	-1.142654

Table 5.3: *Main descriptive statistics of the term spread series over the period 1990-2012.*

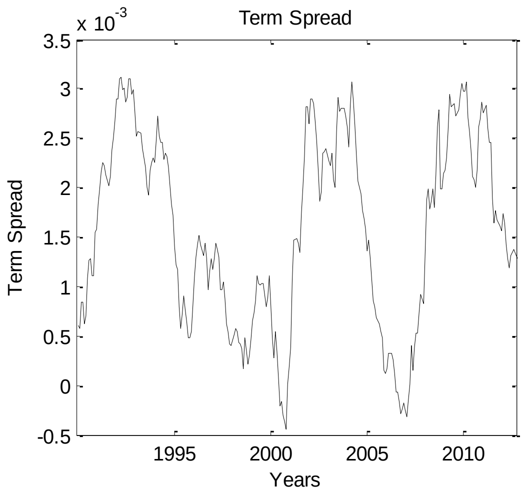

Figure 5.5: *Term spread series over the period 1990-2012.*

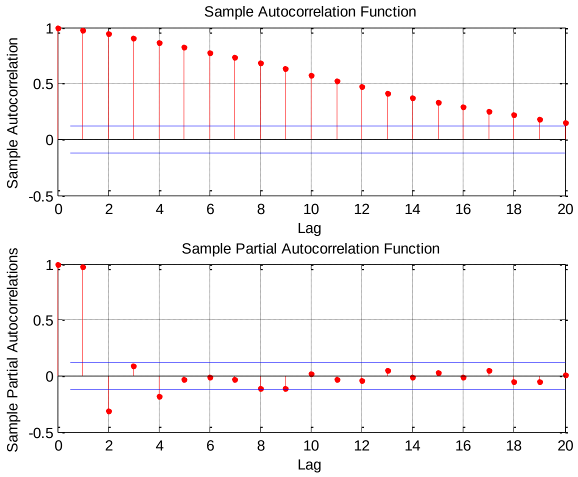

Figure 5.6: *Empirical correlogram of the term spread seires.*

If we look at figure 5.7 we can observe that the empirical density function of the series moves away from the normal probability density function evaluated by using the sample mean and standard deviation. In particular, it appears to have a lower, wider peak around the mean and thinner tails if compared to the normal density. Moreover, when we look at the Normal probability plot we notice a departure of sample quantiles from theoretical ones of the normal distribution. The normality test implemented, clearly rejects the null hypothesis of normality.

Normality Test	
Jarque-Bera	0.0040
Lilliefors	<0.001

Figure 5.4: *Normality tests' P-values for the term spread series..*

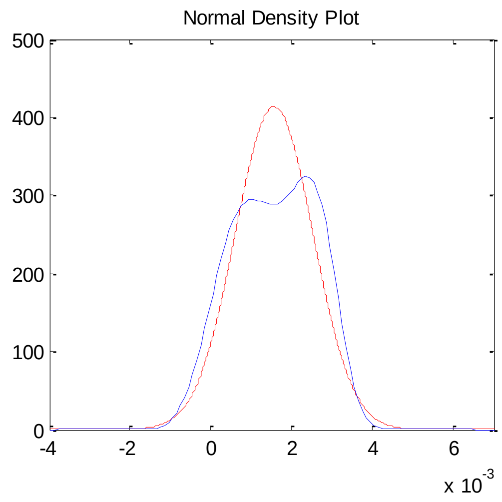

Figure 5.7: *Empirical density function of the term spread series and normal probability density function evaluated by using the sample mean and standard deviation.*

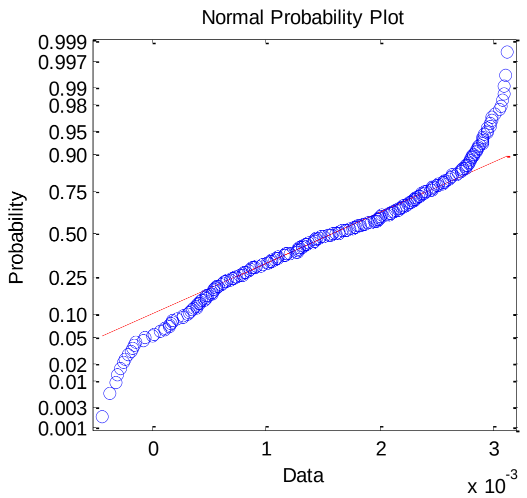

Figure 5.8: *Normal probability plot of the term spread series.*

5.3.3 Credit spread

The credit spread is the difference between the quoted rates of returns on two different investments of different credit quality. It reflects the additional net yield an investor can earn from an asset with more credit risk relative to one with less credit risk. In our analysis we refer to the credit spread as to the difference between the yield to maturity of Baa-rated corporate bonds and Aaa-rated corporate bonds (rated by Moody's Investor Service). We downloaded the data from FRED (Federal Reserve Economic Data) . Since the available data are annualized, we divided the annualized rates by 12 in order to get the monthly rates of return.

Preliminary analysis

The credit spread time series is plotted in Figure 5.9, its mean is positive and this is reasonable since the credit spread is the difference between Aaa-rated bonds yields and Baa bonds yields. The sample autocorrelation function in Figure 5.10 shows many significant lags. When we implement the Dickey-Fuller test with constant, since it is significant, we obtain a value of the t-statistic equal to –4.15, which is greater, in absolute value, than the critical value -3.42, and reject the null hypothesis of presence of unit root at a significance level of 0.05, the credit spread time series can therefore be considered as stationary.

Credit Spread			
Mean	8.032e-004	St. Error	2.133e-005
Minimum	4.553e-004	Variance	1.247e-007
Maximum	0.002801	St. Dev	3.531e-004
1° Quartile	0.000579	Skewnees	3.024166
3° Quartile	0.000920	Kurtosis Excess	11.845541

Table 5.5: *Main descriptive statistics of the credit spread series over the period 1990-2012.*

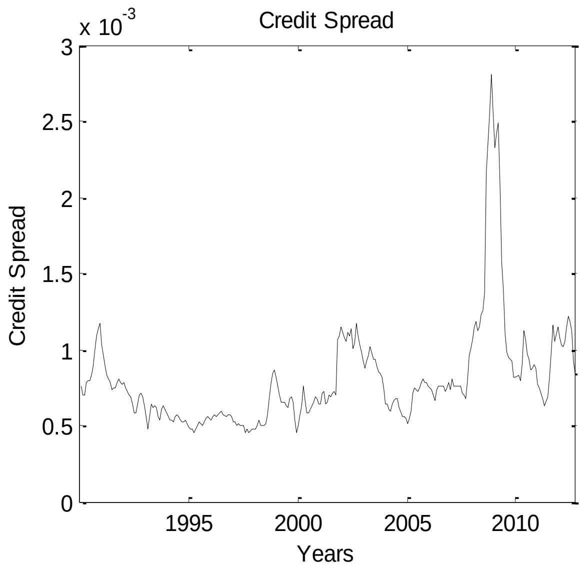

Figure 5.9: *Credit spread series over the period 1990-2012.*

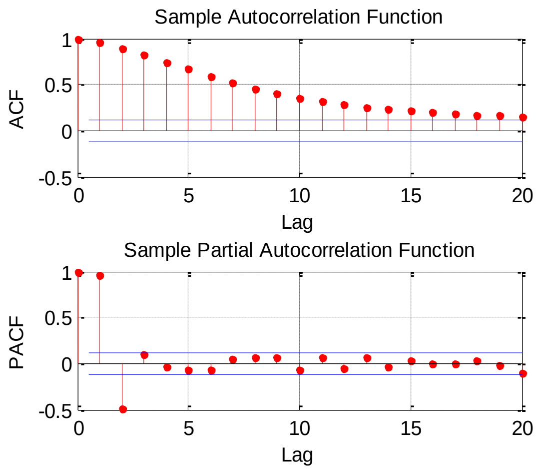

Figure 5.10: *Empirical correlogram of the credit spread series.*

The term spread series has a strong positive skewness of 3.024166 and the excess of kurtosis is equal to 11.845541. Moreover, we can recognize that the empirical density function of the series moves away from the normal probability density function evaluated by using the sample mean and standard deviation. Looking then at the Normal probability plot we can see a departure of sample quantiles from theoretical ones of the normal distribution. The normality test implemented, easily rejects the null hypothesis of normality.

Normality Test	
Jarque-Bera	<0.001
Lilliefors	<0.001

Figure 5.6: *Normality tests' P-values for the credit spread series..*

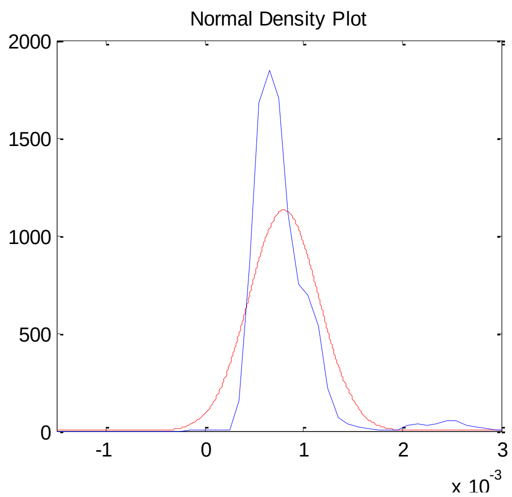

Figure 5.11: *Empirical density function of the credit spread series and normal probability density function evaluated by using the sample mean and standard deviation.*

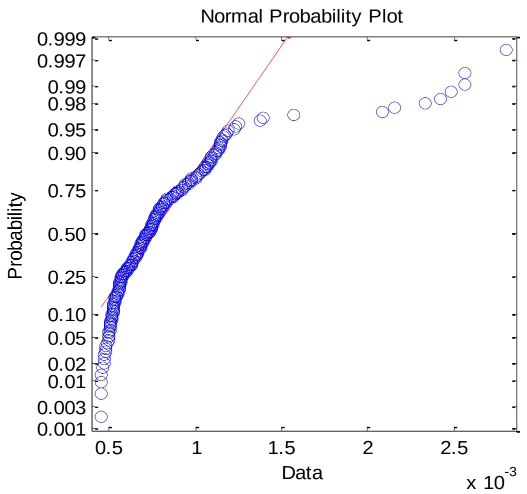

Figure 5.12: *Normal probability plot of credit spread series.*

5.3.4 Risk-free asset

Another variable we use to forecast returns of stock and bond indexes is the short-term interest rate, that we computed starting from the 3-month U.S. Treasury Bill. In the second chapter we analyzed the main properties of this variable, we observed that it was not stationary and not normally distributed.

5.3 Predictability analysis model

To investigate how the evidence of predictability in asset returns affects optimal portfolio choice we analyze a vectorial autoregressive process, VAR. The framework is similar to the one we implemented in the third chapter, the only difference now is that we want to study the predictive effects on stocks and bonds allocation of five predictor variables.

The investor uses a VAR model to forecast returns, where the state vector in the VAR include returns on stock and bond indexes and predictors variables. As we already explained before, this is an advantageous framework for examining how predictability affects portfolio choice: we can indeed compare the optimal allocation of an investor who takes return predictability into account to that of an investor who is blind to it, by only changing the number of predictor variables in the state vector. In the calculations presented in this section, the vector z_t contains seven components: the excess stock index return $r_{1,t}$, the excess bond index return $r_{2,t}$ and five predictor variables: the dividend yield $x_{1,t}$, the VIX index $x_{2,t}$, the term spread $x_{3,t}$, the credit spread $x_{4,t}$ and the risk-free rate $x_{5,t}$. The model takes this form

$$z_t = a + Bx_{t-1} + \varepsilon_t, \tag{5.1}$$

with $z_t' = (r_t, x_t')$, $x_t = (x_{1,t}, ..., x_{n,t})'$, in our analysis the number of predictor variable n is equal to 5, and $\varepsilon_t \sim \text{i.i.d.} N(0, \Sigma)$.

The first two components of z_t, namely $r_{1,t}$ and $r_{2,t}$, are the continuously compounded excess returns over month t of the stock and bond index respectively.

The other five components of z_t make up the vector or explanatory variables x_t. The first two equations in the system specify expected stock and bond returns as a function of the predictor variables. The other equations specify the stochastic evolution of the predictor variables. Referring to our model with five predictor variables, the form is this:

$$\begin{aligned}
r_{1,t+1} &= a_1 + b_{11}x_{1,t} + b_{12}x_{2,t} + b_{13}x_{3,t} + b_{14}x_{4,t} + b_{15}x_{5,t} + \varepsilon_{1,t+1}\,, \\
r_{2,t+1} &= a_2 + b_{21}x_{1,t} + b_{22}x_{2,t} + b_{23}x_{3,t} + b_{24}x_{4,t} + b_{25}x_{5,t} + \varepsilon_{2,t+1}\,, \\
x_{1,t+1} &= a_3 + b_{31}x_{1,t} + b_{32}x_{2,t} + b_{33}x_{3,t} + b_{34}x_{4,t} + b_{35}x_{5,t} + \varepsilon_{3,t+1}\,, \\
x_{2,t+1} &= a_4 + b_{41}x_{1,t} + b_{42}x_{2,t} + b_{43}x_{3,t} + b_{44}x_{4,t} + b_{45}x_{5,t} + \varepsilon_{4,t+1}\,, \\
x_{3,t+1} &= a_5 + b_{51}x_{1,t} + b_{52}x_{2,t} + b_{53}x_{3,t} + b_{54}x_{4,t} + b_{55}x_{5,t} + \varepsilon_{5,t+1}\,, \\
x_{4,t+1} &= a_6 + b_{61}x_{1,t} + b_{62}x_{2,t} + b_{63}x_{3,t} + b_{64}x_{4,t} + b_{65}x_{5,t} + \varepsilon_{6,t+1}\,, \\
x_{5,t+1} &= a_7 + b_{71}x_{1,t} + b_{72}x_{2,t} + b_{73}x_{3,t} + b_{74}x_{4,t} + b_{75}x_{5,t} + \varepsilon_{7,t+1}\,,
\end{aligned} \tag{5.2}$$

where

$$\begin{pmatrix} \varepsilon_{1,t} \\ \vdots \\ \varepsilon_{7,t} \end{pmatrix} \sim N\left(0, \begin{pmatrix} \sigma_{12} & \cdots & \sigma_{17} \\ \vdots & \ddots & \vdots \\ \sigma_{71} & \cdots & \sigma_7^2 \end{pmatrix} \right).$$

Regarding the variance-covariance matrix of contemporaneous innovations Σ we make the same assumptions we already made in chapter 3, that is, it is invertible and not necessarily diagonal; we indeed allow the shocks to be cross-sectionally correlated, but assume that they are homoscedastic and independently distributed over time

As before, the model we handle is not exactly a first order VAR, since all the variables here evaluated do not depend on the lagged value of $r_{1,t}$ and $r_{2,t}$. Basically we analyze a VAR(1) model with some restriction on its parameters , indeed we can write:

$$z_t = a + B_0 z_{t-1} + \varepsilon_t\,, \tag{5.3}$$

Where B_0 is a square matrix and its first two column contains only zeros so that z_t does not depend on $r_{1,t-1}$ and on $r_{2,t-1}$.

$$B_0 = \begin{bmatrix} \begin{pmatrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix} & B \end{bmatrix}.$$

We rewrite the model in a more convenient way:

$$\begin{pmatrix} z_2' \\ \vdots \\ z_T' \end{pmatrix} = \begin{pmatrix} 1 & x_1' \\ 1 & \vdots \\ 1 & x_{T-1}' \end{pmatrix} \begin{pmatrix} a' \\ B' \end{pmatrix} + \begin{pmatrix} \varepsilon_2' \\ \vdots \\ \varepsilon_T' \end{pmatrix}, \tag{5.4}$$

or

$$Z = XC + E, \tag{5.5}$$

where Z is a $(T-1, n+2)$ matrix with the vectors $z_2',..., z_T'$ as rows; X is a $(T-1, n+1)$ matrix with vectors $(1\ x_1'),...,(1\ x_{T-1}')$ as rows, and E is a $(T-1, n+2)$ matrix with vectors $\varepsilon_2',...,\varepsilon_T'$ as rows. Instead C is a $(n+1)(n+2)$ matrix . In this section we study the predictive effect of five predictor variables therefore n is equal to 5 and matrix C takes this form:

$$\begin{pmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 \\ B_{11} & B_{21} & B_{31} & B_{41} & B_{51} & B_{61} & B_{71} \\ B_{12} & B_{22} & B_{32} & B_{42} & B_{52} & B_{62} & B_{72} \\ B_{13} & B_{23} & B_{33} & B_{43} & B_{53} & B_{63} & B_{73} \\ B_{14} & B_{24} & B_{34} & B_{44} & B_{54} & B_{64} & B_{74} \\ B_{15} & B_{25} & B_{35} & B_{45} & B_{55} & B_{65} & B_{75} \end{pmatrix}$$

where the first row contains the intercepts and the other rows contain the coefficients of x_{t-1} .

We write down the problem faced at time T by a buy-and-hold investor with a horizon of $\hat{T}$ months. Since he has no chance to buy or sell assets between time T and horizon $T+\hat{T}$, he is interested only in the distribution of wealth at the end of the holding period. The investor problem therefore concerns the maximization of his expected utility defined over final wealth.

$$\max_{\alpha,\beta} E_T\left(\frac{\left\{W_{T+\hat{T}} = (1-\alpha-\beta)\exp(r_f\hat{T})+\alpha\exp(r_f\hat{T}+R_{1,T+\hat{T}})\ +\beta\exp(r_f\hat{T}+R_{2,T+\hat{T}})\right\}^{1-A}}{1-A}\right). \tag{5.6}$$

The investor calculates the expected utility conditional on his information set at time T, adopting different distributions of cumulative excess returns $R_{T+\hat{T}}$ .These distributions differ in whether they take into account estimation risk or not.

To avoid redundancy we do not describe again how cumulative excess returns are distributed in these two cases, we instead invite you to read section 6 of chapter 3.

5.4 Results

We devote this section to describing the results obtained by implementing the model described above. In order to carry out our analysis we chose to avail ourselves of the interactive environment of numerical computation MATLAB. The employed commands are listed in Appendix B.

The strategy we recur to, is the same one used in the third chapter to see whether predictability in returns has any effect on portfolio choice of a buy-and hold investor. In other words we compare the allocation of an investor who recognizes predictability to that of an investor who is blind to it. The VAR model provides a convenient way of making this comparison because by simply altering the number of predictor variables included in the vector x_t, it simulate investors with different information sets.

In this section we compute the optimal combinations of α and β which maximize the quantity in expression (5.6) for a variety of risk aversion levels A and investment horizons $\hat{T}$, and for different cases where the investor either ignores or accounts for parameter uncertainty.

The results are based on the model $z_t = a + Bx_{t-1} + \varepsilon_t$, where $z_t = (r_{1,t}, r_{2,t} x_t)'$ includes continuously compounded monthly excess stock returns $r_{1,t}$ and bond returns $r_{2,t}$ and a set of five predictor variables x_t, and where $\varepsilon_t \sim$ i.i.d.$N(0, \Sigma)$.

Tables 5.7 and 5.8 present the mean and standard deviation (in parentheses) of the posterior distribution $p(C, \Sigma \mid z)$ for each parameter a, B and Σ.

1990-2012					
a	B				
0.0041 (0.0152)	1.5215 (0.6077)	0.0026 (0.0016)	-30.5827 (12.7341)	-7.2628 (4.6686)	-4.2022 (2.8800)
-0.0129 (0.0112)	0.2651 (0.4448)	0.0027 (0.0012)	-17.9176 (9.3436)	4.8075 (3.4236)	1.0578 (2.1129)
0.0006 (0.0003)	0.9823 (0.0131)	-0.0000 (0.0000)	0.2090 (0.2749)	0.0365 (0.1007)	-0.0332 (0.0621)
1.1017 (0.4127)	-35.5763 (16.4642)	0.7997 (0.04346)	582.1398 (345.094	85.1288 (126.548	82.1823 (78.1225)
-0.0000 (0.0000)	0.0048 (0.0012)	0.0000 (0.0000)	0.8331 (0.0253)	-0.0258 (0.0093)	-0.0177 (0.0057)
-0.0001 (0.0001)	0.0027 (0.0028)	0.0000 (0.0000)	0.0250 (0.0585)	0.9625 (0.0214)	-0-0038 (0.0132)
0.0002 (0.0001)	-0.0050 (0.0022)	-0.0000 (0.0000)	0.0517 (0.0435)	0.0264 (0.0166)	1.0000 (0.0102)

Table 5.7: *Mean and standard deviation (in parenthesis) of parameters* (a, B) *'s posterior distribution.*

Σ						
0.0019 (0.0002)	-0.0002 (0.0001)	-3.4e-05 (0.0000)	-0.0383 (0.0041)	-1.1e-06 (0.0000)	-1.7e-07 (0.0000)	6.4e-07 (0.0000)
	0.0010 (0.0001)	1.8e-06 (0.0000)	0.0065 (0.0024)	4.3e-07 (0.0000)	-3.2e-06 (0.0000)	-5.1e-07 (3.2e-07)
		0.0000 (0.0000)	0.0006 (0.0001)	2.4e-08 (5.4e-09)	1.7e-08 (1.2e-08)	-1.4e-08 (9.3e-09)
			1.4294 (0.1261)	2.9e-05 (6.7e-06)	5.9e-06 (1.5e-05)	-2.9e-05 (0.0000)
				7.6e-09 (6.7e-10)	-3.1e-10 (1.1e-09)	1.8e.09 (8.6e-10)
					4.1e-09 (3.6e-09)	-1.4e-08 (2.1e-09)
						2.4e-08 (2.1e-09)

Table 5.8: *Mean and standard deviation (in parenthesis) of parameter* Σ *'s posterior distribution.*

In the first two rows of the B matrix is summarized the predictive power of the five predictor variables relative to the stock excess returns and to the bond excess returns. We note that the posterior distribution for those coefficients has heterogeneous means, and the standard deviations range from 0.0012 to 12.7341, which obviously appears to be an huge source of parameter uncertainty for the investor. We notice however that standard deviations are higher for those coefficients which advert to bonds predictor variables. Moreover we can see that all the predictor variables exhibit high persistency. The variance matrix shows the strong negative correlation between innovations in stock returns and the first two predictive variables, that are dividend yield and VIX index, estimated here at -0.8084 and -0.7267 respectively; this has an important influence on the distribution of long-horizon returns, even though there are many other effects to take into account since there are 3 others predictor variables. As regarding the correlation between the bond returns and the predictor variables we note that is generally weak, the only one worthy of attention is the negative correlation between the bond returns and the term spread that is -0.4944.

We want to examine how predictability in asset returns and parameter uncertainty affect portfolio choice. To do this, we compute optimal allocation using four different choices for the distribution of future returns. These distributions differ in whether they take into account predictability and estimation risk. In the fourth chapter we explored the issue of parameter uncertainty in the context of i.i.d. returns of stock and bond indexes. Here we want to see whether predictability in returns has any effect on portfolio choice throughout the implementation of a VAR model. In any case the investor may account for parameter uncertainty in the model, and thus use a predictive distribution of the form $p(R_{T+\hat{T}} \mid z)$, or he may ignore parameter uncertainty in the model; in this case we assume that the distribution for future returns are constructed using the posterior means of a , B and Σ given in Tables 5.7 and 5.8 as the fixed values of the parameters.

Ignoring parameter uncertainty

When the investor ignores parameters uncertainty , he solves the maximization problem (5.6), using a distribution for future excess returns conditional on the estimated parameter values and on the observed data of this form $p(Z_{T+\hat{T}} \mid \hat{a}, \hat{B}, \hat{\Sigma})$, which is normally distributed with mean $\hat{\mu}_{sum}$ and variance $\hat{\Sigma}_{sum}$. Since the investor's distribution for future returns depends on the values of the predictor variables at the beginning of the investment horizon x_T , we set the initial value of the predictor variables to its mean in the sample, in order not to consider the impact of the initial values in the portfolio choices, and investigate how the optimal allocation changes with the investor's horizon for these fixed initial values of predictors.

Figure 5.13 shows the optimal portfolio allocation for a buy-and-hold investor, whose preference over terminal wealth are described by a constant relative risk-aversion power utility function. The optimal combinations of α, proportion allocated to the stock index, and β, proportion allocated to the bond index, are plotted against the investment horizon that range from 1 month to 10 years. The graph on the left side is based on relative risk-aversion level of 5 , the one on the right are for $A = 10$.

In each graph the green line represents the percentage α allocated to the stock index, the blue line the percentage β allocated to the bond index.

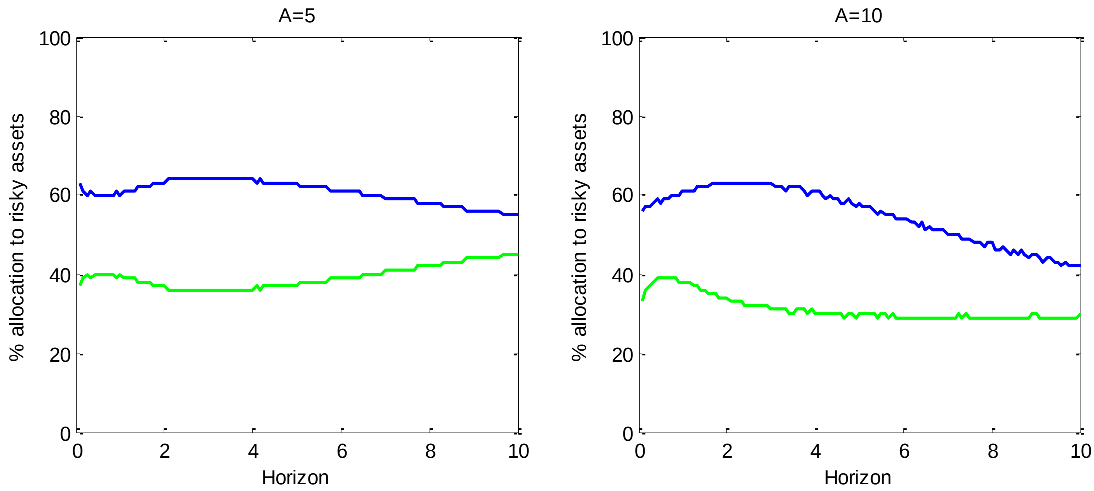

Figure 5.13: *Optimal allocation to risky assets for a buy-and-hold investor with power utility function. The percentage invested in risky assets is plotted against the investment horizon in years. The green line corresponds to the percentage invested in stocks, the blue line to the percentage invested in bonds. The graph on the left side corresponds to a level of risk-aversion of 5, the graph on the right to a level of risk-aversion of 10.*

We can note that the context changes depending on the risk-aversion level of the investor. Indeed, when the risk-aversion level is equal to 5, the green line that represents the optimal percentage α allocated to the stock index, rises as the investment horizon increases, whereas the blue line, that represents the optimal percentage β allocated to the bond index, falls as a function of the investment horizon. On the other hand, when the risk-aversion level of the investor is equal to 10, only the percentage allocated to the bond index decreases with the investment horizon, whereas the percentage invested in the stock index keeps approximately steady. It is important to note then, that when the investor's level of risk-aversion is equal to 5 the percentage allocated to risky assets keeps steady to 100%, independently from the investment horizon. When instead the risk-aversion level increases to 10, the percentage invested in risky assets diminishes as the investment horizon increases; at the beginning it nearly reaches 100% whereas in the end it is approximately around 70%.

This results differ substantially from the ones obtained before, when the investor availed himself of only one risky asset, the stock index; and when only one predictor variable was employed. In the third chapter we found that, in the context of predictability in returns the variance of cumulative returns may grow slower than linearly with the investor's horizon, lowering the perceived long-run risk of stocks and bonds and hence leading to higher allocations to risky assets in the optimal portfolio. In particular, in that case, we could explain the intuition behind this result by the effect of the negative correlation between innovations in stock returns and the dividend yield. On the other hand, now there is no longer a single predictor variable, but five ones. We need therefore to take into account the effect of all the correlations between innovations in stock and bond returns and the predictor variables, in order to explain the evolution of the variance of cumulative returns. In fact, the effects of these correlations may cancel each other out in the conditional variances of cumulative stock and bond returns, therefore not always lowering the perceived long-run risk of risky assets and hence leading to a less evident increase of the risky asset allocation in the optimal portfolio.

Ignoring parameter uncertainty

In this section we try to show how the optimal allocation differs when parameter uncertainty is explicitly incorporated into the investor's decision making framework. Our strategy for understanding the effect of parameter uncertainty is to compare the allocation of an investor who uses predictive distribution to forecast returns with the allocation of an investor who uses instead distribution of returns conditional on fixed parameters $\hat{a}$, $\hat{B}$ and $\hat{\Sigma}$.

Figure 5.14 shows the optimal portfolio allocation for a buy-and-hold investor, whose preference over terminal wealth are described by a power utility function. The optimal combinations of α, proportion allocated to the stock index, and β, proportion allocated to the bond index, are plotted against the investment horizon that range from 1 month to 10 years. The graph on the left side is based on a relative risk-aversion level of 5 , the one on the right is for $A = 10$.

The green lines in the graphs are relative to the stock allocation, whereas the blue lines refer to the bond allocation. The solid lines refer to the case where the investor ignores parameter uncertainty, the dotted lines refer to the cases where the investor accounts for estimation risk.

When we account for predictability and parameter uncertainty together, there is still horizon effect, in other words, the optimal allocation changes as the investment horizon increases. However, it is important to note that is not the kind of horizon effect we expected. In both graphs of Figure 5.14 the share invested in risky assets is strongly affected by the presence of estimation risk. For instance, when the investor's risk-aversion level is 5, the share invested in risky assets shifts from 100%, in the first five years, to 53% in the last month. Moreover, the optimal allocation to stocks and bonds is not monotonic, we can indeed observe that it first rises with the investment horizon , and then it starts falling as the investment horizon grows.

The allocation to risky assets falls even lower than the allocation of an investor who assumes that asset returns are modeled as i.i.d, whether he accounts for parameters uncertainty or not. We need to remind, that most of the means of the posterior distribution for B have large variances, which are a huge source of parameter uncertainty. Moreover we are adding the uncertainty of five different parameters together, not only the uncertainty of the dividend yield as we did in the third chapter.

This effect originate therefore from two different causes: firstly from the investor's uncertainty about the means of stock and bond returns; exactly in the same way of chapter 4, incorporating uncertainty about the means make conditional variances grow faster as the horizon increases, making stocks and bonds look more risky and inducing a lower allocation to risky assets compared to the case where estimation risk is ignored. Secondly, this effect arises from the investor's uncertainty about the predictive power of the predictor variables. It is therefore uncertain also whether the predictor variables does slow the evolution of conditional variance, and hence whether stocks and bonds' riskiness diminish with the horizon. As we explained in the third chapter, the investor acknowledge both that the predictive power may be weaker than the point estimate suggests, and that it may be stronger. These effects go on opposite directions and on net, the investor invest less at longer horizons because he is risk-averse. Moreover, other two effects go on opposite directions, accounting for predictability and incorporating estimation risk; the first one makes risky assets look less risky, the second one makes them look more risky; this therefore lead, to allocations that are not monotonic as a function of the investment horizon.

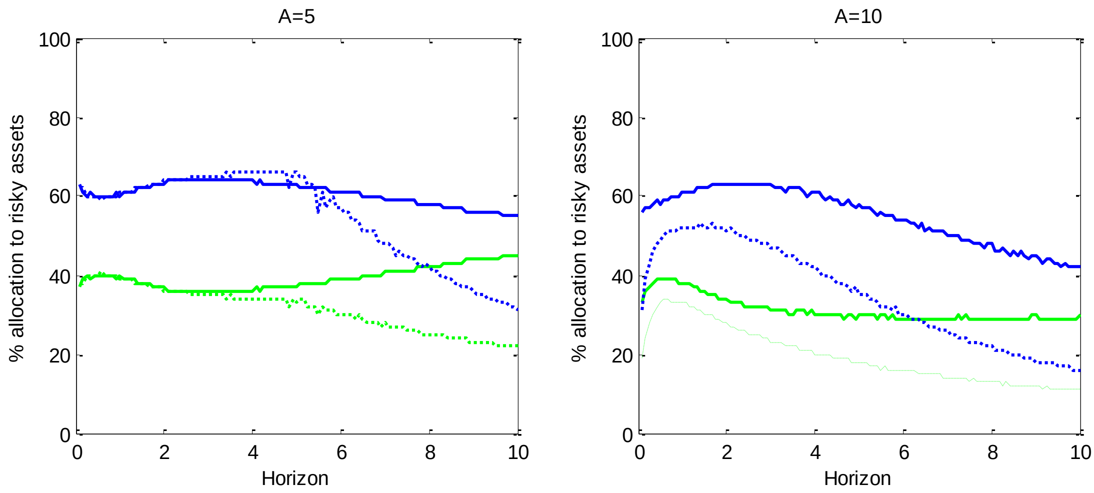

Figure 5.14: *Optimal allocation to stocks for a buy-and-hold investor with power utility function. The percentage invested in stocks is plotted against the investment horizon in years. The green line corresponds to the percentage invested in stocks, the blue line to the percentage invested in bonds. The solid line refers to the case where the investor ignores parameter uncertainty, the dot line to the case where he accounts for it. The graph on the left side corresponds to a level of risk-aversion of 5, the graph on the right to a level of risk-aversion of 10.*

5.5 The role of the predictor variables

We devote this section to analyzing the results on the optimal portfolio allocation for different initial values of the five predictor variables. As we did in the third chapter, we intend to take into account not only the impact of the predictor variables on the conditional variances but also on the mean of cumulative excess returns. This effect has not been taken into account so far because the initial values of the five predictor variables have been kept fixed at its sample mean.

In Figure 5.15, the two graphs on the left show the optimal portfolio allocations when parameter uncertainty is ignored; the graphs on the right incorporate it. Each graph exhibits the optimal stock allocation as a function of the investor's horizon for five different initial values of the predictor variables: the historical mean in our sample, the first and third quartile and the 37.5% and 67.5% percentiles.

Both graphs on the left side illustrate that, for any fixed horizon, the optimal allocation to stocks and bonds is generally, even if not always, higher for higher initial values of the predictor variables. Since the predictors affect the mean of the distribution for future returns, the investor expects higher future returns when their value is high. Besides, we can notice that, when the investor's level of risk-aversion equals 5, the optimal allocation of an investor with 10-year horizon is quite sensitive to the initial value of the predictor variables x_T. So, the various allocation do not converge to a specific value in the long run. This does not happen when the investor's level of risk-aversion is 10, indeed the amount allocated to stocks appears to converge in the long run, even though the percentage invested in bonds is still quite sensitive to the initial value of the predictors.

When we look at the two graphs on right, which refer to the case when parameter uncertainty is incorporated, we notice that the results are extremely different from the previous one. At low value of the predictors, the stock and bond allocations are generally increasing in the investment horizon, whereas those allocations are generally decreasing in the horizon at higher initial value of the predictor variables. The results obtained in the third chapter, when only the dividend yield was affecting the mean and standard deviation of cumulative excess returns, are therefore confirmed in a more elaborated context. Again, the allocation of an investor with a 10-year horizon is less sensitive to the initial value of the predictor variables than the allocation of a one-year horizon investor. In fact, the allocation lines show sign of converging.

It is reasonable to think that the degree of predictability of returns in more distant future months is less than in nearby months, the effect of the initial value of the predictors on future expected returns therefore diminishes as the investment horizon grows.

Moreover, Stambaugh (1999) finds that the various patterns in the optimal assets allocations can be understood to some degree by examining moments of the return distribution, the skewness in particular. Incorporating parameter uncertainty introduce a positive skewness in the predictive distribution for low initial value of the dividend yield, and negative skewness for high initial values. He observes that positive skewness can lead to a higher stock allocation than that obtained with negative skewness, explaining therefore the convergence to a specific value in the long run.

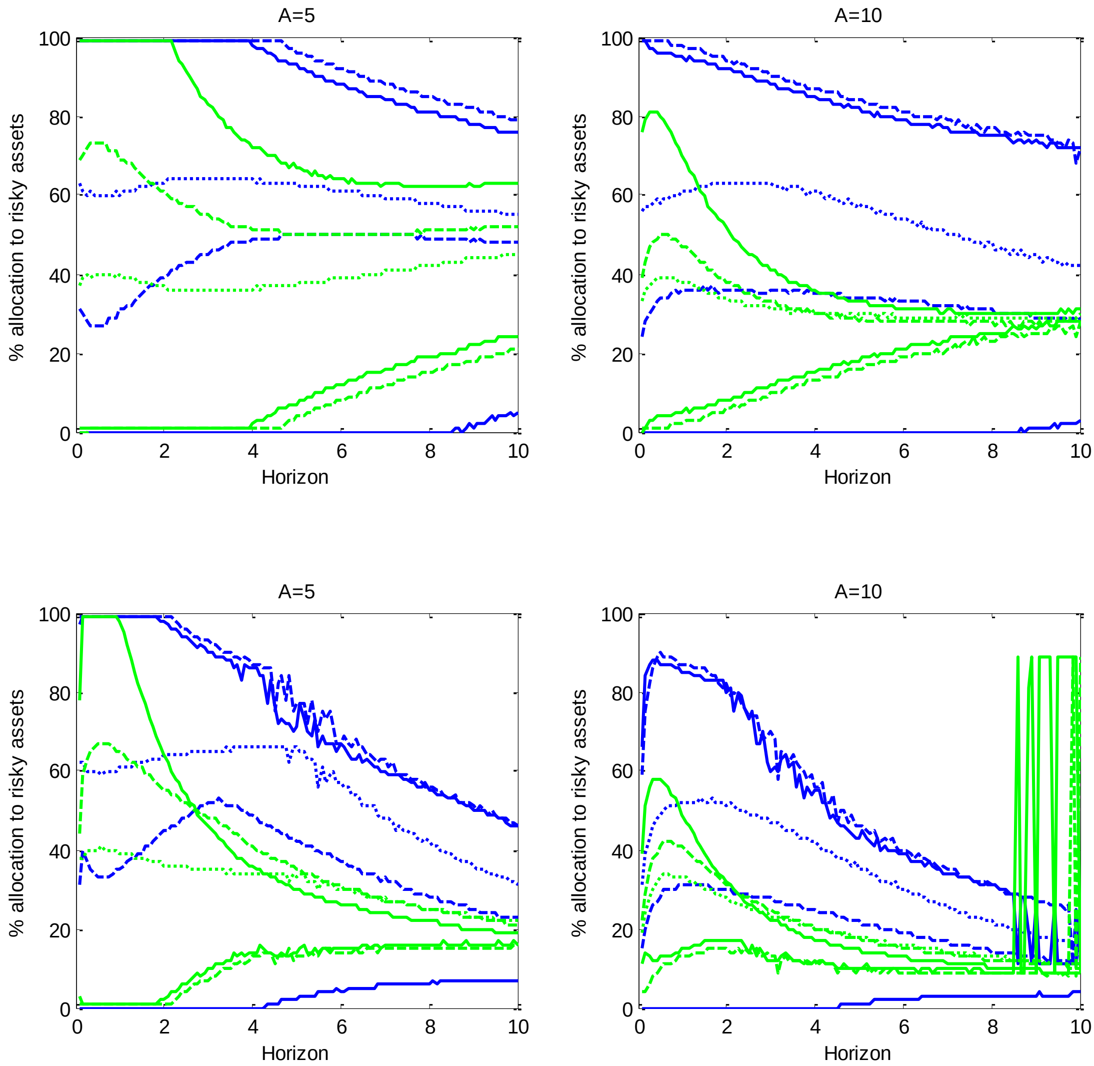

Figure 5.15: *Optimal allocation to risky assets for a buy-and-hold investor with power utility function. The percentage invested in risky assets is plotted against the investment horizon in years. The two graphs on the left ignore parameter uncertainty, the ones on the right account for it. The ten lines within each graph correspond to different initial value of the predictor variables: the mean (solid line), 37.5% and 67.5% percentiles (dashed line),first and third quartiles (dotted line).*

5.6 Other samples results

So far we illustrated the results obtained implementing the models to the sample data for the period January 1990 – November 2012. However, it is important to remind that we carry out those analysis for other three subsample. Hereinafter we briefly describe the key points of the results we obtained. In any case we invite you to take a look at Appendix A, where all the graphs are listed.

5.6.1. Sample 1990-2000

The first sample we consider is the period January 1990 – December 2000.

Figure 5.16 shows the optimal portfolio allocation for a buy-and-hold investor who recognizes predictability. The optimal combinations of α, proportion allocated to the stock index, and β, proportion allocated to the bond index, are plotted against the investment horizon that range from 1 month to 10 years. The graph on the left side is based on a relative risk-aversion level of 5 , the one on the right is for $A = 10$.

The green lines in the graphs are relative to the stock allocation, whereas the blue lines refer to the bond allocation. The solid lines refer to the case where the investor ignores parameter uncertainty, the dotted lines refer to the cases where the investor accounts for estimation risk.

The main feature that immediately strike the viewer is that the amount invested in stocks and the one invested in bond are inverted compared to the result obtained in the full sample. In fact, when the investor's level of risk-aversion is equal to 5, he allocates almost 100% to stock, whereas when his level of risk-aversion equals 10 he invests around 70% on stocks after 10 years. In this case all the correlation between innovations in stock returns and the predictor variables are negative. They can therefore affect conditional variances of cumulative stock returns, making stocks look relatively less risky at longer horizon and increasing their optimal weight in the investor's portfolio. Moreover, as we observed in the previous section, the initial values of the predictor variables is another important factor that can influence the investor's optimal allocation.

When the investor accounts for predictability and parameter uncertainty together, his behavior is similar in essence, to the one illustrated for the full period (1990-2012). The large uncertainty about the estimated parameters make the allocation to risky asset substantially fall with the horizon.

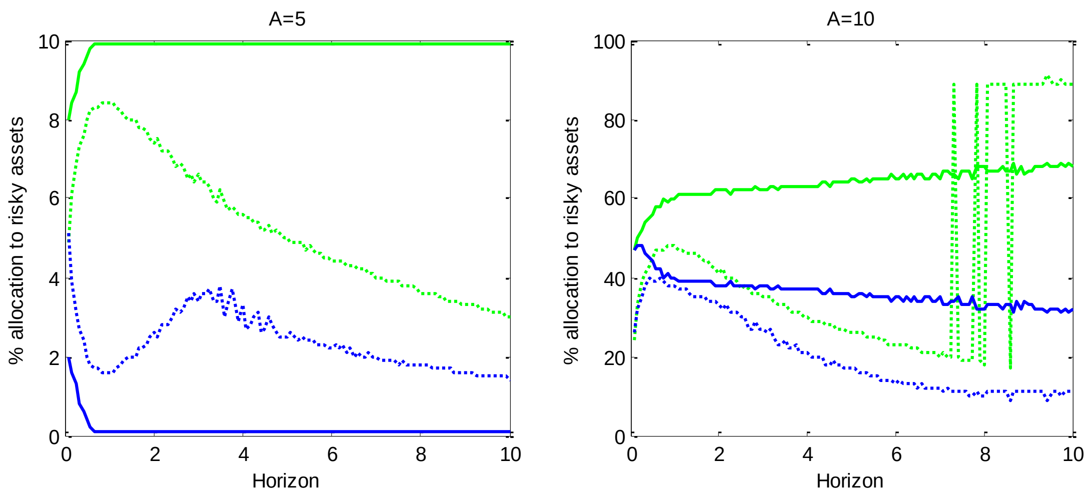

Figure 5.16: *Optimal allocation to stocks for a buy-and-hold investor with power utility function. The percentage invested in stocks is plotted against the investment horizon in years. The green line corresponds to the percentage invested in stocks, the blue line to the percentage invested in bonds. The solid lines refers to the cases where the investor ignores parameter uncertainty, the dotted line to the cases where he accounts for it. The graph on the left side corresponds to a level of risk-aversion of 5, the graph on the right to a level of risk-aversion of 10.*

5.6.2. Sample 2002-2006

The second sample we analyze is the period January 2002 – December 2006.

Figure 5.17 shows the optimal portfolio allocation for a buy-and-hold investor who recognizes predictability. The optimal combinations of α, proportion allocated to the stock index, and β, proportion allocated to the bond index, are plotted against the investment horizon that range from 1 month to 10 years. The graph on the left side is based on a relative risk-aversion level of 5 , the one on the right is for $A = 10$.

The green lines in the graphs are relative to the stock allocation, whereas the blue lines refer to the bond allocation. The solid lines refer to the case where the investor ignores parameter uncertainty, the dotted lines refer to the cases where the investor accounts for estimation risk

As in the previous case, we observe that the amount allocated to stocks is greater than the amount allocated to bonds, compared to the result obtained in the full sample. Here, the percentage allocated to the stock index nearly reaches 100% after 3 years, even when the investor's level of risk-aversion equals 10. If we observe the correlation between innovations in stock returns and the predictor variables we note that they are not all negative. Again, we would need to take into account the effect of the initial values of the predictor variable.

Looking at the lines that refer to the case when the investor accounts for predictability and estimation risk together, we note that when his level of risk-aversion equals 10, the amount invested in risky assets keeps steady around 100% after the fourth year, even if the combination of bonds and stocks appears to be variable. Instead when the investor's level of risk- aversion equals 5 the allocation to risky asset considerably fall.

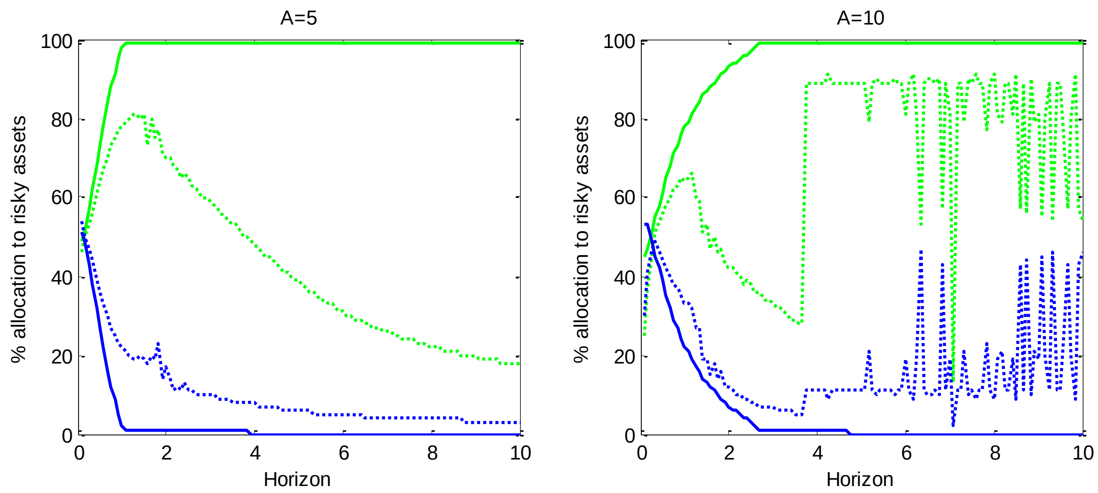

Figure 5.17: *Optimal allocation to stocks for a buy-and-hold investor with power utility function. The percentage invested in stocks is plotted against the investment horizon in years. The solid line refers to the case where the investor ignores parameter uncertainty, the dot line to the case where he accounts for it. The graph on the left side corresponds to a level of risk-aversion of 5, the graph on the right to a level of risk-aversion of 10.*

5.6.3 Sample 2007-2012

The last sample we consider is the period January 2007 – November 2012.

Figure 5.18 shows the optimal portfolio allocation for a buy-and-hold investor who recognizes predictability. The optimal combinations of α, proportion allocated to the stock index, and β, proportion allocated to the bond index, are plotted against the investment horizon that range from 1 month to 10 years. The graph on the left side is based on a relative risk-aversion level of 5 , the one on the right is for $A = 10$.

The green lines in the graphs are relative to the stock allocation, whereas the blue lines refer to the bond allocation. The solid lines refer to the case where the investor ignores parameter uncertainty, the dotted lines refer to the cases where the investor accounts for estimation risk

In this sample, as it happened for the full one, the amount invested in stocks is generally smaller than the amount invested in bonds. However, when the investor's level of risk aversion equals 5, the percentage allocated to the stock index rises more clearly with the investment horizon, compared to the full sample. In fact, around the seventh year the amount invested in stock even exceed the amount allocated to bond index. The percentage allocated to risky assets hold steady to 100% when the investor's level of risk aversion is equal to 5; when instead it increases to 10, the percentage invested in risky assets diminishes as the investment horizon increases.

When the investor accounts for predictability and parameter uncertainty together, his behavior appear to be considerably sensitive to estimation risk. The large uncertainty about the estimated parameters make the allocation to risky asset substantially fall with the horizon when the investor's level of risk aversion is equal to 5; whereas when it equals 10, the amount invested in risky asset, and in particular in stocks, reaches 100% after the third year.

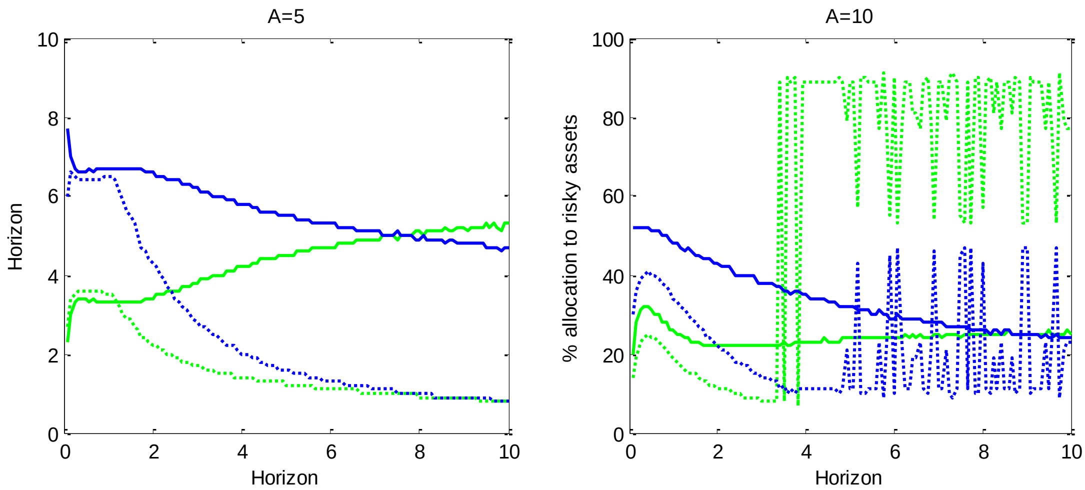

Figure 5.18: *Optimal allocation to stocks for a buy-and-hold investor with power utility function. The percentage invested in stocks is plotted against the investment horizon in years. The solid line refers to the case where the investor ignores parameter uncertainty, the dot line to the case where he accounts for it. The graph on the left side corresponds to a level of risk-aversion of 5, the graph on the right to a level of risk-aversion of 10.*

CDS in place of Dividend Yield

A further analysis we carried out, is to explore the implications of replacing the predictor variable dividend yield, which has a long tradition among practitioners and academics, by measure of credit risk, the CDS of the U.S. Banking sector. Since the data available do not cover the full sample, we decided to implement this study only for the last subsample.

In Figure 5.19 is displayed the optimal portfolio allocation for a buy-and-hold investor who recognizes predictability. The optimal combinations of α, proportion allocated to the stock index, and β, proportion allocated to the bond index, are plotted against the investment horizon that range from 1 month to 10 years. The graph on the left side is based on a relative risk-aversion level of 5 , the one on the right is for $A=10$.

The blue lines in the graphs are relative to the stock allocation, whereas the green lines refer to the bond allocation. The solid lines refer to the case where the investor ignores parameter uncertainty, the dotted lines refer to the cases where the investor accounts for estimation risk.

The similarity of these graphs with the graphs above, where the dividend yield was incorporated in the model, is obvious. Whether the investor takes into account parameter uncertainty or not, the results obtained by replacing the dividend yield with the CDS are identical, in essence, to the ones obtained before. It therefore seems that when the investor avail himself of a heterogeneous set of variable, the role of the dividend yield can easily be replaced by another variable such as the CDS.

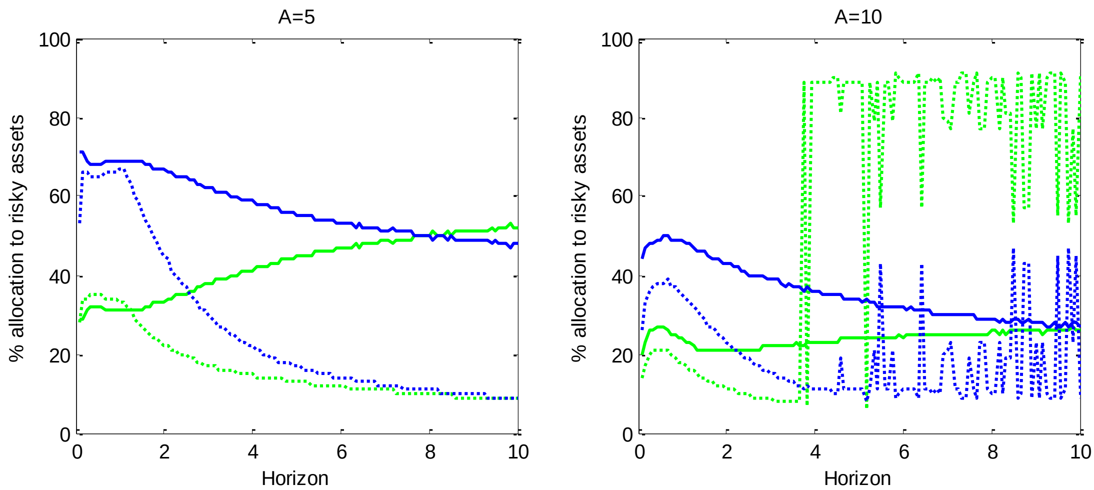

Figure 5.18: *Optimal allocation to stocks for a buy-and-hold investor with power utility function. The percentage invested in stocks is plotted against the investment horizon in years. The solid line refers to the case where the investor ignores parameter uncertainty, the dot line to the case where he accounts for it. The graph on the left side corresponds to a level of risk-aversion of 5, the graph on the right to a level of risk-aversion of 10.*

Chapter six

Portfolio allocation under loss aversion

6.1 Introduction

In this chapter we investigate how the optimal allocation changes when investors' preferences are described by a different utility function.

The first sections are devoted to showing the main critiques to the Expected Utility Theory and to illustrate the experimental contributions that led to the emergence of Behavioral Finance.

We then introduce the Prospect Theory, a behavioral economic theory that tries to model real-life choices, rather than optimal decisions, availing itself of methods originated from psychology. According to this theory a *loss aversion* function is implemented to explore the portfolio choice.

In the last paragraph the results about the optimal portfolio allocation when investors' preferences are described by a loss aversion function are given

6.2 Critiques to the Expected Utility theory

A crucial element of any model trying to understand asset prices or trading behavior is an assumption about investor preferences, or about how investors evaluate risky gambles. The majority of models assume that investors evaluate gambles according to the expected utility framework. This theory, introduced by Von Neumann and Morgenstern in 1944, has been generally accepted as a normative model of rational choice, and widely applied as a descriptive model of economic behavior. They show that if preferences satisfy a number of plausible axioms, then they can be represented by the expectation of a utility function. However there is now general agreement that this theory does not provide an adequate description of individual

choice: experimental work has shown that decision makers systematically violate Expected Utility theory when choosing among risky gambles. It may be that Expected Utility theory is a good approximation to how people evaluate a risky gamble like the stock market, even if it does not explain attitudes to the kinds of gambles observed in experimental settings. However, the difficulty the Expected Utility method has encountered in trying to explain basic facts about the stock market suggests that it may be worth taking a closer look at the experimental evidence. Indeed,
recent work in behavioral finance has argued that some of the lessons we learn from violations of Expected Utility are central to understanding a number of financial phenomena.
The first inconsistency of actual observed choices with the predictions of expected utility theory is demonstrated by Maurice Allais (1953), who, throughout his paradox, shows that people underweight outcomes that are merely probable in comparison with outcomes that are obtained with certainty. On the contrary, according to the expectation principle, the utility of a risky prospect is linear in outcome probabilities.

Daniel Kahneman and Amos Tversky (1979) find that, contrary to expected utility theory, people treat gains and losses differently and typically over- or underweight true probabilities. In their research they list some major phenomena of choice, which violate the standard model:

The certainty effect, that is the tendency to underweight outcomes that are merely probable relative to outcomes that are considered certain. However this does not mean that certainty is generally desirable, rather, it appears that certainty increases the aversiveness of losses as well as the desirability of gains. Indeed, in the positive domain, it contributes to a risk averse preference for sure gain over a larger gain that is merely probable; in the negative domain it leads to a risk seeking preference for loss that is merely probable over a smaller loss that is certain.

The reflection effect, that reverses the preference order of decision makers. That is, they usually prefer smaller gains with higher probability, whereas they prefer larger losses with lower probability. This effect causes therefore risk aversion in the positive domain and risk seeking in the negative domain . Williams reported data where a translation of outcomes produces a dramatic shift from risk aversion to risk seeking. Moreover a review by Fishburn and Kochenberger documents the prevalence of risk seeking in choices between negative prospectus.

The isolation effect, that is the disposition to disregard components that the alternatives share, and focus on the components that distinguish them, in order to simplify the choice between alternatives. This approach to choice problems may produce inconsistent preferences because a pair of prospects can be decomposed into common and distinctive components in more than one way, and different decompositions sometimes lead to different preferences (framing effect). This violates the description invariance assumed by the rational theory of choice, which asserts that equivalent formulation of a choice problem should give rise to the same preference order.

A huge amount of evidence for anomalies in human behavior has been found , the field of behavioral finance has evolved attempting to understand and explain how emotions and cognitive errors influence investors and decision-making process. The common belief in this field is that the study of psychology and social sciences can explain many stock market anomalies and shed light on the efficiency of financial markets.

6.3 Behavioral Finance

The traditional finance approach tries to understand financial markets using models in which agents are “rational”. Rationality carries two main consequences. Firstly, when decision makers receive new information, they update their beliefs correctly, in the manner described by Bayes’ law. Secondly, given their beliefs, agents make choices that are normatively acceptable, in the sense that they are consistent with Subjective Expected Utility .

This traditional framework is simple, and it would be very satisfying if its predictions were supported by the data. Unfortunately, it has become clear that basic facts about the aggregate stock market, the cross-section of average returns and individual trading behavior are not easily understood in this context.

Behavioral finance is a new approach to financial markets that has arisen, at least in part, in response to the difficulties faced by the traditional modus operandi. In general, it argues that some financial phenomena can be better understood using models in which some agents are not fully rational. More specifically, it analyzes what happens when we loosen one, or both, of the two principle that underlie individual rationality. In some behavioral finance models, agents fail to update their beliefs accurately. In other models, agents apply Bayes' law properly but make choices that are normatively controversial.

Surveys and empirical researches suggest that individuals do not always follow the traditional assumptions about rational economic decision-making. This point of view is consistent with the fundamental economic proposition that people can and do try to maximize their self-interest, but it also recognizes that such decisions are often sub-optimal, given available information. These anomalies have led to the emergence of a new approach to financial markets, Behavioral Finance. It was developed in the 50s, but only towards the end of the '70s has acquired the status of theory thanks to relevant empirical studies.

Among the various behavioral factors that usually influence agents' choice, we illustrate the most common ones, identified by behavioral finance:

Heuristic decision-making: *"Heuristics are simple rules of thumb which have been proposed to explain how people make decisions, come to judgments and solve problems, typically when facing complex problems or incomplete information. These rules work well under most circumstances, but in certain cases lead to systematic cognitive biases"* Kahneman (2011). Tversky defined heuristic as a strategy, which can be applied to a variety of problems, that usually yields a correct solution. People often use these shortcuts to reduce complex problem solving to more simple operations.

Framing effects : another bias in decision-making is a result of the fact that many participants are easily conditioned by the way in which investment question are illustrated to them. If a number of different investment options are illustrated, issues such as numbering and the order in which they appear will affect the choice made. Benartzi and Thaler (1999) find that simple changes in the way information is displayed can affect individuals‘ choices.

Overconfidence: overconfidence is the tendency for people to overestimate their knowledge, capabilities and the accuracy of their information, for that reason investment decisions become based on conjecture rather than fundamental value. A large experimental literature finds that individuals are usually overconfident (see for example Fischoff, 1982), that is, they believe their knowledge is more accurate than it actually is.

However, we do not have to think to Behavioral Finance as an alternative model to the traditional theory, but rather , considering that traditional approaches can explain the majority of phenomena, we need to think to Behavioral Finance as an opportunity to interpret, by analyzing the real investors’ behavior, those anomalies that are not fully comprehended by traditional theory.

In response of the great amount of evidence for anomalies, since 80s there has been an explosion of work on so-called non expected utility theories, all of them trying to do a better job of explaining the real behavior of decision makers. Some of the best known models include weighted-utility theory (Chew and MacCrimmon, 1979), implicit expected utility (Chew, 1989 and Dekel,1986), disappointment aversion (Gul, 1991) and probably the most relevant one, Prospect theory, originating from the work of Kahneman & Tversky (1979). We devote the next section to illustrating the main ideas that are the foundation of Prospect theory.

6.4 Prospect theory

Among the alternatives to the Expected Utility approach, Prospect theory is considered the most successful at capturing experimental results. This theory was developed by Daniel Kahneman and Amos Tversky in 1979 as a psychologically more realistic description of preferences compared to expected utility theory. Its goal is to capture people's attitudes to risky gambles as parsimoniously as possible. Indeed, Tversky and Kahneman (1986) argue that normative methods are doomed to failure, because people make choices that are simply impossible to justify on normative grounds, in that they violate dominance or invariance.

An essential feature of the present theory is that the carriers of value are changes in wealth or welfare, rather than final states. This assumption is compatible with basic principles of perception and judgment. Our perceptual apparatus is attuned to the evaluation of changes or differences rather than to the evaluation of absolute magnitudes. However, the emphasis on changes as the carriers of value should not be taken to imply that the value of a particular change is independent of initial position. Indeed, value should be treated as a function in two arguments: the asset position that serves as reference point, and the magnitude of the change (positive or negative) from that reference point.

According to Tversky and Kahneman's approach, the value function υ replaces the utility function developed by Neumann and Morgnestern, and takes this form

$$\upsilon(x) = \begin{cases} (x - x_0)^{\alpha_1} & \text{if} \quad x \geq x_0 \\ -\beta(x_0 - x)^{\alpha_2} & \text{if} \quad x < x_0 \end{cases} \tag{6.1}$$

This formulation has some important features. Firstly, utility is defined over gains and losses relative to a reference point x_0 rather than over final wealth positions, an idea first proposed by Markowitz (1952) and which has been implicitly accepted in most experimental measurements utility. Secondly, the S shape of the value function υ, namely its concavity in the domain of gains and convexity in the domain of losses. The shape of the value function depends on the parameters' values. The parameter β is the coefficient of loss aversion, a measure of the relative sensitivity to gains and losses, if set greater than 1 it allows to indicate the greater sensitivity to losses than to gains; α_1 measure the level of risk aversion for gains; α_2 measures the level of risk seeking for losses. Several values of α_1, α_2 , and β are used in financial literature, Tversky and Kahneman (1992) use experimental evidence to estimate $\alpha_1 = \alpha_2 = 0.88$, $\beta = 2.25$. Instead Gemmill, Hwang and Salmon (2005) set $\alpha_1 = 0.85$ and $\alpha_2 = 0.95$.

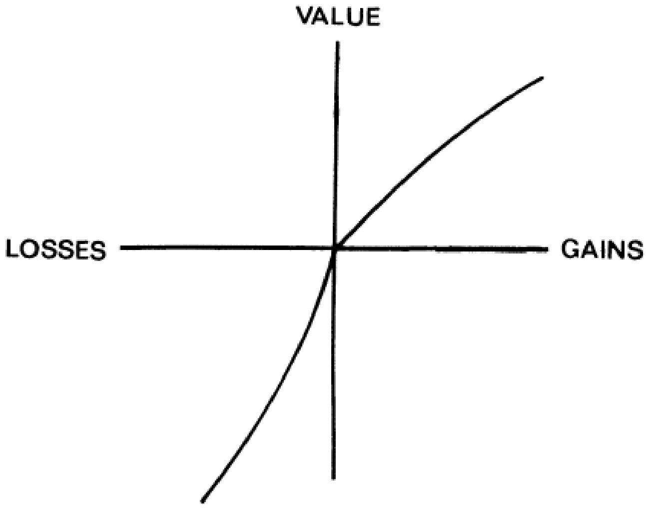

Figure 6.1: *an example of value function.*

As we already said, the υ function is concave above the reference point and convex below the reference point, a discontinuity point is therefore placed in correspondence to the reference point x_0 . These conditions reflect the principle of diminishing sensitivity : the impact of change diminishes with the distance from the reference point.

The υ function also is steeper for losses than for gains, this is implied by the principle of loss aversion according to which losses loom larger than corresponding gains. . In other words, the aggravation that one experiences in losing a sum of money appears to be greater than the pleasure associated with gaining the same amount.

In the classical theory, the utility of an uncertain prospect is the sum of the utilities of the outcomes, each weighted by its probability. On the other hand, in Prospect theory the value of each outcome is multiplied by a decision weight not by an additive probability. This weighting scheme is a monotonic transformation of outcome probabilities , however decision weights are not probabilities, they do not obey to the probability axioms and they should not be understand as a measure of degree or belief. They measure the influence of events on the desirability of prospects, and not merely the perceived likelihood of these events. The two scales coincide if the expectation principle holds, but not otherwise. Decision makers use subjective weights that overestimate or underestimate the true probability p_i. If a gamble promises outcome x_i with probability p_i people assign the gamble the value $V(G)=\sum_i \pi_i\,\upsilon(x_i)$. The weight depends on the cumulative distribution function of the gamble and are set equal to $\pi_i = w(P_i) - w(P_i^*)$, where P_i is the probability that the gamble will yield an outcome at least as good as x_i, and P_i^* is the probability that it will yield an outcome strictly better than x_i, w denote the nonlinear transform on the cumulative distribution of G

Tversky and Kahneman have suggested the following one parameter approximation in order to obtain the decision weights.

$$w(P)=\frac{P^{\gamma}}{(P^{\gamma}+(1-P)^{\gamma})^{1/\gamma}}. \tag{6.2}$$

and estimated $\gamma = 0.65$

6.5 Long horizon asset allocation

This section is dedicated to the study of the portfolio allocation problem for a buy-and-hold investor whose preferences are described by a loss aversion function developed in Tversky and Kahneman's (1992) Prospect theory.

The investor is assumed to adopt a buy-and-hold strategy, he has therefore no chance to buy or sell assets between time T and horizon $T+\hat{T}$; and he can choose to allocate his wealth between three assets: the equity index, the bond index and the risk-free asset.

In this chapter the investor's preferences are described by a loss aversion function, rather than by a power utility function. Utility is assigned to gains and losses achieved at time $T+\hat{T}$, defined relative to a reference point identified as the initial wealth W_T ,that we fix as equal to 1 for the sake of simplicity. This approach is completely different from the one developed by the expected utility theory, indeed in prospect theory values are attached to changes rather than to final wealth. On the other hand, according to the expected utility theory, utility is maximized over terminal wealth $W_{T+\hat{T}}$, independently from the level of initial wealth.

The loss aversion function takes this form:

$$\upsilon(W_{T+\hat{T}}) = \begin{cases} (W_{T+\hat{T}} - W_T)^{\alpha_1} & \text{if} \quad W_{T+\hat{T}} \geq W_T \\ -\beta(W_T - W_{T+\hat{T}})^{\alpha_2} & \text{if} \quad W_{T+\hat{T}} < W_T \end{cases} \tag{6.3}$$

The parameter β is the coefficient of loss aversion and measures the relative sensitivity to gains and losses; α_1 measures the level of risk seeking for losses; α_2 measure the level of risk aversion for gains.

We choose not to replace objective probabilities by decision weights in the portfolio allocation problem, as contemplated by Aït-Sahalia and Brandt (2001) and Berkelaar, Kouwenberg and Post (2004). Bernatzi and Thaler (1995) find that the loss aversion function is the main determinant of Prospect theory whereas the specifical functional forms of the value function and weighting functions are not critical. We therefore focus our attention on the effect that the loss aversion function has on portfolio choices.

The problem faced by the investor is the same one explained in chapter 4 and 5, the only change is that now he maximizes his utility using a loss aversion function in place of the power utility function. The investor calculates the expected utility conditional on his information set at time T adopting different distributions of cumulative excess returns $R_{T+\hat{T}}$ .These distributions differ in whether they take into account predictability and estimation risk.

Assuming that excess returns are i.i.d., so that $z_t = a + \varepsilon_t$, with $z_t' = (r_{1,t}, r_{2,t})$, $a' = (a_1, a_2)$ and $\varepsilon_t \sim \text{i.i.d.}N(0,\Sigma)$, two distributions can be used depending on whether the investor accounts or ignores parameters uncertainty .

On the other hand, if we allow for predictability in excess returns we can use a VAR model to study the predictive effects on stocks and bonds allocation of a set of five predictor variables. The model takes this form $z_t = a + Bx_{t-1} + \varepsilon_t$, with $z_t' = (r_{1,t}, r_{2,t}, x_t')$, $x_t = (x_{1,t}, ..., x_{n,t})'$ and $\varepsilon_t \sim \text{i.i.d.}N(0,\Sigma)$. The distribution of

cumulative returns conditional on the data available at time T is then normal with mean μ_{sum} and variance Σ_{sum}, where mean and variance are calculated in different way depending on whether we incorporate estimation risk or not.
To avoid redundancy, we do not illustrate again cumulative excess returns distributions , we instead invite you to read sections 4.4 and 3.6.

6.6 Results

This section presents the results of our analysis when investor's preferences are described by a loss aversion function. In order to implement the model we chose to avail ourselves of the interactive environment of numerical computation MATLAB. The employed commands are listed in Appendix B.

Our objective is to show how the portfolio allocation of a buy-and hold investor changes as a function of the investment horizon; and how the optimal combination of α and β, changes depending on whether parameter uncertainty is taken into account or ignored and if the investor recognizes predictability or is blind to it.

We use two different form of value function in order to compute the expected utility. The first one, is the loss aversion function used by Barkelaar, Kouwenbera and Post (2004) and sets $\alpha_1 = \alpha_2 = 0.88$ and $\beta = 2.25$. The second one has been used by Barberis, Huang and Santos (2001) and has $\alpha_1 = \alpha_2 = 1$ and $\beta = 2.25$. Here the investor is risk neutral for gains and losses, but he is much more distressed by losses than he is happy with equivalent gains.

Figure 6.2 shows the optimal portfolio allocation for a buy-and-hold investor whose preferences are described by a loss aversion function, and when he ignores predictability of assets returns, . The optimal combinations of α, proportion allocated to the stock index, and β, proportion allocated to the bond index, are plotted against the investment horizon that range from 1 month to 10 years. The graph on the left side refers to the case where $\alpha_1 = \alpha_2 = 0.88$, whereas the one on the right to the case where the investor is risk neutral, thus $\alpha_1 = \alpha_2 = 1$. The green lines represent the percentage allocated to the stock index, the blue ones represent instead the percentage allocated to the bond index. The dashed lines stand for when estimation risk is taken into account.

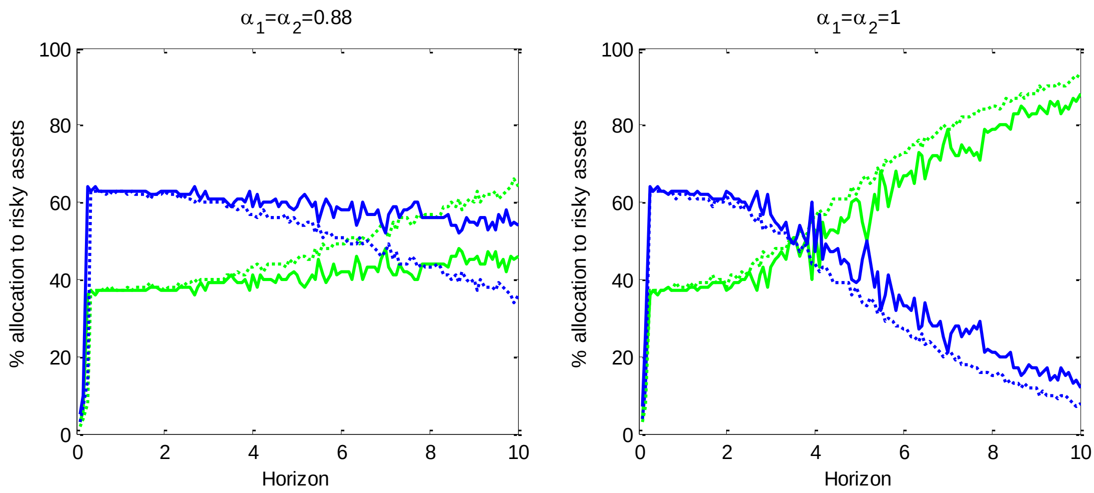

Figure 6.2: *Optimal allocation to risky assets for a buy-and-hold investor under loss aversion. The percentage invested in stocks is plotted against the investment horizon in years. The green line corresponds to the percentage invested in stocks, the blue line to the percentage invested in bonds. The solid lines refers to the cases where the investor ignores parameter uncertainty, the dotted line to the cases where he accounts for it. The graph on the left side corresponds to the case where* $\alpha_1 = \alpha_2 = 0.88$ *, the one on the right to the case where* $\alpha_1 = \alpha_2 = 1$

When we look at Figure 6.2 a distinguishing feature just leaps out of it, that is, the share allocated to the risky assets considerably changes as the investment horizon $\hat{T}$ increases. Although we are in the context of i.i.d. returns we can observe the remarkable presence of horizon effect. When the investor uses the entire sample, and parameters are $\alpha_1 = \alpha_2 = 0.88$ the percentage invested in stocks in the first month is 3% and in the third month it already reaches 36% , thereafter it keeps on growing as a function of the investment horizon. The percentage invested in bonds is instead 5% in the first month , by the end of the third month it grows to 63% and around twenty months after it starts falling up to 54%. The horizon effect is much clearer in the graph related to the risk neutral investor. Here the percentage allocated to the stock index is 36% in the third month, but it rises to 86% by the end of tenth years. Meanwhile ,the percentage allocated to the bond index fall from 63% to 12%.

The share invested in risky assets starting from the third month is always 100% irrespective of the level of risk aversion of the investor. However, we notice that when $\alpha_1 = \alpha_2 = 1$ and individuals are risk neutral for gains and losses, the percentage invested in stocks grows substantially compared to the case when $\alpha_1 = \alpha_2 = 0.88$.

All that clash with the results obtained employing a power utility function, according to which, when returns are i.i.d, the portfolio allocation holds steady irrespective of the investment horizon $\hat{T}$. Our results are also confirmed by Benartzi and Thaler (1995), they find that when we are in a loss aversion context, the attractiveness of risky asset will depend on the time horizon of the investor. The longer the investor intends to hold the asset, the more attractive the risky asset will appear.

When parameter uncertainty is taken into account, the share invested in risky assets is again 100% for the most part of the investment horizons, however the percentage invested in bonds is smaller compared to the one the investor allocates when he ignores estimation risk. Vice-versa for the share allocated to stocks.

Figure 6.3 shows the optimal portfolio allocation for a buy-and-hold investor whose preferences are described by a loss aversion function, and when he takes into account predictability of assets returns, . The optimal combinations of α, proportion allocated to the stock index, and β, proportion allocated to the bond index, are plotted as a function of the investment horizon that range from 1 month to 10 years. The graph on the left side refers to the case where $\alpha_1 = \alpha_2 = 0.88$, whereas the one on the right to the case where the investor is risk neutral, thus $\alpha_1 = \alpha_2 = 1$. The green lines represent the percentage allocated to the stock index, the blue ones represent instead the percentage allocated to the bond index. The solid lines refer to the case where the investor ignores parameter uncertainty, the dotted lines refer to the cases where the investor accounts for estimation risk.

Looking at the graphs, we note that the investor allocates all his wealth in risky assets already from the third month, both in the case where $\alpha_1 = \alpha_2 = 0.88$ and in the case where $\alpha_1 = \alpha_2 = 1$. Moreover now he allocates a larger percentage to stocks than to bonds as the horizon increases. When $\alpha_1 = \alpha_2 = 0.88$ he indeed allocates almost 100% to stocks and 0% to bonds, starting from the seventh year. When instead he is risk-neutral but loss averse, he starts to allocate 100% to stocks even before the fifth year. When estimation risk is taken into account the share invested in risky assets keeps steady at 100%, but the combination of bonds and stocks seems to be variable especially when the investment horizon is longer than 9 years.

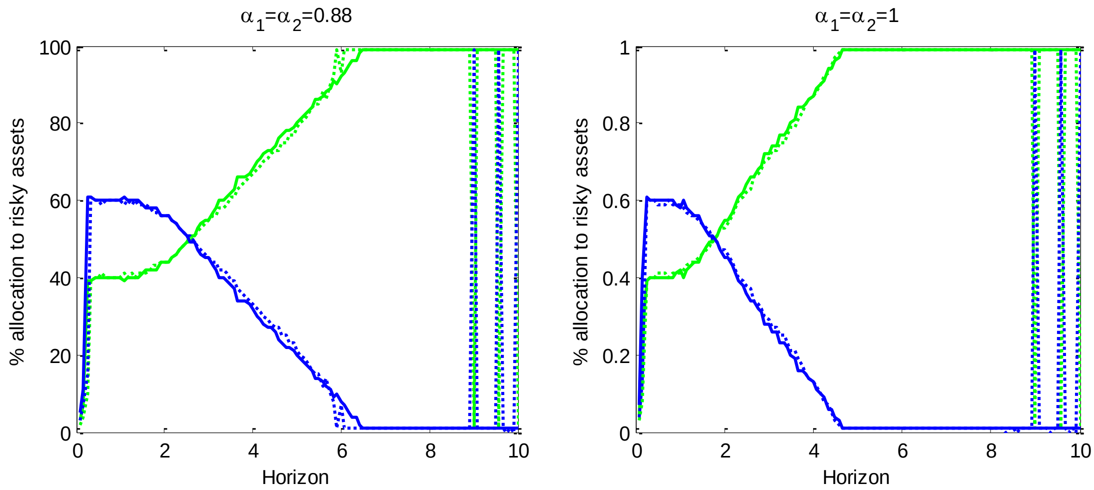

Figure 6.3: *Optimal allocation to risky assets for a buy-and-hold investor under loss aversion. The percentage invested in stocks is plotted against the investment horizon in years. The green line corresponds to the percentage invested in stocks, the blue line to the percentage invested in bonds. The solid lines refers to the cases where the investor ignores parameter uncertainty, the dotted line to the cases where he accounts for it. The graph on the left side corresponds to the case where* $\alpha_1 = \alpha_2 = 0.88$ *, the one on the right to the case where* $\alpha_1 = \alpha_2 = 1$

Conclusions

In this work we investigated the issue of portfolio choices for investors with long horizons. In particular, given the evidence of predictability in asset returns of recent financial research, we examined the effects of this predictability for investors seeking to make portfolio allocation decisions. Our study reveals that portfolio allocations for short- and long-horizon investors can be very different in the context of predictable returns.

For the most part of our work we assumed that investor's preferences over terminal wealth were described by a constant relative risk-aversion power utility function.
We started out our analysis by considering the case where the investor was allowed to choose how to allocate his wealth only between two assets: the stock index and the risk-free asset. When asset returns are assumed to be i.i.d. with normal innovations, and the parameters in the model are treated as if known with complete precision, we observe that the optimal allocation is independent of the horizon, remaining identical to the short run.

On the other hand, we observe that when parameter uncertainty is explicitly incorporated into the investor's decision making framework, by using a Bayesian approach, the stock allocation falls as the horizon increases, parameter uncertainty can therefore introduce horizon effect even in the context of i.i.d. model returns. This extra uncertainty means the variance of the distribution for cumulative returns increase faster than linearly with the horizon, making stocks appear riskier to long-horizon investors

Afterwards we considered the impact of predictability implementing a VAR model, an important aspect of this analysis is that in constructing optimal portfolios, we accounted for the fact that the true extent of predictability in returns is highly uncertain.
When we ignore the estimation risk we observe that the optimal allocation to equities for a long-horizon investor is much higher than for a short-horizon investor. In the context of predictability in returns the variance of cumulative stock returns may grow more slowly than linearly with the investor's horizon, which is the case when asset returns are modeled as i.i.d., lowering the perceived long-run risk of stocks and hence leading to higher allocations to stocks in the optimal portfolio.
However when we accounted for predictability and parameter uncertainty together, we still find horizon effect, although the long-horizon allocation is not nearly as high as when we ignore estimation risk. We can deduce that incorporating parameter uncertainty can considerably reduce the size of the horizon effect. Therefore a long-horizon investor who ignores parameter uncertainty may over-allocate to stocks by a sizeable amount.

We then devoted the majority of our work to examining in what way the optimal portfolio allocation changes when investors have the opportunity to choose how to allocate their wealth among three different assets, instead of the previous two: a stock index, a bond index, and a risk-free asset.
We firstly assumed i.i.d. modeled returns and we observed that an investor ignoring the uncertainty about the mean and variance of assets returns would allocate the same amount to stocks and to bonds regardless of his investment horizon. Independently from the time horizon and from the risk-aversion level then, the percentage allocated to the bond index is always greater than the one allocated to the stock index.
Accounting for estimation risk instead, the investor's distribution for long-horizon returns incorporates an extra degree of uncertainty, involving an increase in its variance.

We then investigated the predictability of excess stock and bond returns availing ourselves of a set of five predictor variables commonly used in literature for the portfolio choice problems. In this case the allocation to stocks, and in general to risky assets, does not rise so dramatically at long horizons as in the case where only two assets and one predictor variable were included in the model. In particular, the amount allocated to the bond index tends to fall as the horizon increases, whereas the percentage invested in the stock index rises slightly or keeps approximately steady, depending on the risk aversion level. When we consider the predictive power of five variables instead of the sole dividend yield, the change of the conditional variance over time appears not to be as obvious as before, since now the effect of the five variables influences its form.
When we account for predictability and parameter uncertainty together, we note that the horizon effect is still present, however the share invested in risky assets is strongly affected by the presence of estimation risk. The allocation to risky assets falls even lower than the allocation of an investor who assumes that asset returns are modeled as i.i.d, whether he accounts for parameters uncertainty or not. This effect arises therefore from two different causes: firstly from the investor's uncertainty about the means of stock and bond returns. Secondly, from the investor's uncertainty about the predictive power of the predictor variables.

When we employ a loss aversion function, instead of the common power utility function, in order to describe the investor's preferences, the optimal portfolio allocation changes dramatically. Even in the context of i.i.d. returns we can observe remarkable presence of horizon effects. All that clashes with the results obtained employing a power utility function, according to which, when returns were i.i.d, the portfolio allocation held steady irrespective of the investment horizon. When we take parameter uncertainty into account, we find that the percentages invested in bonds and stocks changes slightly compared to the case where the investor does not account for estimation risk .

When we consider the effects of predictability, we note that the percentage allocated to stocks is increasingly large, until it reaches 100% at a investment horizon of seven and even five years, depending on the risk aversion level. When estimation risk is taken into account the share invested in risky assets is still the same, but the combination of bonds and stocks seems to be variable.

Our results suggest that portfolio calculations can be seriously deceptive if the allocation framework ignores the uncertainty surrounding parameters evaluation. Moreover, we observe that parameter uncertainty makes the optimal allocation much less sensitive to the initial value of the predictor variables. This suggest that studies which ignore uncertainty about parameters may lead the investor to take positions in stocks which may be both too large and too sensitive to the predictor variables selected.

This work makes it possible to extend the model and their framework to examine other issues of interest to investors. We could indeed change the assets included in the model or select a different set of predictor variable. We could introduce variation in conditional volatilities and conditional means. Finally we could consider time-variation in the studied parameters.

An intriguing extension of what we have handled in this work concerns the study of the dynamic problem faced by an investor who rebalances optimally at regular intervals. This investment strategy, better approximation of reality, refers to the hedging demands, originally treated by Merton (1973).

Appendix A

In addition to the sample examined throughout the work (January 1990 – November 2012), we carried out all the analysis for other three subsample: January 1990 – December 1999, January 2002 – December 2006, January 2007 – November 2012. Hereinafter we list the most meaningful graphs of each samples.

Sample 1990 – 2000

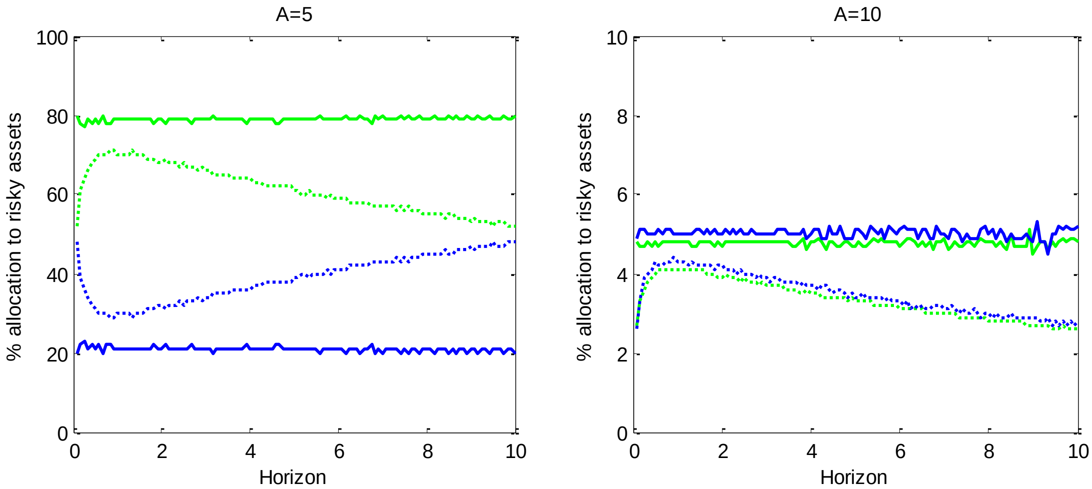

Figure A.1: *Optimal allocation to risky assets for a buy-and-hold investor with power utility function when he does not take into account predictability. The percentage invested in risky assets is plotted against the investment horizon in years. The green line corresponds to the percentage invested in stocks, the blue line to the percentage invested in bonds. The graph on the left side corresponds to a level of risk-aversion of 5, the graph on the right to a level of risk-aversion of 10.*

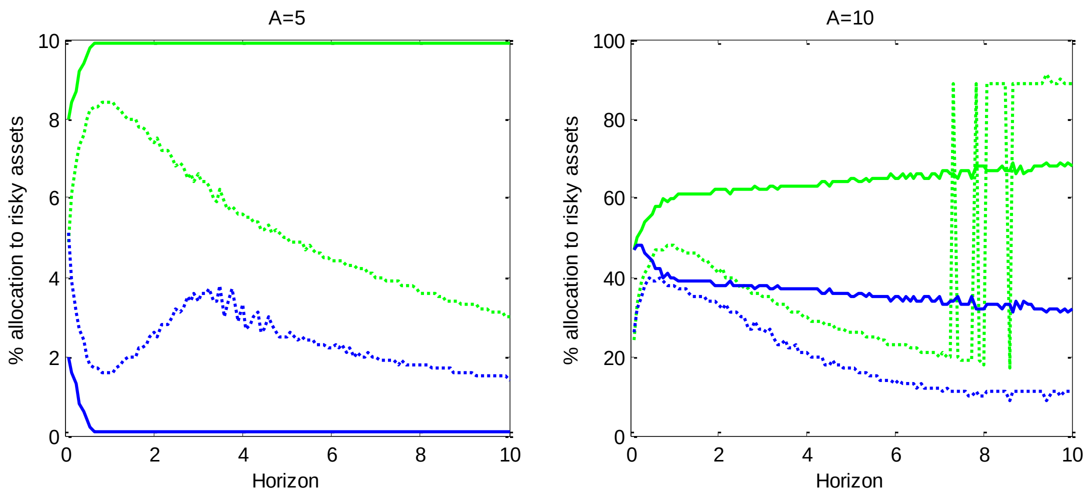

Figure A.1: *Optimal allocation to risky assets for a buy-and-hold investor with power utility function when he does take into account predictability. The percentage invested in risky assets is plotted against the investment horizon in years. The green line corresponds to the percentage invested in stocks, the blue line to the percentage invested in bonds. The graph on the left side corresponds to a level of risk-aversion of 5, the graph on the right to a level of risk-aversion of 10.*

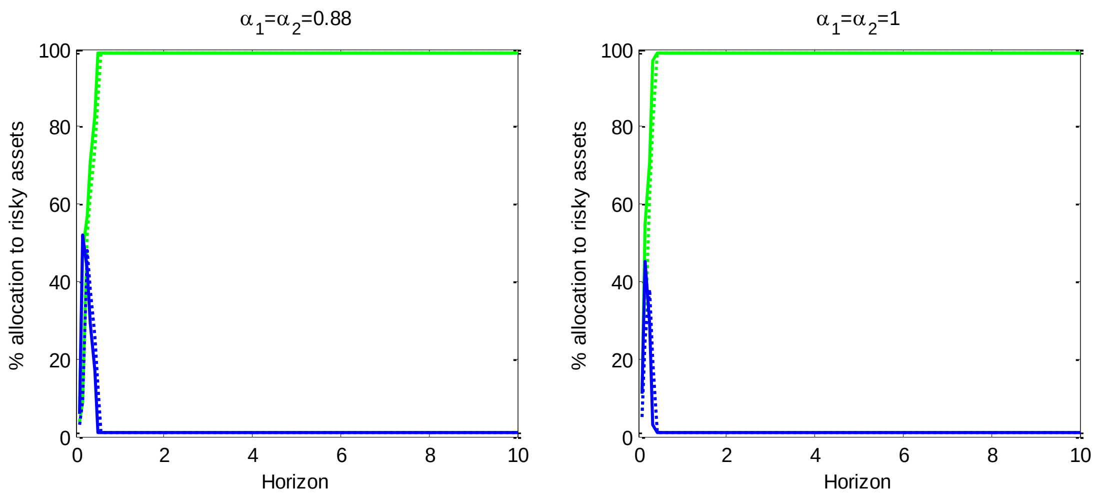

Figure A.1: *Optimal allocation to risky assets for a buy-and-hold investor with loss aversion function when he does not take into account predictability. The percentage invested in risky assets is plotted against the investment horizon in years. The green line corresponds to the percentage invested in stocks, the blue line to the percentage invested in bonds. The graph on the left side corresponds to the case where* $\alpha_1 = \alpha_2 = 0.88$ *, the one on the right to the case where* $\alpha_1 = \alpha_2 = 1$

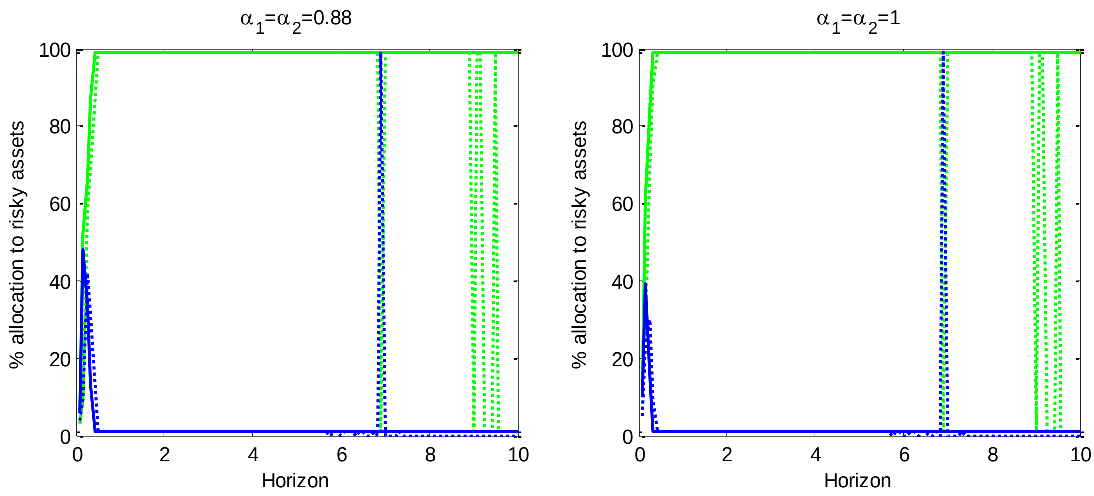

Figure A.1: *Optimal allocation to risky assets for a buy-and-hold investor with loss aversion function when he does take into account predictability. The percentage invested in risky assets is plotted against the investment horizon in years. The green line corresponds to the percentage invested in stocks, the blue line to the percentage invested in bonds. The graph on the left side corresponds to the case where* $\alpha_1 = \alpha_2 = 0.88$ *, the one on the right to the case where* $\alpha_1 = \alpha_2 = 1$

Sample 2002 – 2006

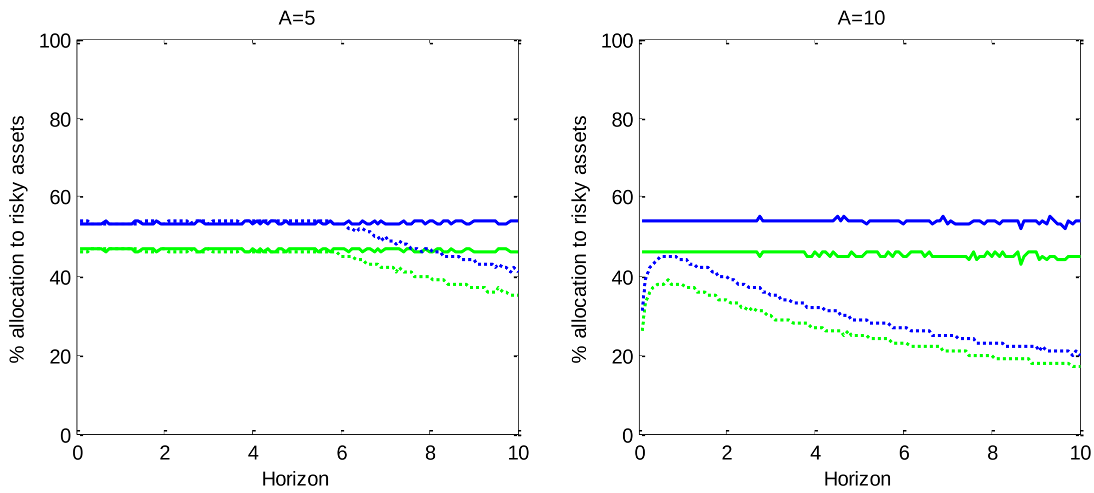

Figure A.1: *Optimal allocation to risky assets for a buy-and-hold investor with power utility function when he does not take into account predictability. The percentage invested in risky assets is plotted against the investment horizon in years. The green line corresponds to the percentage invested in stocks, the blue line to the percentage invested in bonds. The graph on the left side corresponds to a level of risk-aversion of 5, the graph on the right to a level of risk-aversion of 10.*

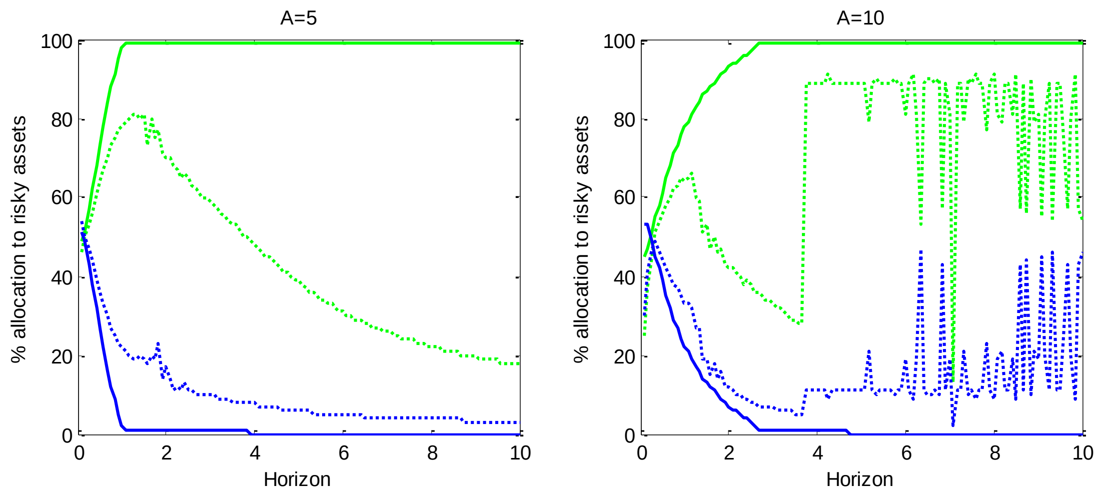

Figure A.1: *Optimal allocation to risky assets for a buy-and-hold investor with power utility function when he does take into account predictability. The percentage invested in risky assets is plotted against the investment horizon in years. The green line corresponds to the percentage invested in stocks, the blue line to the percentage invested in bonds. The graph on the left side corresponds to a level of risk-aversion of 5, the graph on the right to a level of risk-aversion of 10.*

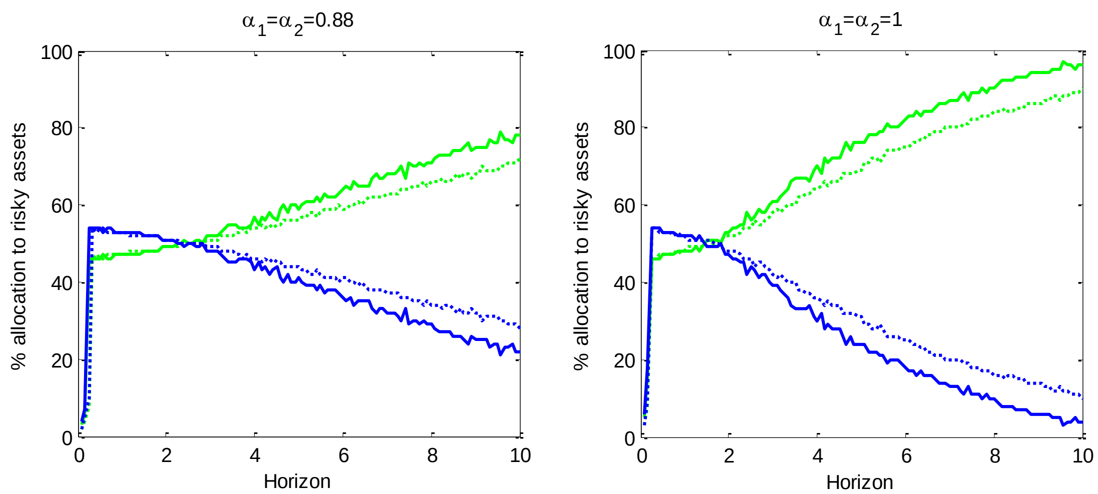

Figure A.1: *Optimal allocation to risky assets for a buy-and-hold investor with loss aversion function when he does not take into account predictability. The percentage invested in risky assets is plotted against the investment horizon in years. The green line corresponds to the percentage invested in stocks, the blue line to the percentage invested in bonds. The graph on the left side corresponds to the case where* $\alpha_1 = \alpha_2 = 0.88$ *, the one on the right to the case where* $\alpha_1 = \alpha_2 = 1$

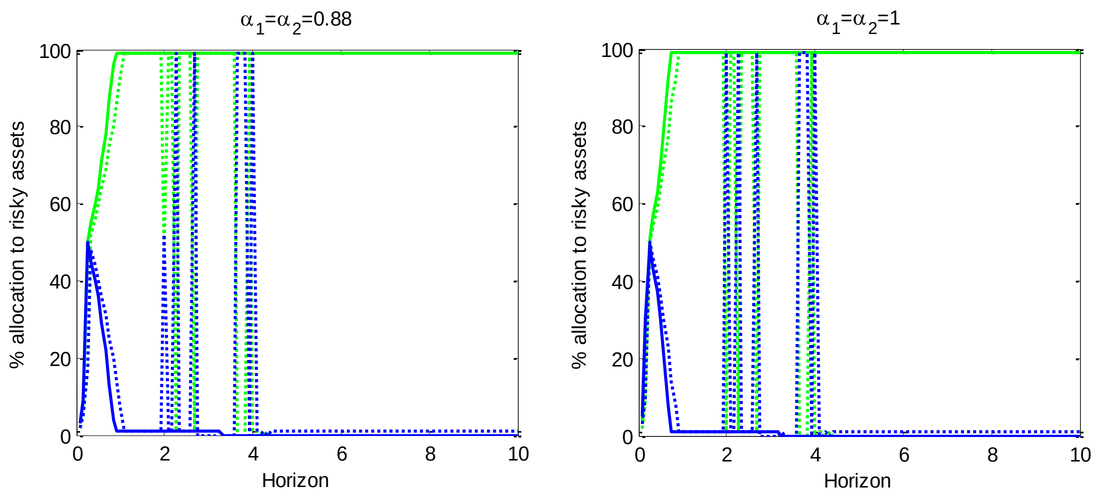

Figure A.1: *Optimal allocation to risky assets for a buy-and-hold investor with loss aversion function when he does take into account predictability. The percentage invested in risky assets is plotted against the investment horizon in years. The green line corresponds to the percentage invested in stocks, the blue line to the percentage invested in bonds. The graph on the left side corresponds to the case where* $\alpha_1 = \alpha_2 = 0.88$ *, the one on the right to the case where* $\alpha_1 = \alpha_2 = 1$

Sample 2006 - 2012

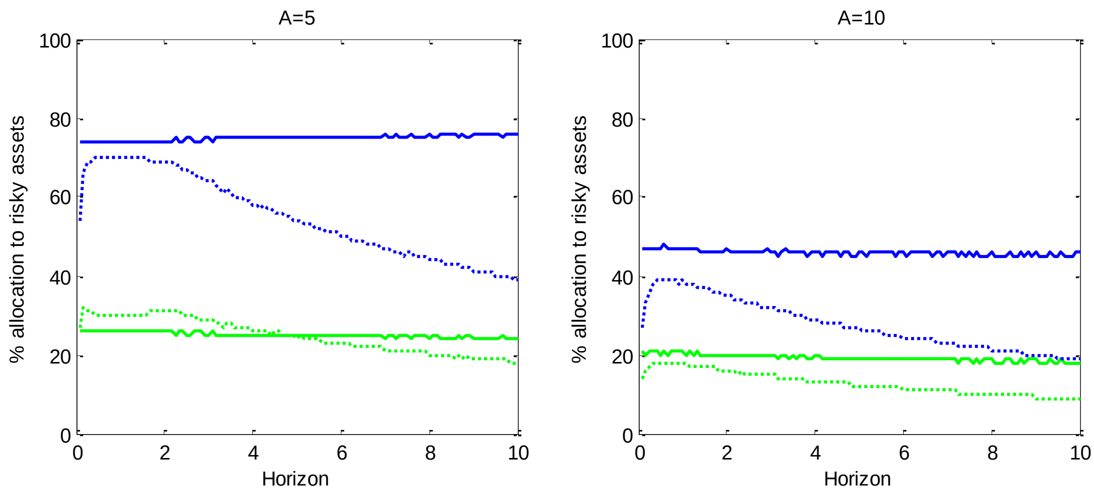

Figure A.1: *Optimal allocation to risky assets for a buy-and-hold investor with power utility function when he does not take into account predictability. The percentage invested in risky assets is plotted against the investment horizon in years. The green line corresponds to the percentage invested in stocks, the blue line to the percentage invested in bonds. The graph on the left side corresponds to a level of risk-aversion of 5, the graph on the right to a level of risk-aversion of 10.*

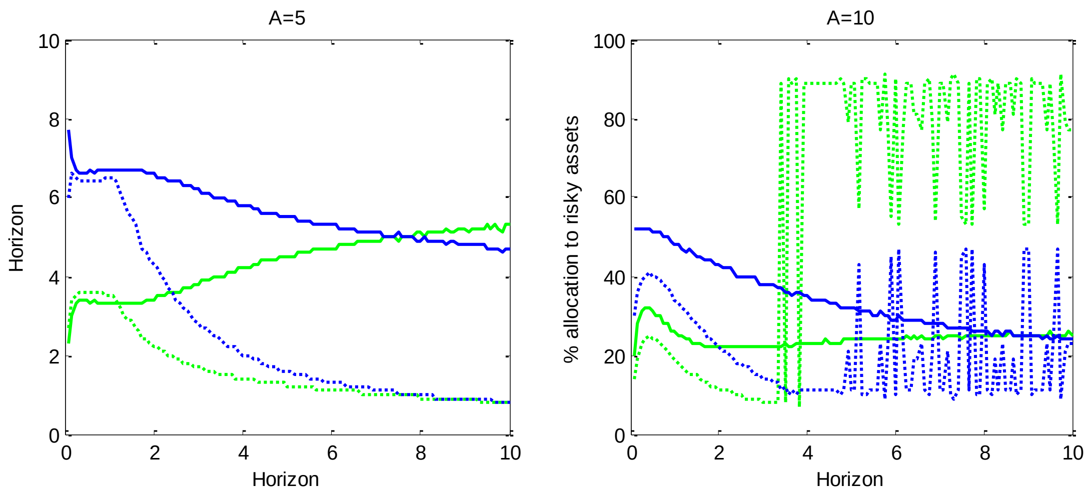

Figure A.1: *Optimal allocation to risky assets for a buy-and-hold investor with power utility function when he does take into account predictability. The percentage invested in risky assets is plotted against the investment horizon in years. The green line corresponds to the percentage invested in stocks, the blue line to the percentage invested in bonds. The graph on the left side corresponds to a level of risk-aversion of 5, the graph on the right to a level of risk-aversion of 10.*

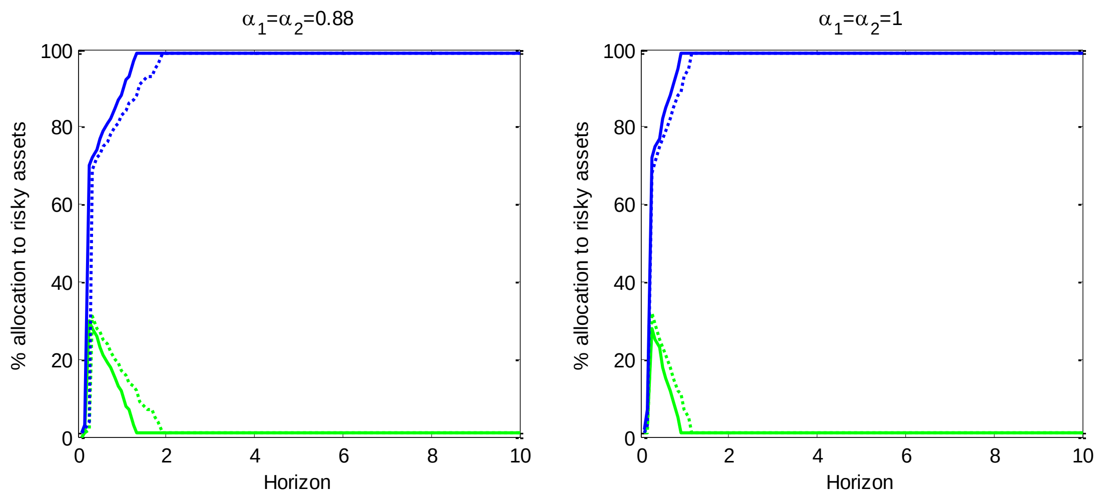

Figure A.1: *Optimal allocation to risky assets for a buy-and-hold investor with loss aversion function when he does not take into account predictability. The percentage invested in risky assets is plotted against the investment horizon in years. The green line corresponds to the percentage invested in stocks, the blue line to the percentage invested in bonds. The graph on the left side corresponds to the case where* $\alpha_1 = \alpha_2 = 0.88$ *, the one on the right to the case where* $\alpha_1 = \alpha_2 = 1$

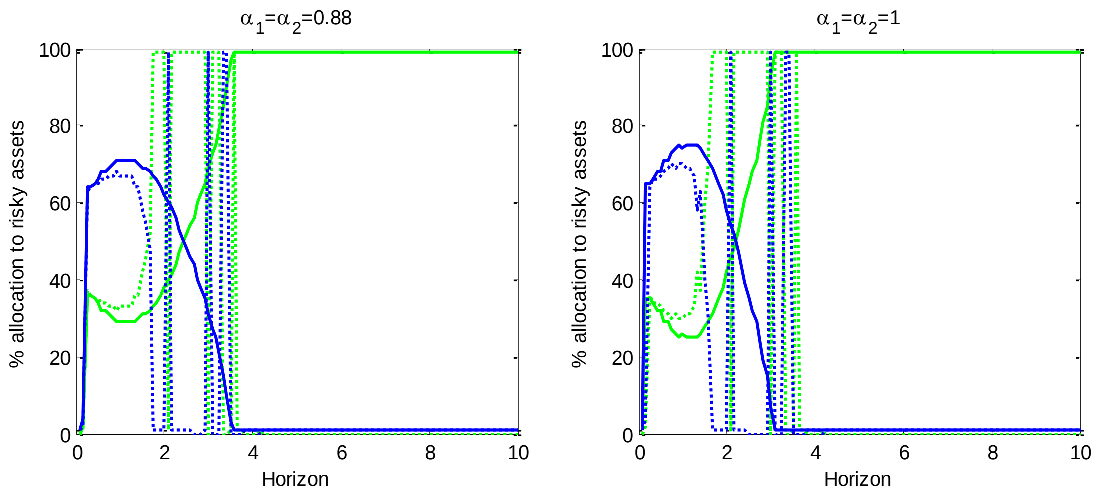

Figure A.1: *Optimal allocation to risky assets for a buy-and-hold investor with loss aversion function when he does take into account predictability. The percentage invested in risky assets is plotted against the investment horizon in years. The green line corresponds to the percentage invested in stocks, the blue line to the percentage invested in bonds. The graph on the left side corresponds to the case where* $\alpha_1 = \alpha_2 = 0.88$ *, the one on the right to the case where* $\alpha_1 = \alpha_2 = 1$

Appendix B

Portfolio allocation with parameter uncertainty

Here, the commands we used to implement the analysis of chapter 2 are listed.

We assumed i.i.d. excess returns of the form $r_t = \mu + \varepsilon_t$, where $\varepsilon_t \sim \text{i.i.d.}N(0,\sigma^2)$.

```
% uploading data
load dati.mat
num_sample=200000
TB3MS;                        % 3-months treasury bill
ri;                           % S&P 500 Stock Price Index
div = dy;                     % S&P 500 DY m=length(ri);

% Continuously compounded stock returns (included dividend payments)
wdr = zeros((m-1),1);
for i = 2:m;
      wdr(i) = log(ri(i)/ri(i-1));
```

```
end

rf = log(1+TB3MS/1200);
rfree = mean(rf);
rt = wdr-rfree;              % stock excess returns
mu=mean(rt);
sigma2=var(rt);
n=length(rt);
omega = [0:0.01:0.99]' ; % stock percentage

% Power Utility Function
function [util] = U(x,R)
      util = (x.^(1-R))/(1-R);
end



% PARAMETER UNCERTAINTY

a = (n-1)/2;
b = ((n-1)*sigma2)/2;
c = 1/b;

x = gamrnd(a,c,[num_sample,1]);
marg = 1./x  ;           % sigma2|r distribution, Inverse-Gamma

W = zeros(100,num_sample);
EU5_1 = zeros(100,1);
utilita5_1 = zeros (100,num_sample);
EU10_1 = zeros(100,1);
utilita10_1 = zeros (100,num_sample);
omegamax5_1 = zeros(120,1);
omegamax10_1 = zeros(120,1);
maxcal5_1 = zeros(120,1);
maxcal10_1 = zeros(120,1);

for t=1:120
      C = randn(num_sample,1);
      sd = sqrt(marg/n);
      condiz = mu+sd.*C;      % distribuzione di mu|sigma2,r
      Z = randn(num_sample,1);
      sdd = sqrt(t*marg);
      RT = t*condiz+sdd.*Z;   % distribuzione extrarendimenti
      w1 = (1-omega)*exp(t*rfree);
      W1=repmat(w1,1,num_sample);
      W = W1+omega*(exp(t*rfree+RT))';
      utilita5_1 = U(W,5);
      EU5_1 = mean(utilita5_1,2);
      [maxcal5_1(t),ind]=max(EU5_1);
      omegamax5_1(t)=omega(ind);
      utilita10_1 = U(W,10);
      EU10_1 = mean(utilita10_1,2);
      [maxcal10_1(t),ind]=max(EU10_1);
      omegamax10_1(t)=omega(ind) ;
end
```

```
% NO UNCERTAINTY

mup = mean(condiz);                 % posterior mean of mu
sigma2p = mean(marg);               % posterior mean of sigma2

W = zeros(100,num_sample);
EU5_2 = zeros(100,1);
utilita5_2 = zeros (100,num_sample);
EU10_2 = zeros(100,1);
utilita10_2 = zeros (100,num_sample);
omegamax5_2 = zeros(120,1);
omegamax10_2 = zeros(120,1);
maxcal5_2 = zeros(120,1);
maxcal10_2 = zeros(120,1);

for t = 1:120
      sd = sqrt(t*sigma2p);
      RT = t*mu + sd .* randn(num_sample,1);     %
RT~N(t*mup,t*sigma2p)
      w1 = (1-omega)*exp(t*rfree);
      W1=repmat(w1,1,num_sample);
      W = W1+omega*(exp(t*rfree+RT))';
      utilita5_2 = U(W,5);
      EU5_2 = mean(utilita5_2,2);
      [maxcal5_2(t),ind]=max(EU5_2);
      omegamax5_2(t)=omega(ind);
      utilita10_2 = U(W,10);
      EU10_2 = mean(utilita10_2,2);
      [maxcal10_2(t),ind]=max(EU10_2);
      omegamax10_2(t)=omega(ind) ;

end
```

Resampling

Here, the commands we used to implement the analysis of section 2.6 are listed. In this case the normality assumption of cumulative excess returns conditional on past data is loosened.

```
% UNCERTAINTY

a = (n-1)/2;
b = ((n-1)*sigma2)/2;
c = 1/b;
x = gamrnd(a,c,[num_samples,1]);
marg = 1./x ;                       % sigma2|r distribution Inverse-
Gamma

C = randn(num_samples,1);
sd = sqrt(marg/n);
condiz = mu+sd.*C ;                 % mu|sigma2,r distribution

sdd = sqrt(marg);
RT = zeros(num_samples,1);
omegamax5_3 = zeros(120,1);
omegamax10_3 = zeros(120,1);
```

```
ZTi = zeros(num_samples,1);
RTi = zeros(num_samples,1);

UT =(rt-mu)/(sqrt(sigma2));         % standardized returns.

for t=1:120
      ZTi = randsample(UT,num_samples,true);
      RTi = (ZTi.*sdd)+condiz;
      RT = RT+RTi;                          % resampled cumulative
returns
      W = zeros(100,num_samples);
      w1 = (1-omega)*exp(t*rfree);
      W1 = repmat(w1,1,num_samples);
      W  = W1+omega*(exp(t*rfree+RT))';
      utilita5_3 = zeros(100,num_samples);
      utilita5_3 = U(W,5);
      EU5_3 = zeros(100,1);
      EU5_3 = mean(utilita5_3,2);
      [maxcal5_3(t),ind]=max(EU5_3);
      omegamax5_3(t)=omega(ind);
      utilita10_3 = zeros(100,num_samples);
      utilita10_3 = U(W,10);
      EU10_3 = zeros(100,1);
      EU10_3 = mean(utilita10_3,2);
      [maxcal10_3(t),ind]=max(EU10_3);
      omegamax10_3(t)=omega(ind);
end

% NO UNCERTAINTY

RT = zeros(num_samples,1);
W = zeros(100,num_samples);
utilita5_4 = zeros(100,num_samples);
EU5_4 = zeros(100,1);
omegamax5_4 = zeros(120,1);
utilita10_4 = zeros(100,num_samples);
EU10_4 = zeros(100,1);
omegamax10_4 = zeros(120,1);

for t=1:120
      RTi = randsample(rt,num_samples,true);
      RT = RT+RTi;                          % resampled cumulative
returns
      w1 = (1-omega) * exp(t*rfree);
      W1=repmat(w1,1,num_samples);
      W = W1+omega*(exp(t*rfree+RT))';
      utilita5_4 = U(W,5);
      EU5_4 = mean(utilita5_4,2);
      [maxcal5_4(t),ind]=max(EU5_4);
      omegamax5_4(t)=omega(ind);
      utilita10_4 = U(W,10);
      EU10_4 = mean(utilita10_4,2);
      [maxcal10_4(t),ind]=max(EU10_4);
      omegamax10_4(t)=omega(ind) ;
end
```

Portfolio allocation with predictable returns

Here, the commands we used to implement the analysis of chapter 3 are listed.

We assumed predictable excess returns, the model we implemented takes therefore this form: $z_t = a + Bx_{t-1} + \varepsilon_t$, with $z_t' = (r_t, x_t')$ and $\varepsilon_t \sim \text{i.i.d.} N(0, \Sigma)$.

```
m_div = mean(div);
Z = [rt(2:n),div(2:n)];
I = ones((n-1),1);
X = [I,div(1:(n-1))];
D = inv(X'*X);
Chat = D*X'*Z;
S = (Z-X*Chat)'*(Z-X*Chat);
vecChat = Chat(:);


% Functions we implement in order to derive the predictive
distribution

function [varcov] = sposta(x)
      varcov=[x([1]),x([2]);x([2]),x([3])];
end

function [varcov] = sposta2(x)
      varcov = [x([1]),x([2]);x([3]),x([4])];
end

% Raising a matrix to a power
function [pot] = potenza(x,n);
if (n==0)
      pot=x^0;
elseif (n==1)
      pot=x;
else  pot = x;
     for (i=1:(n-1))
            pot = pot*x;
     end
end

% Mean of the predictive distribution
% B0^0 (t-1)* B0^1 +1* B0^(t-1)
function [totsum] = polinomio(x,n);
totsum = zeros(2,2);
for (i=1:n)
      sum = i*potenza (x,(n-i));
      totsum=totsum + sum;
end
end

function [totsum] = sommamatrix(x,n) ;
totsum = zeros(2,2);
for (i=1:n)
      sum = potenza(x,i);
      totsum = totsum + sum;
end
```

```
function [totsum] = sommamatrix2(x,n)
totsum = zeros(2,2);
for (i=0:(n-1))
      sum = potenza(x,i);
      totsum = totsum + sum;
end

% variance matrix of predictive distribution
function [totalsum] = sigmaric(x,t,sigma);
totalsum = zeros(2,2);
if (t==1)
      totalsum = totalsum
else
    for (n=1:(t-1))
            f = sommamatrix2(x,(n+1));
            sum = f * sigma * f';
            totalsum = totalsum +sum;
    end
end


 % NO UNCERTAINTY

cond = zeros(num_sample,4);
totalsomma1 = zeros(num_sample,1);
totalsomma2 = zeros(num_sample,1);
totalsommacov = zeros(num_sample,1);

for i = 1:num_sample
      sigmainv =  wishrnd(inv(S),(n-3));
      sigma = inv(sigmainv);
      totalsomma1(i) = sigma(1,1);
      totalsomma2(i) = sigma(2,2);
      totalsommacov(i) = sigma(1,2);
      varcov = kron(sigma,D);
      cond(i,:) = mvnrnd(vecChat',varcov);
end

% posterior means of parameters
a = [m_prev([1]);m_prev([3])];
B0 = zeros(2,2);
B0(1,2) = m_prev([2]);
B0(2,2) = m_prev([4]);
sigma = zeros(2,2);
sigma(1,1) = mean(totalsomma1);
sigma(1,2) = mean(totalsommacov);
sigma(2,1) = mean(totalsommacov);
sigma(2,2) = mean(totalsomma2);

% z_t starting value
ZZ = [Z((n-1),1);m_div] ;

RTT = zeros(num_sample,2);
RT = zeros(num_sample,1);
W = zeros(100,num_sample);
EU5_5 = zeros(100,1);
utilita5_5 = zeros (100,num_sample);
EU10_5 = zeros(100,1);
```

```
utilita10_5 = zeros (100,num_sample);
omegamax5_5 = zeros(120,1);
omegamax10_5 = zeros(120,1);
maxcal5_5 = zeros(120,1);
maxcal10_5 = zeros(120,1);

for t = 1:120
      musum = (polinomio(B0,t)*a+sommamatrix(B0,t)*ZZ);

      totalsum = zeros(2,2);
      sigmasum = sigma + sigmaric(B0,t,sigma);

      RTT = mvnrnd(musum',sigmasum,num_sample) ;
      RT = RTT(:,1
      w1 = (1-omega)*exp(t*rfree);
      W1=repmat(w1,1,num_sample);
      W = W1+omega*(exp(t*rfree+RT))';

      utilita5_5 = U(W,5);

      EU5_5 = mean(utilita5_5,2);

      [maxcal5_5(t),ind]=max(EU5_5);
      omegamax5_5(t)=omega(ind);


      utilita10_5 = U(W,10);
      EU10_5 = mean(utilita10_5,2);

      [maxcal10_5(t),ind]=max(EU10_5);
      omegamax10_5(t)=omega(ind) ;
end


% UNCERTAINTY

cond = zeros(num_sample,4);
a_c = zeros(num_sample,2);
B0_c = zeros(num_sample,4);
mupred = zeros(120,2);
ZT = zeros (120,num_sample);

ZZ = [Z((n-1),1);m_div] ;

for i = 1:num_sample
      sigmainv =  wishrnd(inv(S),(n-3));
      sigma = inv(sigmainv);
      varcov = kron(sigma,D);

      cond(i,:) = mvnrnd(vecChat',varcov);
      a_c(i,:) = [cond(i,1);cond(i,3)];
      B0_c(i,:) = [0;cond(i,2);0;cond(i,4)];
      totsigmaricors = zeros(2,2);

for t = 1:120
      mupred(t,:) = polinomio(sposta2(B0_c(i,:)),t)*a_c(i,:)'+
      sommamatrix(sposta2(B0_c(i,:)),t)* ZZ ;
```

```
      f = sommamatrix2(sposta2(B0_c(i,:)),t);
      sigmaricors = f * sigma * f';
      totsigmaricors = totsigmaricors+sigmaricors;
      sigmapred = sigma + totsigmaricors;
      zt = mvnrnd(mupred(t,:)',sigmapred);
      ZT(t,i) =  zt(1,1);
end
end


W = zeros(100,num_sample);
EU5_6 = zeros(100,1);
utilita5_6 = zeros(100,num_sample);
omegamax5_6 = zeros(120,1);
EU10_6 = zeros(100,1);
utilita10_6 = zeros(100,num_sample);
omegamax10_6 = zeros(120,1);
maxcal5_6 = zeros(120,1);
maxcal10_6 = zeros(120,1);


for t = 1:120
      w1 = (1-omega) * exp(t*rfree);
      W1 = repmat(w1,1,num_sample);
      W = W1+omega*(exp(t*rfree+ZT(t,:)'))';
      utilita5_6 = U(W,5);
      EU5_6 = mean(utilita5_6,2);
      [maxcal5_6(t),ind] = max(EU5_6);
      omegamax5_6(t) = omega(ind);
      utilita10_6 = U(W,10);
      EU10_6 = mean(utilita10_6,2);
      [maxcal10_6(t),ind] = max(EU10_6);
      omegamax10_6(t) = omega(ind);
end
```

Stock and bond portfolio allocation under uncertainty

Here, the commands we used to implement the analysis of chapter 4 are listed.

We assumed i.i.d. stock and bond excess returns of the form. $r_t = a + \varepsilon_t$, with $r_t' = (r_{1,t}, r_{2,t})$, $a' = (a_{r_1}, a_{r_2})$ and $\varepsilon_t \sim \text{i.i.d.} N(0, \Sigma)$

In addition we also inserted the commands for the optimal portfolio allocation under loss aversion that we handled in chapter 6.

```
% Continuously compounded bond returns

b = TR20YR;                   % 20-Yr treasury bond

wdb = zeros((m-1),1);
for i = 2:m;
      wdb(i) = log(b(i)/b(i-1));
end
```

```
wdb = wdb(2:m);
bt = wdb - rfree ;
% Predictor variables

div;
rf;
aaa;                          % aaa rated bonds
baa;                          % baa rated bonds
vix,                          % vix
tl = TB10YR_prova;
tl = tl(2:m);
rfl = log(1+tl/1200);
ts = rfl - rf;                % term spread
cs = baa - aaa;               % credit spread
delta                   % vector of combination of alpha and beta
alpha = delta(:,1);           % percentage allocated to stocks
beta  = delta(:,2);           % percentage allocated to bonds



% Functions we used in order to derive the distributions.

% Loss aversion function, alpha_1 = alpha_2 = 0.88, beta = 2.25

function [loss] = loss_aversion_case1(x)
loss = ((x-1).^0.88);
end

function [loss] = loss_aversion_case2(x)
loss = (-2.25.*((1-x).^0.88));
end

% Loss aversion function, alpha_1 = alpha_2 = 1, beta = 2.25

function [loss2] = loss_aversion2_case1(x);
loss2 = (x-1);
end

function [loss2] = loss_aversion2_case2(x);
loss2 = (-2.25*(1-x));
end

function [varcov7] = sposta7(x)
varcov7 =
[x([1]),x([2]),x([3]),x([4]),x([5]),x([6]),x([7]);x([8]),x([9]),x([1
0]),x([11]),x([12]),x([13]),x([14]);x([15]),x([16]),x([17]),x([18]),
x([19]),x([20]),x([21]);x([22]),x([23]),x([24]),x([25]),x([26]),x([2
7]),x([28]);x([29]),x([30]),x([31]),x([32]),x([33]),x([34]),x([35]);
x([36]),x([37]),x([38]),x([39]),x([40]),x([41]),x([42]);x([43]),x([4
4]),x([45]),x([46]),x([47]),x([48]),x([49])];
end

% Mean of the predictive distribution
* B0^0 (t-1)* B0^1 +1* B0^(t-1)
function [totsum7] = polinomio7(x,n);
totsum7 = zeros(7,7);
for (i=1:n)
    sum = i*potenza (x,(n-i));
```

```
    totsum7=totsum7 + sum;
end
end

function [totsum7] = sommamatrix7(x,n) ;
totsum7 = zeros(7,7);
for (i=1:n)
    sum = potenza(x,i);
    totsum7 = totsum7 + sum;
end
```

% Used to derive variance matrix of predictive distribution

```
function [totsum7] = sommamatrix7_2(x,n)
totsum7 = zeros(7,7);
for (i=0:(n-1))
    sum = potenza(x,i);
    totsum7 = totsum7 + sum;
end

function [totalsum7] = sigmaric7(x,t,sigma);
totalsum7 = zeros(7,7);
if (t==1)
    totalsum7 = totalsum7 ;
else
    for (n=1:(t-1))
        f = sommamatrix7_2(x,(n+1));
        sum = f * sigma * f';
        totalsum7 = totalsum7 +sum;
    end
end



% UNCERTAINTY

Z = [rt(2:n),bt(2:n)];
I = ones((n-1),1);
X = I;
D = inv(X'*X);
Chat = D*X'*Z;
S = (Z-X*Chat)'*(Z-X*Chat);
vecChat = Chat(:);

cond = zeros(1,2);
a_c = zeros(1,2);
B0_c = zeros(1,4);
mupred = zeros(1,2);
ZZ = [Z((n-1),1);Z((n-1),2)] ;
W = zeros(5149,1);

for i = 1:num_samples
      sigmainv =  wishrnd(inv(S),(n-2));
      sigma = inv(sigmainv);
      varcov = kron(sigma,D);
      cond = mvnrnd(vecChat',varcov);
      a_c = [cond(1),cond(2)];
      B0_c = [0,0,0,0];
```

```
      totsigmaricors = zeros(2,2);

    for t = 1:120
            mupred = (polinomio(sposta(B0_c),t)*a_c'+
            sommamatrix(sposta(B0_c),t)* ZZ)' ;
            f = sommamatrix2(sposta(B0_c),t);
            sigmaricors = f * sigma * f';
            totsigmaricors = totsigmaricors+sigmaricors;
            sigmapred = sigma + totsigmaricors;
            zt = mvnrnd(mupred,sigmapred);
            w1 = (1-alpha-beta)*exp(t*rfree);
            W = w1+delta*(exp(t*rfree+zt))';
            utilita5(:,t) =  U(W,5);
            utilita10(:,t) = U(W,10);
            bigger  = W(W >= 1);
            smaller = W(W < 1);
            loss_new_a(W >= 1)= loss_aversion_case1(bigger);
            loss_new_a(W < 1)  = loss_aversion_case2(smaller);
            loss_new_b(W >= 1) = loss_aversion2_case1(bigger);
            loss_new_b(W < 1)  = loss_aversion2_case2(smaller);
            loss_a(:,t) = loss_new_a;
            loss_b(:,t) = loss_new_b;
    end
      utilita_laA_1 = utilita_laA_1+ loss_a;
      utilita_laB_1 = utilita_laB_1+ loss_b;
      utilita5_1 = utilita5_1 + utilita5;
      tilita10_1 = utilita10_1 + utilita10;

end
        EU5_1 = utilita5_1 / num_samples;
        EU10_1 = utilita10_1 / num_samples;
        EU_laA_1 = utilita_laA_1 / num_samples;
        EU_laB_1 = utilita_laB_1 / num_samples;
        [maxcal5_1,ind]=max(EU5_1,[],1);
        betamax5_1 = beta(ind);
        alphamax5_1 = alpha(ind);
        deltamax5_1= delta(ind,:);
        [maxcal10_1,ind]=max(EU10_1,[],1);
        betamax10_1 = beta(ind);
        alphamax10_1 = alpha(ind);
        deltamax10_1 = delta(ind,:);
        [maxcal_laA_1,ind]=max(EU_laA_1,[],1);
        betamax_laA_1 = beta(ind);
        alphamax_laA_1 = alpha(ind);
        deltamax_laA_1= delta(ind,:);
        [maxcal_laB_1,ind]=max(EU_laB_1,[],1);
        betamax_laB_1 = beta(ind);
        alphamax_laB_1 = alpha(ind);
        deltamax_laB_1 = delta(ind,:);

% NO UNCERTAINTY

cond = zeros(num_samples,2);
totalsomma1 = zeros(num_samples,1);
totalsomma2 = zeros(num_samples,1);
totalsomma12 = zeros(num_samples,1);
for i = 1:num_samples
      sigmainv =  wishrnd(inv(S),(n-2=));
      sigma = inv(sigmainv
```

```
      totalsomma1(i) = sigma(1,1);
      totalsomma2(i) = sigma(2,2);
      totalsomma12(i) = sigma(1,2);
      varcov = kron(sigma,D
      cond(i,:) = mvnrnd(vecChat',varcov);
end

m = mean(cond,1);
st_mean = std(cond,1)
% parameters' posterior means
a = [m(1);m(2)];
B0 = zeros(2,2);
sigma = zeros(2,2);
sigma(1,1) = mean(totalsomma1);
sigma(1,2) = mean(totalsomma12);
sigma(2,1) = mean(totalsomma12);
sigma(2,2) = mean(totalsomma2);

% z_t starting value
ZZ = [Z((n-1),1);Z((n-1),2)] ; %%% AGGIUNGO L'ULTIMO VALORE DEL BOND

delta ;
alpha = delta(:,1);
beta = delta(:,2);

for t = 1:120
      EU5_2 = zeros(5149,1);
      utilita5_2 = zeros (5149,1);
      EU10_2 = zeros(5149,1);
      utilita10_2 = zeros (5149,1);
      EU_laA_2 = zeros(5149,1);
      utilita_laA_2 = zeros (5149,1);
      EU_laB_2 = zeros(5149,1);
      utilita_laB_2 = zeros (5149,1);
      musum = (polinomio(B0,t)*a+sommamatrix(B0,t)*ZZ);
      totalsum = zeros(2,2);
      sigmasum = sigma + sigmaric(B0,t,sigma);
      RT = mvnrnd(musum',sigmasum,num_samples)
      w1 = (1-alpha-beta)*exp(t*rfree);

    for i = 1:num_samples
            rtt = RT(i,:);
            W = w1+delta*(exp(t*rfree+rtt))';
            utilita5_2 = utilita5_2 + U(W,5);
            utilita10_2 = utilita10_2 + U(W,10);
            bigger  = W(W >= 1);
            smaller = W(W < 1);
            loss_new_a(W >= 1) = loss_aversion_case1(bigger);
            loss_new_a(W < 1)  = loss_aversion_case2(smaller);
            loss_new_b(W >= 1) = loss_aversion2_case1(bigger);
            loss_new_b(W < 1)  = loss_aversion2_case2(smaller);
            utilita_laA_2 = utilita_laA_2+ loss_new_a;
            utilita_laB_2 = utilita_laB_2+ loss_new_b;
    end

    EU5_2 = utilita5_2 / num_samples;
    EU10_2 = utilita10_2 / num_samples;
    EU_laA_2 = utilita_laA_2 / num_samples;
    EU_laB_2 = utilita_laB_2 / num_samples;
```

```
    [maxcal5_2(t),ind]=max(EU5_2);
    betamax5_2(t) = beta(ind);
    alphamax5_2(t) = alpha(ind);
    deltamax5_2(t,:)= delta(ind,:);
    [maxcal10_2(t),ind]=max(EU10_2);
    betamax10_2(t) = beta(ind);
    alphamax10_2(t) = alpha(ind);
    deltamax10_2(t,:) = delta(ind) ;
    [maxcal_laA_2(t),ind]=max(EU_laA_2);
    betamax_laA_2(t) = beta(ind);
    alphamax_laA_2(t) = alpha(ind);
    deltamax_laA_2(t,:)= delta(ind,:);
    [maxcal_laB_2(t),ind]=max(EU_laB_2);
    betamax_laB_2(t) = beta(ind);
    alphamax_laB_2(t) = alpha(ind);
    deltamax_laB_2(t,:)= delta(ind,:);
end

Z = [rt(2:n),bt(2:n),div(2:n),vix(2:n),cs(2:n),ts(2:n),rf(2:n)];
I = ones((n-1),1);
X = [I,div(1:(n-1)),vix(1:(n-1)),cs(1:(n-1)),ts(1:(n-1)),rf(1:(n-
1))];
D = inv(X'*X);
Chat = D*X'*Z;
S = (Z-X*Chat)'*(Z-X*Chat);
vecChat = Chat(:);
```

Portfolio allocation with predictable returns and five predictor variables

Here, the commands we used to implement the analysis of chapter 5 are listed.

We assumed predictable excess returns, the model we implemented takes therefore this form: $z_t = a + Bx_{t-1} + \varepsilon_t$, with $z_t' = (r_t, x_t')$, $x_t = (x_{1,t}, ..., x_{n,t})'$, and $\varepsilon_t \sim \text{i.i.d.} N(0, \Sigma)$.

In addition we also inserted the commands for the optimal portfolio allocation under loss aversion that we handled in chapter 6.

```
% NO UNCERTAINTY

m_div = mean(div);
m_vix = mean(vix);
m_ts = mean(ts);
m_cs = mean(cs);
m_rf = mean(rf);
m_pe = mean(pe);

% k = predictor variables
k=5;
```

```
for i = 1:num_samples
sigmainv =  wishrnd(inv(S),(n-k-2));
sigma = inv(sigmainv);
totalsomma1(i) = sigma(1,1);
totalsomma2(i) = sigma(2,2);
totalsomma3(i) = sigma(3,3);
totalsomma4(i) = sigma(4,4);
totalsomma5(i) = sigma(5,5);
totalsomma6(i) = sigma(6,6);
totalsomma7(i) = sigma(7,7);
totalsomma12(i) = sigma(1,2);
totalsomma13(i) = sigma(1,3);
totalsomma14(i) = sigma(1,4);
totalsomma15(i) = sigma(1,5);
totalsomma16(i) = sigma(1,6);
totalsomma17(i) = sigma(1,7);
totalsomma23(i) = sigma(2,3);
totalsomma24(i) = sigma(2,4);
totalsomma25(i) = sigma(2,5);
totalsomma26(i) = sigma(2,6);
totalsomma27(i) = sigma(2,7);
totalsomma34(i) = sigma(3,4);
totalsomma35(i) = sigma(3,5);
totalsomma36(i) = sigma(3,6);
totalsomma37(i) = sigma(3,7);
totalsomma45(i) = sigma(4,5);
totalsomma46(i) = sigma(4,6);
totalsomma47(i) = sigma(4,7);
totalsomma56(i) = sigma(5,6);
totalsomma57(i) = sigma(5,7);
totalsomma67(i) = sigma(6,7);
varcov = kron(sigma,D);
cond(i,:) = mvnrnd(vecChat',varcov);
end

m_prev = mean(cond,1);
st_mean_prev = std(cond,1)

% parameters' posterior means
a =
[m_prev(1);m_prev(7);m_prev(13);m_prev(19);m_prev(25);m_prev(31);m_p
rev(37)];
B0 = zeros(7,7);



% z_t starting value
ZZ = [Z((n-1),1);Z((n-1),2);m_div; m_vix; m_cs; m_ts; m_rf] ; %

RT = zeros(num_samples,7) ;
W = zeros(5149,1);
alpha  ;
beta = alpha(:,1);
omega = alpha(:,2);

loss_new_a = zeros(5149,1);
loss_new_b = zeros(5149,1);
for  t = 1:120
    musum = (polinomio7(B0,t)*a+sommamatrix7(B0,t)*ZZ);
```

```
    totalsum7 = zeros(7,7);
    sigmasum = sigma + sigmaric7(B0,t,sigma);
    RT = mvnrnd(musum',sigmasum,num_samples)
    RT = RT(:,1:2);
    EU5_5 = zeros(5149,1);
    utilita5_5 = zeros (5149,1);
    EU10_5 = zeros(5149,1);
    utilita10_5 = zeros (5149,1);
    EU_laA_5 = zeros(5149,1);
    utilita_laA_5 = zeros (5149,1);
    EU_laB_5 = zeros(5149,1);
    utilita_laB_5 = zeros (5149,1);
    w1 = (1-beta-omega)*exp(t*rfree);

    for i = 1:num_samples
            rtt = RT(i,:);
            W = w1+alpha*(exp(t*rfree+rtt))';
            utilita5_5 = utilita5_5 + U(W,5);
            utilita10_5 = utilita10_5 + U(W,10);
            bigger  = W(W >= 1);
            smaller = W(W < 1);
            loss_new_a(W >= 1) = loss_aversion_case1(bigger);
            loss_new_a(W < 1)  = loss_aversion_case2(smaller);
            loss_new_b(W >= 1) = loss_aversion2_case1(bigger);
            loss_new_b(W < 1)  = loss_aversion2_case2(smaller);
            utilita_laA_5 = utilita_laA_5+ loss_new_a;
            utilita_laB_5 = utilita_laB_5+ loss_new_b;
    end

    EU5_5 = utilita5_5 / num_samples;
    EU10_5 = utilita10_5 / num_samples;
    EU_laA_5 = utilita_laA_5 / num_samples;
    EU_laB_5 = utilita_laB_5 / num_samples;
    [maxcal5_5(t),ind]=max(EU5_5);
    omegamax5_5(t) = omega(ind);
    betamax5_5(t) = beta(ind);
    alphamax5_5(t,:)= alpha(ind,1:2);
    [maxcal10_5(t),ind]=max(EU10_5);
    omegamax10_5(t) = omega(ind);
    betamax10_5(t) = beta(ind);
    alphamax10_5(t,:) = alpha(ind,1:2) ;
    [maxcal_laA_5(t),ind]=max(EU_laA_5);
    omegamax_laA_5(t) = omega(ind);
    betamax_laA_5(t) = beta(ind);
    alphamax_laA_5(t,:)= alpha(ind,:);
    [maxcal_laB_5(t),ind]=max(EU_laB_5);
    omegamax_laB_5(t) = omega(ind);
    betamax_laB_5(t) = beta(ind);
    alphamax_laB_5(t,:)= alpha(ind,:);
end



% UNCERTAINTY

cond = zeros(1,42);
a_c = zeros(1,7);
B0_c = zeros(1,49);
mupred = zeros(1,7);
```

```
ZZ = [Z((n-1),1);Z((n-1),2);m_div; m_vix; m_cs; m_ts; m_rf] ;
W = zeros(5149,1);
loss_b = zeros(5149,120);

for i = 1:num_samples
      sigmainv =  wishrnd(inv(S),(n-k-2));
      sigma = inv(sigmainv);
      varcov = kron(sigma,D);
      cond = mvnrnd(vecChat',varcov);
      a_c =
[cond(1),cond(7),cond(13),cond(19),cond(25),cond(31),cond(37)];
            B0_c =
      [0,0,cond(2),cond(3),cond(4),cond(5),cond(6),0,0,cond(8),cond(
      9),cond(10),cond(11),cond(12),0,0,cond(14),cond(15),cond(16),c
      ond(17),cond(18),0,0,cond(20),cond(21),cond(22),cond(23),cond(
      24),0,0,cond(26),cond(27),cond(28),cond(29),cond(30),0,0,cond(
      32),cond(33),cond(34),cond(35),cond(36),0,0,cond(38),cond(39),
      cond(40),cond(41),cond(42)];
      totsigmaricors = zeros(7,7);

for t = 1:120
            mupred = (polinomio7(sposta7(B0_c),t)*a_c'+
            sommamatrix7(sposta7(B0_c),t)* ZZ)' ;    %%%% IL
            f = sommamatrix7_2(sposta7(B0_c),t);
            sigmaricors = f * sigma * f';
            totsigmaricors = totsigmaricors + sigmaricors;
            sigmapred = sigma + totsigmaricors;
            zt = mvnrnd(mupred,sigmapred);
            zt = zt (1,1:2);
            w1 = (1-beta-omega)*exp(t*rfree);
            W = w1+alpha*(exp(t*rfree+zt))'
            utilita5(:,t) =  U(W,5);
            utilita10(:,t) = U(W,10);
            bigger  = W(W >= 1);
            smaller = W(W < 1);
            loss_new_a(W >= 1)= loss_aversion_case1(bigger);
            loss_new_a(W < 1)  = loss_aversion_case2(smaller);
            loss_new_b(W >= 1) = loss_aversion2_case1(bigger);
            loss_new_b(W < 1)  = loss_aversion2_case2(smaller);
            loss_a(:,t) = loss_new_a;
            loss_b(:,t) = loss_new_b;
end
      utilita_laA_6 = utilita_laA_6+ loss_a;
      utilita_laB_6 = utilita_laB_6+ loss_b;
      utilita5_6 = utilita5_6 + utilita5 ;
      utilita10_6 = utilita10_6 + utilita10;
end

 EU5_6 = utilita5_6 / num_samples;
 EU10_6 = utilita10_6 / num_samples;
 EU_laA_6 = utilita_laA_6 / num_samples;
 EU_laB_6 = utilita_laB_6 / num_samples;
 [maxcal5_6,ind]=max(EU5_6,[],1);
 omegamax5_6 = omega(ind);
 betamax5_6 = beta(ind);
 alphamax5_6= alpha(ind,:);
 [maxcal10_6,ind]=max(EU10_6,[],1);
 omegamax10_6 = omega(ind);
```

```
betamax10_6 = beta(ind);
alphamax10_6 = alpha(ind,:);
[maxcal_laA_6,ind]=max(EU_laA_6,[],1);
omegamax_laA_6 = omega(ind);
betamax_laA_6 = beta(ind);
alphamax_laA_6= alpha(ind,:);
[maxcal_laB_6,ind]=max(EU_laB_6,[],1);
omegamax_laB_6 = omega(ind);
betamax_laB_6 = beta(ind);
alphamax_laB_6 = alpha(ind,:);
```

Bibliography

Aït-Sahalia, Y. and Brandt, M.W. (2001). Variable selection for portfolio choice. *Journal of Finance*, **56**, 1297-1351.

Bodie, Z. (1995). On the risk of stocks in the long run, *Financial Analysts Journal* 51, 18–22

Barberis, N. (2000). Investing for the long run when returns are predictable. *Journal of Finance*, **55**, 225-264.

Barberis, N., Huang, M. and Santos, T. (2001). Prospect theory and asset prices. *Quarterly Journal of Economics*, **116**, 1-53.

Barberis, N. and Thaler, R. (2003). A survey of behavioral finance. *Handbook of the Economics of Finance*. North-Holland, Amsterdam.

Bawa, V.S., Brown, S.J. and Klein, R.W. (1979). Estimation risk and optimal portfolio choice. North-Holland, Amsterdam.

Bell, D. (1982). Regret in decision making under uncertainty. *Operations Research*, **30**, 961-981.

Benartzi, S. and Thaler, R. (1995). Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics*, **110**, 73-92.

Berkelaar, A.B., Kouwenberg, R. and Post, T. (2004). Optimal portfolio choice under loss aversion. *Review of Economics and Statistics*, **86**, 973-987.

Brennan, M.J., Schwartz, E.S. and Lagnado, R. (1997). Strategic asset allocation. *Journal of Economic Dynamics and Control*, **21**, 1377-1403.

Campbell, J.Y. (1987). Stock returns and the term structure. *Journal of Financial Economics*, **18**, 373-399.

Campbell, J.Y. (1991). A variance decomposition for stock returns. *Economic Journal*, **101**, 157-179.

Campbell, J.Y. and Schiller R. (1988). The dividend-price ratio, expectations of future dividends and discount factors. *Review of Financial Studies*, **1**, 195-227.

Campbell, J.Y. and Schiller R. (1988). Stock prices, earnings and expected dividends. *Journal of Finance*, **43**, 661-676.

Campbell, J.Y. and Viceira , L.M. (1998). Who should buy long-term bonds?. *Working paper*, Harvard University.

Campbell, J.Y. and Viceira, (2002). Strategic asset allocation: portfolio choice for long-term investors. Oxford University, New York.

Caporin, M. (2012). Equity and CDS Sector Indices: Dynamic Models and Risk Hedging.

Chacko, G. and Viceira, L.M. (1999). Dynamic consumption and portfolio choice with stochastic volatility in incomplete markets. NBER *Working paper, no. 7377*.

Chew, S. (1989). Axiomatic utility theories with the betweenness property. *Annals of Operations Research*, **19**, 273-298.

Chew, S. and MacCrimmon, K. (1979). Alpha-nu choice theory: an axiomatization of expected utility. *Working Paper*, University of British Columbia.

Fama, E.F. and French, K.R. (1988). Dividend yields and expected stock returns. *Journal of Financial Economics*, **22**, 3-27.

Fama, E.F. and French, K.R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics*, **25**, 23-49.

Ferson, W.E. and Harvey, R. (1991). The variation of economic risk premiums. *Journal of Political Economy*, **99**, 385-415.

Gaiofatto, E. (2009). Investing for the long run: predictability and loss aversion.

Gemmill, G., Hwang, S. and Salmon, M. (2006). Performance measurement with loss aversion. *Journal of Asset Management*, **7**, 190-207.

Hodrick, R. (1992). Dividend yields and expected stock returns: Alternative procedures for inference and measurement, *Review of Financial Studies* 5, 357–386.

Kahneman, D. and Tversky, A. (1979). Prospect Theory: an analysis of decision under risk. *Econometrica*, **47**, 263-291.

Kan, R. and Zhou, G. (2007). Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis*, **42**, 621-656.

Kandel, S. and Stambaugh, R.F. (1996). On the predictability of stock returns: an asset allocation perspective. *Journal of Finance*, **51**, 385-424.

Keim, D.B. and Stambaugh, R.F. (1986). Predicting returns in the stock and bond markets. *Journal of Financial Economics*, **17**, 357-390.

Lo, A.W. and MacKinlay, A.C. (1988). Stock market prices do not follow random walks: evidence from a simple specification test. *Review of Financial Studies*, **1**, 41-66.

Lynch, Anthony W. (2001). Portfolio choice and equity characteristics: Characterizing the hedging demands induced by return predictability. *Journal of Financial Economics*, **62**, 67-130.

Markovitz, H. (1952). Portfolio selection. Journal of Finance, **7**, 77-91.

Merton, R. (1969). Lifetime portfolio selection: the continous-time case. *Review of Economics and Statistics,* **51**, 247-257.

Merton, R. (1973). An intertemporal capital asset pricing model. *Econometrica*, **41**, 867-887.

Samuelson, P. (1969). Lifetime portfolio selection by dynamic stochastic programming. *Review of Economics and Statistics,* **51**, 239-246.

Siegel, J. (1994). Stocks for the long run. McGraw-Hill.

Stambaugh, R.F. (1999). Predictive regressions. *Journal of Financial Economics*, **54**, 375-421.

Tversky, A. and Kahneman, D. (1992). Advanced in Prospect Theory: cumulative representation of uncertainty. *Journal of risk and uncertainty*, **5**, 297-323.

Zellner, A. (1971). An introduction to bayesian inference in econometrics. John Wiley and Sons, New York.