NTNU
Innovation and Creativity

Financial Risk, Risk Appetite and the Macroeconomic Environment

Petter Haugen

Master of Science in Physics and Mathematics
Submission date: June 2006
Supervisor: Jacob Laading, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

Problem Description

This thesis seeks to establish a methodology to reveal whether the risk appetite - the willingness of investors to bear risk - is dependent on the macroeconomic environment and, if present, to quantify this dependency. To do so a case study is carried out and the collected data is used in a multi-resolution risk analysis of the relationship between actual financial risk on one hand and business cycles and the market risk on the other.

Assignment given: 03. February 2006
Supervisor: Jacob Laading, MATH

Preface

This is my master thesis written at NTNU, Dept. of Industrial Mathematics, during the spring of 2006.

Several persons have assisted different parts of this thesis. First and foremost I would like to thank my supervisor Jacob Laading; it was him who introduced me to finance in his lectures during the fall of 2005 and this thesis could not have been written without his help and guidance.

I also would like to thank Egil Matsen at Dept. of Economics (NTNU) for recommendations on relevant literature and Torbjørn Eika at SSB for his elaborative comments on his and P.R. Johansen's article which I have used extensively in this thesis.

A special thanks go to Anders Pederstad and Rannveig Eriksen as they assisted the proofreading of the thesis.

All errors and omissions are the sole responsibility of the author.

Petter Haugen
Trondheim, June 30, 2006.

Abstract

This thesis seeks to establish a methodology to reveal whether the risk appetite held by investors is dependent on the macroeconomic environment and, if present, to quantify this dependency.

To do so a generic model is built and a case study is carried out with data from DnBNOR. The available data consists of the daily profit and losses together with the number and volume of transactions made in a currency portfolio owned by DnBNOR and some selected timeseries on exchange rates, all against NOK. Also, timeseries on the gross national product and consumer price index are collected from Statistics Norway (SSB).

In the process of building the model, the thesis sets out the theoretical foundation for different risk measurement concepts and gives a presentation of the theory on business cycles as this is used to classify and measure the macroeconomic environment. The model is built with a Bayesian approach and implemented in `WinBUGS`. The use of Bayesian statistics is motivated by different time resolution of the data; some of the data is observed every day while other parts are observed each quarter.

The thesis' main idea is to decompose the relevant part of the economy in one micro-economic and one macroeconomic state. The microeconomic state is unique for each day while the macroeconomic state accounts for one quarter; together they give the expected risk appetite for each day.

In this way the impact from the macroeconomic state is quantified and the results show that the macroeconomic state *is* statistically significant for the risk appetite. As this is a case study one needs more data and research before any universal valid conclusions can be made.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The objective of this thesis is to establish a methodology whose intention is to reveal whether the risk appetite held by investors is dependent on the macroeconomic environment and, if present, to quantify this dependency. The dependency is of interest, both to academics and practitioners, as some suggest that changes in asset prices are due to exogenous changes in the risk appetite. There also exist trading strategies that are based on the idea that it is possible to quantify the movements in asset prices that are due to changes in risk appetite.[1]

The literature does not always make clear distinctions between the terms "risk appetite" and "risk aversion". The trend is that the financial community uses "risk appetite" while academics prefer "risk aversion". This thesis will in particular discuss the definition used by Gai & Vause (2006) who attempt to make a clear distinction between the two terms. This distinction implies that the risk appetite, *the willingness of investors to bear risk*, is affected by both *the degree of which investors dislike uncertainty about their future consumption possibilities*, i.e. their risk aversion, and *the level of this uncertainty* which is determined by the macroeconomic environment.

To answer whether the risk appetite is dependent on the macroeconomic environment, a statistical model is built and applied with data collected from one portfolio in a case study. The model is based upon the idea that it is possible to describe the economy relevant for the portfolio at hand with the use of two states; one describing the macroeconomic environment and one describing the microeconomic environment. Both of these may or may not affect the risk appetite held by the portfolio administrator. This idea is incorporated by a Bayesian approach where several microeconomic states belong to the same macroeconomic state. The latter is described by a business cycle estimated with a Hodrick-Prescott filter.

Because this thesis do not follow a standard approach it gives a thorough presentation of all steps taken towards the complete model. All programmed code for all treatment and adjustment made to any part of the data can be found in the appendix or in the text itself. The thesis assumes that readers have some background in economics and statistics, although all key concepts are introduced in a way that does not demand a specific background.

The empirical findings in this thesis partly support Gai & Vause's (2006) assumption as the risk appetite is found to be dependent on the macroeconomic environment. At the same time the causality behind this dependency is questioned as one might suspect that the macroeconomic environment does not only affect the *level* of uncertainty surrounding

[1] See for instance Misina (2003).

future consumption, but perhaps also the *degree* of which investors dislikes such uncertainty, i.e. their risk aversion. However, as this is a case study is it impossible to state any universal valid conclusions.

The following chapter concentrates on market risk and how it should be measured. Chapter 3 lay out the theoretical foundations used to classify and measure the macroeconomic environment in the terms of business cycles. The collected data is presented in chapter 4 before chapter 5 constructs the model and explains its statistical attributes. The latter chapter also contains some theoretical aspects on Bayesian statistics and risk concepts such as risk aversion and risk appetite. In chapter 6 the results from the model is presented and interpreted. Chapter 7 concludes and states the necessary reservations about the findings. Figure 1.1 illustrates the structure of this thesis with its main parts and model approach.

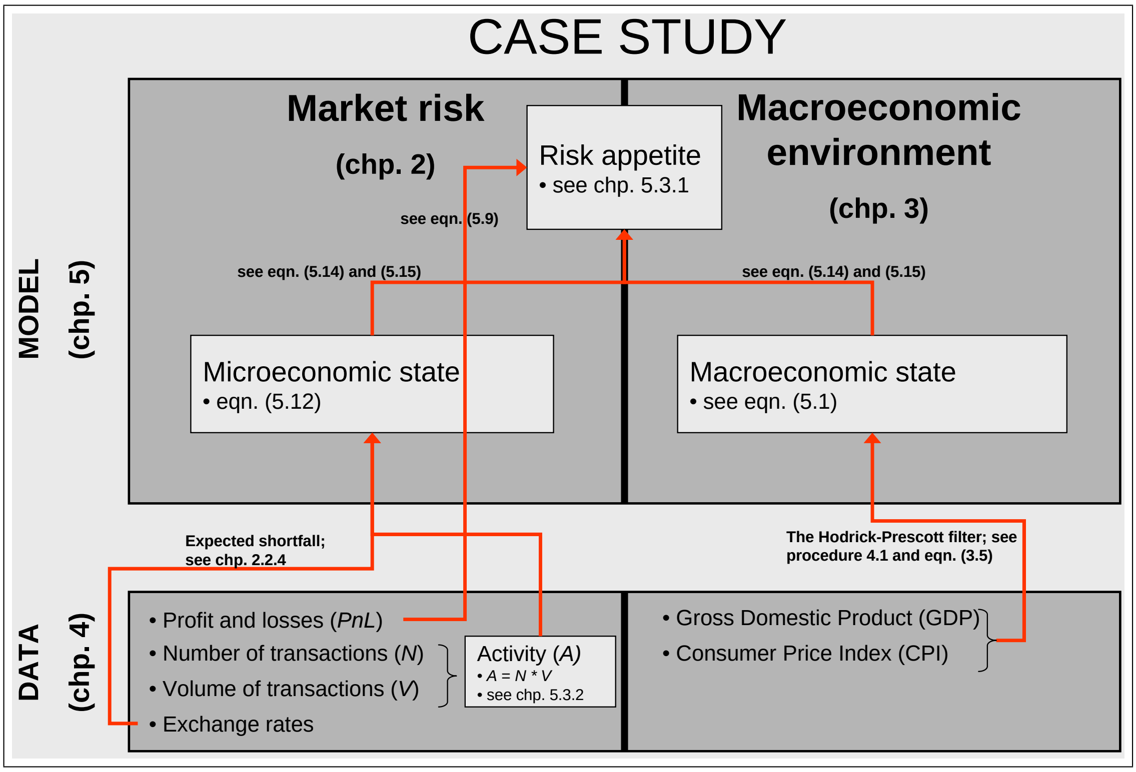


Figure 1.1: The structure of the thesis with its main parts and model approach.

Chapter 2

Market Risk

As this thesis sets out to establish a possible dependency between the risk appetite and the macroeconomic environment, there is an apparent need to establish an understanding of financial risk and how it should be defined and measured. This is the intention in the present chapter.

This chapter starts out with an overview of different approaches to risk measurement; risk measures based on the loss distribution receives most attention. After measures such as Value at Risk (VaR) and Expected Shortfall (ES) is defined and exemplified, a quick introduction on how to estimate different risk measures is given. Definitions and discussions surrounding risk concepts such as risk appetite, risk aversion and risk premium is left out here as it is introduced in the building of the model in chapter 5.1.

2.1 Financial Risk

John F. Kennedy once pointed out that

> *"When written in Chinese the word crisis is composed to two characters. One represents danger, and the other represents opportunity".*[1]

This is very much like the concept of risk; there is an upside and a downside. In finance, like in every aspect of life, one tries to minimize the downside while maximizing the upside; one wants to minimize the risk and maximize the returns.

Financial risk is often divided into three main categories: Market risk, credit risk and operational risk. The former is *the risk of a change in the value of a portfolio due to changes in the value of its underlying components.* These components can be stocks, bonds, exchange rates, etc. Credit risk is *the risk that your counter party will not be able to fulfil his promised repayments*, also called the default of the borrower. Operational risk is defined by The Basel Committee on Banking Supervision as *"the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events."*[2] The three categories are not mutually exclusive, and in addition two other types of risk are likely to be present: Model risk and liquidity risk. The former is the risk related to using an inaccurate or misspecified model for measuring risk. Liquidity risk is the risk,

[1] According to `http://quotations.about.com/`.

[2] This is a group within The Bank for International Settlements and the publication where this definition is written can be found at `http://www.bis.org/publ/bcbs118.htm`.

roughly speaking, that one cannot sell or buy an investment quickly enough (to minimize a loss) due to lack of marketability.

It is desirable to measure risk for a number of reasons. First and foremost because it can, by definition, result in big losses. Regulation of risk is also demanded by law as the society today is depending on stable financial institutions; there are several examples on how devastating the lack of risk management can be, see e.g. Jorion (2001, chapter 2) for a collection. And in some businesses, in particular the insurance business, risk measuring is an absolute necessity because the companies' income (the premiums) results from a transaction of risk from a counter part.

2.2 Different Risk Measures

According to McNeil, Frey & Embrechts (2005) there are four different categories of approaches to risk measurement: the notional-amount approach, factor-sensitivity measures, risk measures based on the loss distribution and risk measures based on scenarios.

The notional-amount approach is the simplest of the four. This approach measure risk as the sum of the notional values in the portfolio where each notional value can possibly be weighted with a factor that determines how risky each of the underlying securities are. Its advantage is the simplicity, but the disadvantages are several. First, this kind of risk measure will not take netting effects into account; e.g. risk stemming from short and long positions in the same security will be counted separately. Diversification effects in the portfolio are also ruled out. The approach also has problems with risk connected to derivatives, whereas the notional amount of the underlying and the value of the derivative position can differ widely.

Factor-sensitivity measures give the change in portfolio value for a given predetermined change in one of the underlying risk factors. The "Greeks", famous from option pricing, are examples of factor-sensitivity measures.[3] The advantage of factor-sensitivity measures is that they give important information about how robust the value of a portfolio is with respect to changes in risk factors, but they do not actually measure the overall riskiness of a position. They also, like the notional-amount approach, have problems aggregating risk; e.g. one cannot add the Greeks from the derivatives in a portfolio.

Risk measures based on loss distribution are the most modern approach and subject of much recent research. They measure risk as the statistical quantities describing the loss distribution of the portfolio over some predetermined time horizon. Examples on this "loss-distribution-approach" are Value-at-Risk (VaR) and Expected Shortfall (ES), both derived and explained below. This way of measuring risk overcomes most of the disadvantages of the other two methods above; it measures risk in a single statistic and can also be aggregated across portfolios.[4] The use of loss distributions is justified as the worrying part of risk is without doubt connected to losses and not so much the profit one can make. Other facts supporting the choice of the loss distribution approach, is that it make sense on all levels of aggregation, it reflects netting and diversification effects and it can be compared across portfolios. The loss distribution is formalized in a mathematical notation below as it will be used extensively throughout this thesis.

Scenario-based risk measures are, as the names says, concerned about future scenarios with insulated or simultaneous changes in the risk factors; e.g. a $x\%$ rise in an exchange

[3]A concise explanation of the Greeks can be found in Crouchy, Galai & Mark (2001, page 186) while Wilmott (2001, chapter 10) gives a more detailed presentation of the subject.

[4]This does not apply to VaR calculated the in traditional way, see chapter 2.2.3 below.

rate with a $y\%$ fall in some stock market index. In this way a worst case scenario can be found and the risk, measured as the possible loss, emerges. The risk of the portfolio is then measured as the maximum loss of the portfolio under all scenarios, where the most extreme ones can be weighted down if appropriate.

2.2.1 The Loss Distribution

Let the value, measured at time t, of a portfolio be denoted by V_t. The loss of a portfolio over the period from t to $t+1$ can then be defined as

$$L_{t+1} = -(V_{t+1} - V_t). \tag{2.1}$$

In financial literature it is common to regard the portfolio's value to be dependent on one or more *risk factors*. If the d numbers of risk factors are denoted as the random vector $\boldsymbol{Z}_t = (Z_{t,1}, Z_{t,2}, \ldots, Z_{t,d})'$, then the value of a portfolio at time t can be written as

$$V_t = f(t, \boldsymbol{Z}_t), \tag{2.2}$$

where $\boldsymbol{Z}_t$ is assumed observable at time t. How to choose the risk factors and the function f will depend on the portfolio at hand; an example of such a risk factor can be the price on a stock in a portfolio at time t, S_t, such that $\boldsymbol{Z}_t = (S_{t,1}, S_{t,2}, \ldots, S_{t,d})'$ when the portfolio has d stocks.

Further, it is the changes in these risk factors that will determine the risk (of losses). This makes it convenient to define a new random vector $\boldsymbol{X}_t = \boldsymbol{Z}_t - \boldsymbol{Z}_{t-1}$ such that the loss over one period now becomes

$$L_{t+1} = -(f(t+1, \boldsymbol{Z}_{t+1}) - f(t, \boldsymbol{Z}_t)) = -(f(t+1, \boldsymbol{Z}_t + \boldsymbol{X}_{t+1}) - f(t, \boldsymbol{Z}_t)) \tag{2.3}$$

because $\boldsymbol{X}_{t+1} = \boldsymbol{Z}_{t+1} - \boldsymbol{Z}_t$. A first-order approximation, L^{Δ}_{t+1}, to the loss in equation (2.3) is

$$L^{\Delta}_{t+1} = -\left(f_t(t, \boldsymbol{Z}_t) + \sum_{i=1}^{d} f_{Z_i}(t, \boldsymbol{Z}_t) X_{t+1,i}\right) \tag{2.4}$$

provided that the function f is differentiable. The subscripts to f denote partial derivatives. Equation (2.4) is handy because the losses now becomes a linear function of the risk factors. But still, this is a approximation and its quality depends on the time resolution (the changes in the risk factors are likely to be small if the time horizon is short) and, of course, to what degree the portfolio's value actually is linear in the risk factors. One can also note that the first term, $f_t(t, \boldsymbol{Z}_t)$, is often dropped when the horizon is short. An application of equation (2.4) is found below in example 2.1.

Example 2.1 Consider a portfolio consisting of stocks in d companies. The price for stock i at time t is denoted $S_{t,i}$ and the number of stocks in company i is λ_i. The value of the portfolio at time t is therefore $V_t = \sum_{i=1}^{d} \lambda_i S_{t,i}$. Standard financial theory says that the risk factors are the log transformed stocks prices such that $\boldsymbol{Z}_t = (\ln S_{t,1}, \ldots, \ln S_{t,d})$, which makes $V_t = \sum_{i=1}^{d} \lambda_i e^{Z_{t,i}}$ and $\boldsymbol{X}_{t+1,i} = \ln S_{t+1,i} - \ln S_{t,i} = \ln(S_{t+1,i}/S_{t,i})$. Then the loss of the portfolio at time t is

$$
\begin{aligned}
L_{t+1} &= -(V_{t+1} - V_t) \\
&= -\left(\sum_{i=1}^{d} \lambda_i S_{t+1,i} + \sum_{i=1}^{d} \lambda_i S_{t,i}\right) \\
&= -\sum_{i=1}^{d} \lambda_i S_{t,i}(e^{X_{t+1,i}} - 1),
\end{aligned}
$$

and the linearized loss becomes

$$
\begin{aligned}
L_{t+1}^{\Delta} &= -\left(\frac{\partial}{\partial t} f(t, Z_{t,i}) + \sum_{i=1}^{d} \frac{\partial}{\partial Z_i} f(t, Z_{t,i})\right) \\
&= -\left(0 + \sum_{i=1}^{d} \lambda_i S_{t,i} X_{t+1,i}\right).
\end{aligned}
$$

Here the time horizon is assumed so small that $f_t(t, Z_{t,i}) = 0$. Instead of counting the stocks with λ, one can write $w_{t,i} = \frac{\lambda_i S_{t,i}}{V_t}$, i.e. w is the weight (proportion) of stock i in the portfolio at time t, such that

$$
L_{t+1}^{\Delta} = -V_t \sum_{i=1}^{d} w_{t,i} X_{t+1,i}.
$$

Further, suppose that $\boldsymbol{X}$ comes from a distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ. Then, using the expectancy operator $\mathbb{E}(\cdot)$ and general rules for the mean and variance of linear combinations, one get

$$
\begin{aligned}
\mathbb{E}(L_{t+1}^{\Delta}) &= -V_t \sum_{i=1}^{d} w_{t,i} \mathbb{E}(X_{t+1,i}) \\
&= -V_t \boldsymbol{w}' \boldsymbol{\mu}
\end{aligned}
$$

and

$$
\begin{aligned}
\text{var}(L_{t+1}^{\Delta}) &= \text{var}(-V_t \boldsymbol{w}' \boldsymbol{X}) \\
&= V_t^2 \boldsymbol{w}' \Sigma \boldsymbol{w}.
\end{aligned}
$$

2.2.2 Coherent Risk Measures

"A coherent risk measure" is an attempt to answer the question "What is a good risk measure?". It originates from Artzner, Delbaen, Eber & Heath (1999) who present four desirable properties for measures of risk; if all four are met the risk measure in question is labelled "coherent".

The four properties are posted as axioms. Here, the presentation of these axioms is done with a notation similar to the one used by McNeil et al. (2005).[5] In order to present these axioms in a consistent way, a formal definition of a risk measure is needed:[6]

[5] This notation differs from Artzner et al.'s (1999) notation as the former concentrate on *losses* while the last on *the future value*, of a portfolio.

[6] This is the definition used by McNeil et al. (2005).

Definition 2.1 Let Δ represent the time horizon and fix some probability space $(\Omega, \mathcal{F}, P)$.[7] Denote by $L_0(\Omega, \mathcal{F}, P)$ the set of all random variables on the measurable space $(\Omega, \mathcal{F})$. Financial risks are represented by a set $\mathcal{M} \subset L_0(\Omega, \mathcal{F}, P)$ of random variables, which is interpreted as portfolio losses over a time horizon Δ. Further, assume that $\mathcal{M}$ is a convex cone, i.e. that $L_1, L_2 \in \mathcal{M}$ implies that $L_1 + L_2 \in \mathcal{M}$ and $\lambda L_1 \in \mathcal{M}$ for every $\lambda > 0$. Then, risk measures are real-valued functions $\rho : \mathcal{M} \to \mathbb{R}$.

The interpretation of $\rho(L)$ is *the capital that should be added to a position with loss given by L, so that position again becomes acceptable to an external or internal risk controller*. The axioms that a risk measure should possess to be called coherent, are listed below (axioms 2.1 - 2.4).

Axiom 2.1 For all $L \in \mathcal{M}$ and every $l \in \mathbb{R}$ we have $\rho(L + l) = \rho(L) + l$. This axiom is called "translation invariance".

Axiom 2.2 For all $L_1, L_2 \in \mathcal{M}$ we have $\rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2)$. This axiom is called "subadditivity".

Axiom 2.3 For all $L \in \mathcal{M}$ and every $\lambda \geq 0$ we have $\rho(\lambda L) = \lambda \rho(L)$. This axiom is called "positive homogeneity".

Axiom 2.4 For all $L_1, L_2 \in \mathcal{M}$ such that $L_1 \leq L_2$ almost surely[8] we have $\rho(L_1) \leq \rho(L_2)$. This axiom is called "monotonicity".

Definition 2.2 A risk measure ρ whose domain includes the convex cone $\mathcal{M}$ is called coherent if it fulfills axioms 2.1-2.4.

Axiom 2.1 is quite intuitive; adding or subtracting a sure initial amount l to a position with loss L alters the capital requirements with exactly l. The short example 2.2 illustrates this:

Example 2.2 Consider a position which has experienced a loss L. Adding an amount $\rho(L)$ to this position results in the adjusted loss $\tilde{L} = L - \rho(L)$ such that $\rho(\tilde{L}) = \rho(L) - \rho(L) = 0$.

Because axiom 2.2 rules out Value at Risk (see below) as a coherent risk measure, this is the most debated one. But there are several aspects that speaks in favour of this axiom: First; if you merge two separate portfolios, standard theory tells you that this should not create additional risk (but it could *reduce* overall risk). Second; if a risk measure fails to comply with axiom 2.2, it could not been used in centralized risk management; this is exemplified in example 2.3. Third; if a risk measure did not comply with axiom 2.2, it could be possible for a broker to meet demands from internal or external risk managers by dividing the portfolio in as many sub-portfolios needed because this would *reduce* the aggregated risk.

Example 2.3 A risk manager is responsible to oversee that the risk of the overall loss $\rho(L) = \rho(L_1 + L_2) \leq M$ where M is an upper limit. If the risk measure fulfils axiom 2.2 he can do this by allocating upper risk limits, M_1 and M_2, to each of his two brokers such that $M_1 + M_2 \leq M$ and $\rho(L_i) \leq M_i,\ i = 1, 2$. Then $\rho(L_1 + L_2) \leq M_1 + M_2 \leq M$.

If axiom 2.2 holds, it is easy to verify axiom 2.4: Subadditivity implies that

[7] The probability space with its components is defined in appendix A.1.

[8] "Almost surely" can be read as "with probability one"; this terminology is used in much of the literature, see for instance Øksendal (2003).

$$\rho(nL) = \rho(L + \cdots + L) \leq \rho(L) + \rho(L) + \cdots + \rho(L) = n\rho(L)$$

for $n \in \mathbb{N}$. Because there are no netting or diversification between losses in this portfolio, it would seem reasonable to claim that equality should hold in the equation above, leading to positive homogeneity.

Axiom 2.4 states that positions leading to higher losses in every state of the world requires more capital. One can also note that subadditivity and positive homogeneity together implies that ρ is convex, i.e. that

$$\rho(\lambda L_1 + (1-\lambda)L_2) \overset{\text{Axiom 2.2}}{\leq} \rho(\lambda L_1) + \rho((1-\lambda)L_2) \overset{\text{Axiom 2.3}}{=} \lambda\rho(L_1) + (1-\lambda)\rho(L_2) \quad (2.5)$$

for all $L_1, L_2 \in \mathcal{M}$ and $\lambda \in [0,1]$. Some argue that for large values of λ it would be reasonable to have that $\rho(\lambda L) \geq \lambda\rho(L)$ to penalize a concentration of risk and the possible liquidity problems. This is Föllmer & Schied's (2002) idea:

> *"In many situations, however, the risk of a position might increase in a nonlinear way with the size of the position. For example, an additional liquidity risk may arise if a position is multiplied by a large factor. This suggests to relax the conditions of positive homogeneity and of subadditivity and to require, instead of* [axiom 2.2] *and* [axiom 2.3], *the weaker property of convexity* [equation (2.5)]. *Convexity means that diversification does not increase the risk, i.e., the risk of a diversified position* [$\lambda L_1 + (1-\lambda)L_2$] *is less or equal to the weighted average of the individual risks."*

where the brackets [...] mean that their notation has been adjusted to fit the one used in this thesis.

Below this thesis will present different risk measures and first out is one that is very popular among practitioners; Value at Risk.

2.2.3 The Value-at-Risk (VaR) Measure

VaR, as we know it today, is quite new. It was born in Global Derivatives Study Group (1993) and soon became a branch standard. Roughly speaking it tells you, in monetary terms, *how much of your portfolio you risk to loose with a given confidence level in a given time period.* The formal definition is:

Definition 2.3 Given the confidence level $\alpha \in (0,1)$, VaR_α of a portfolio at confidence level α is the smallest amount l such that the probability that the loss L exceeds l is no larger than $(1-\alpha)$. In mathematical terms:

$$VaR_\alpha = \inf\{l \in \mathbb{R} : \; P(L > l) \leq 1 - \alpha\} = \inf\{l \in \mathbb{R} : \; F_L(l) \leq \alpha\} \quad (2.6)$$

where $F_L(l) = P(L \leq l)$.

Definition 2.3 states that VaR_α is simply a quantile of the loss distribution. Typical values for α is 0.95 or 0.99 and the time horizon, which is not explicitly stated in the definition but implied in the loss distribution, is often chosen to one day or one week when applied to market risk (more on the choice of time horizon and time scaling can be found in appendix A.2). In credit and operation risk management the time horizon normally is longer, often one year. Example 2.4 together with figure 2.1 shows how the VaR_α measure is computed when the risk factors, and therefore also the losses, are normally distributed.

Example 2.4 Assume that the loss distribution is normally distributed, i.e. $F_L \sim N(\mu, \sigma^2)$, over some (unspecified) time horizon Δ. Then

$$VaR_\alpha = \mu + \sigma\Phi^{-1}(\alpha) \tag{2.7}$$

where μ is the expected loss, σ^2 its variance and Φ is the distribution function (making $\Phi^{-1}(\alpha)$ the α-quantile of Φ). This can be seen from the fact that

$$F_L(VaR_\alpha) = P(L \leq VaR_\alpha = l) = P\left(\frac{L-\mu}{\sigma} \leq \Phi^{-1}(\alpha)\right) = \Phi(\Phi^{-1}(\alpha)) = \alpha$$

.

The procedure is illustrated in figure 2.1, with $\mu = 0$ and $\sigma^2 = 1$, where the VaR_α measure is shown as the vertical full line. Here $VaR_{\alpha=0.99}$ is 2.326, i.e. if the losses are measured in million dollars this says that *there is a* 1 *% risk that the loss over the period* Δ *will be greater than* 2.326 *MILL USD.*[9]

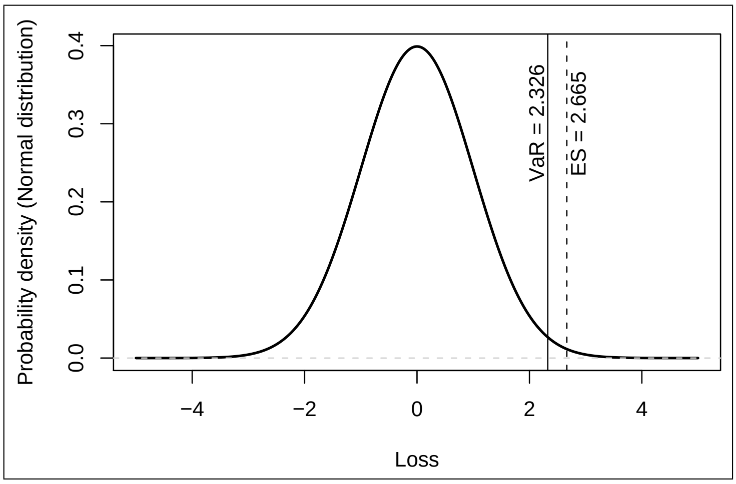


Figure 2.1: Illustration of the risk measures Value at Risk (see example 2.4) and Expected Shortfall (see equation (2.9)). Assumes standard normally distributed losses, $L \sim N(0,1)$ and uses $\alpha = 0.99$.

Another thing one should note is that VaR_α does not give any information about the distribution of losses *beyond* the VaR_α-value; that is also the reason for VaR_α *not* beeing a coherent risk measure as it does not comply with axiom 2.2, subadditivity.[10] A graphical illustration of a loss distribution that points the necessity of coherency is found in figure 2.2. A measure that *does* classify as a coherent risk measure is the Expected shortfall (ES) measure.

2.2.4 The Expected Shortfall (ES) Measure

The ES measure is based on the VaR measure, but unlike VaR it succeeds to take the whole distribution into account. Definition 2.4 defines ES.[11]

[9] This is not completely true; McNeil et al. (2005) points out that this interpretation does not account for model or liquidity risk.

[10] Though, in (McNeil et al. 2005, Theorem 6.8, page 242) it is shown that if the risk factors is *elliptically distributed*, then the VaR_α measure will be coherent.

[11] This is McNeil et al.'s (2005) definition.

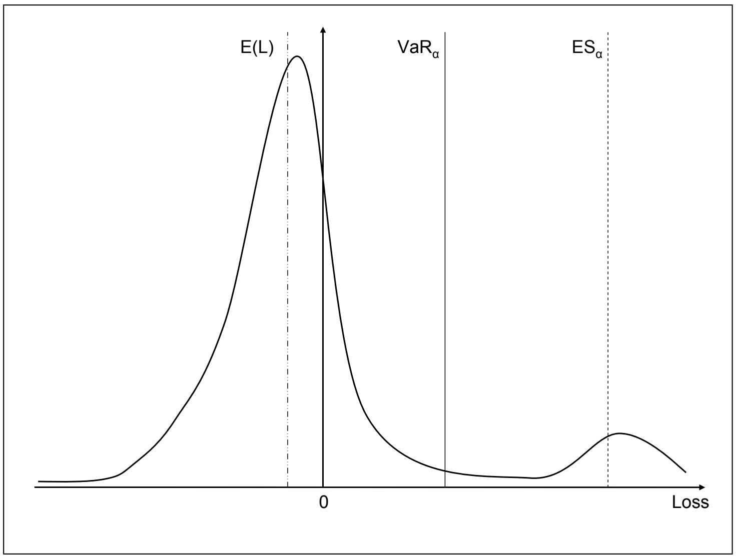


Figure 2.2: Stylized example of a loss distribution that makes the VaR_α measure non-subadditive. Here the loss distribution has a "bump" in its right tail which makes the ES_α measure more trustworthy than the VaR_α measure with the same confidence level.

Definition 2.4 For a loss L with $E(|L|) < \infty$ and distribution function F_L the expected shortfall at confidence level $\alpha \in (0,1)$ ES_α is defined as

$$ES_\alpha = \frac{1}{1-\alpha}\int_\alpha^1 q_u(F_L)\mathrm{d}u \tag{2.8}$$

where $q_u(F_L)$ is the (general) quantile function of the loss distribution F_L.

ES is closely connected to VaR; this becomes obvious when the quantile function is written $q_u(F_L) = VaR_u(L)$. Then the ES measure is written as

$$ES_\alpha = \frac{1}{1-\alpha}\int_\alpha^1 VaR_u \mathrm{d}u.$$

So, instead of fixing the α, one calculates the average of VaR_α over all levels $\alpha \leq u \leq 1$. And this opens for an alternative interpretation of ES_α as *the average loss, given that this loss is greater than* VaR_α. The ES_α measure, given that $F_L \sim N(0,1)$, becomes

$$ES_\alpha = \mu + \sigma\frac{\phi(\Phi^{-1}(\alpha))}{1-\alpha} \tag{2.9}$$

where ϕ is the density function for the standard normal distribution. Its value is shown in figure 2.1 as the vertical dotted line.

2.2.5 Other Risk Measures

Besides VaR and ES there exist other risk measures based on the loss distribution such as its *variance* and *lower and upper partial movements.*

The former has been, and still is, extensively used in both theory and practice; this is to a large degree due to the work on portfolio theory by Harry M. Markowitz.[12] The variance is well understood as an analytical tool, but it has some drawbacks as it e.g. assumes that the second moment of the loss (or "profit-and-loss-distribution", PnL) exists and that the distribution is symmetrical. Later in this thesis, a variance measure will be used for the risk appetite; see equation (5.9).

[12] See for instance Markowitz (1952).

Partial moments measures of risk focus on the upper and lower part of a distribution. With the sign convention in this thesis it is natural to look at the upper tail of the loss distribution and therefore put the attention on the upper partial moments, UMP:

$$UPM(k,q) = \int_q^\infty (l-q)^k \mathrm{d}F_L(l)$$

where q is a reference point and k gives the weight on the losses deviations from q.

Two special cases are obtained when one choose $k = 0 \Rightarrow UPM(k,q) = \int_q^\infty \mathrm{d}F_L(l) = P(L \geq q)$; with $k = 2$ and $q = E(L)$ a risk measure called the upper semivariance of L is obtained.[13]

2.3 Estimating Market Risk

There are three standard methods for estimating market risk measures; the variance-covariance method; historical simulation and the Monte Carlo approach, each with its strengths and weaknesses. These methods are presented below.

2.3.1 The Variance-Covariance Method

The variance-covariance method is a parametric method which assumes that the risk factor changes is multivariate normally distributed, $\boldsymbol{X}_{t+1} \sim N_d(\boldsymbol{\mu}, \Sigma)$, where $\boldsymbol{\mu}$ is the mean vector and Σ is the variance (or variance-covariance) matrix. Further, it assumes that the linearization in equation (2.4) is sufficiently accurate such that

$$L_{t+1}^{\Delta} = -(c_t + \boldsymbol{b}_t' \boldsymbol{X}_{t+1}) \tag{2.10}$$

for some constant c_t and constant vector $\boldsymbol{b}_t$, both assumed known at time t. This means that

$$L_{t+1}^{\Delta} \sim N(-c_t - \boldsymbol{b}_t'\boldsymbol{\mu}, \boldsymbol{b}_t'\Sigma\boldsymbol{b}_t) \tag{2.11}$$

which makes it is easy to calculate VaR_α and ES_α from equations (2.7) and (2.9) respectively.

To make a practical procedure out of this, estimates on the expected loss and the covariance matrix are needed. Such estimates can be found using historical risk factor changes, $\boldsymbol{X}_{t-n+1}, \ldots, \boldsymbol{X}_t$, and unbiased estimators of $\boldsymbol{\mu}$ and Σ.

The strength of the variance-covariance method is its simplicity, but the linearization of losses in respect of the risk factor changes and the normality assumption makes it a crude method for prediction as both these assumptions generally are inadequate approximations.

An application of the variance-covariance method can be found in chapter 5.3.2 below.

2.3.2 Historical Simulation

The historical simulation procedure is different from the variance-covariance method as it does not assume any particular distribution for the risk factor changes. Instead it concentrates on the *empirical* distribution of data $\boldsymbol{X}_{t-n+1}, \ldots, \boldsymbol{X}_t$. Based on these historical observations one can construct historically simulated losses $\tilde{L}_s$:

[13] For papers concerned with partial moments, see Unser (2000) and Price, Price & Nantell (1982).

$$\{\tilde{L}_s = -(f(t+1, \boldsymbol{Z}_t + \boldsymbol{X}_s) - f(t, \boldsymbol{Z}_t)) : \ s = t-n+1, \dots, t\} \tag{2.12}$$

where n is the number of historical observations. Then these values $(\tilde{L}_s)$ show what would happen to the current portfolio if the risk factor changes on day s were to recur.

Normally VaR_α and ES_α are estimated with the use of empirical quantile estimation; for instance, if $n = 1000$ one would calculate $\tilde{L}_{t=1}, \dots, \tilde{L}_{t=1000}$, sort them and use tenth largest value as an estimate of $VaR_{\alpha=0.99}$. The $ES_{\alpha=0.99}$ would be the mean of the ten largest values.

This method is easy to implement and assumes nothing about the distribution of the risk factor changes. But, its downside is the need for very large quantities of data.

2.3.3 The Monte Carlo Method

The first step in the Monte Carlo approach is to fit the collected data on risk factor changes to a parametric model. Then one draws realizations from this model and in this way one gets an empirical distribution of the loss in the next period. More precisely:

Algorithm 2.1 The Monte Carlo Method

Fit the data $\boldsymbol{X}_{t-n+1}, \dots, \boldsymbol{X}_t$ to some parametric distribution $D(\boldsymbol{\theta})$
for $i = 1, \dots, m$ **do**
  Draw $\boldsymbol{X}_{t+1}^{(i)} \sim D(\boldsymbol{\theta})$
  Calculate $\tilde{L}_{t+1}^{(i)} = -(f(t+1, \boldsymbol{Z}_t + \boldsymbol{X}_{t+1}^{(i)}) - f(t, \boldsymbol{Z}_t))$
end for
return $\tilde{\boldsymbol{L}}_{t+1}$

where $\tilde{\boldsymbol{L}}_{t+1}$ is the vector of all realizations with length m. Then one can use this vector to estimate VaR_α and ES_α as in the historical simulation method.

Because this method assumes some parametric model the results will of course depend on just how good this model fits reality. It also has large computational costs as every simulation requires a revaluation of the portfolio in question. But, on the bright side, this method allows m to be chosen much larger than n in the historical simulation method such that the estimated empirical quantile will be more accurate.

2.4 Relevant Literature

There are several textbooks on risk management on the market. This chapter is mainly based on McNeil et al. (2005). This is relatively technical and requires some statistical and mathematical background from the reader. Less technical, but still useful, presentations can be found in e.g. Jorion (2001), Crouchy et al. (2001) and Holton (2003). Glasserman (2004) provides a introduction to simulation techniques used in finance.

One important issue not treated in the text above, that still deserves a comment, is the difference between the unconditional and conditional loss distribution. The latter gives origin to e.g. (G)ARCH[14] models which, in general, has shown to provide more accurate predictions than models build on the unconditional loss distribution (see for instance McNeil et al. (2005, chapter 2.3.6)). An introduction to GARCH models is given in for example Brooks (2002, chapter 8) and McNeil et al. (2005, chapter 4).

[14] (Generalized) AutoRegressive Conditionally Heteroscedastic.

Chapter 3

The Macroeconomic Environment

As the last chapter concentrated on different aspects surrounding market risk measurement, this chapter will concentrate on how to define and measure the macroeconomic environment. This is essential to answer the question posed in the introduction: *Is the risk appetite dependent on the macroeconomic environment, in what way and to what extent.*

There is no standard approach to define and measure "the macroeconomic environment". This thesis argues that one possible solution is to use the theory and terminology from the research on *business cycles* as this provides an generic, but precise, terminology of the macroeconomic performance of an economy.

The analysis of business cycles is an important and much debated field within economics and is subject of extensive research. This chapter will not give a full business cycle analysis, but limit itself to a more descriptive decomposition of the Norwegian gross domestic product (GDP)[1] by using the framework from the theory on business cycles. The results from this decomposition will be used as explanatory variables in the forthcoming model.

3.1 Terminology and Definitions

Even though this thesis not will produce a complete business cycle analysis, some terminology must be introduced. This thesis follows Johansen & Eika (2000) who defines the different states of the economy as shown in figure 3.1.

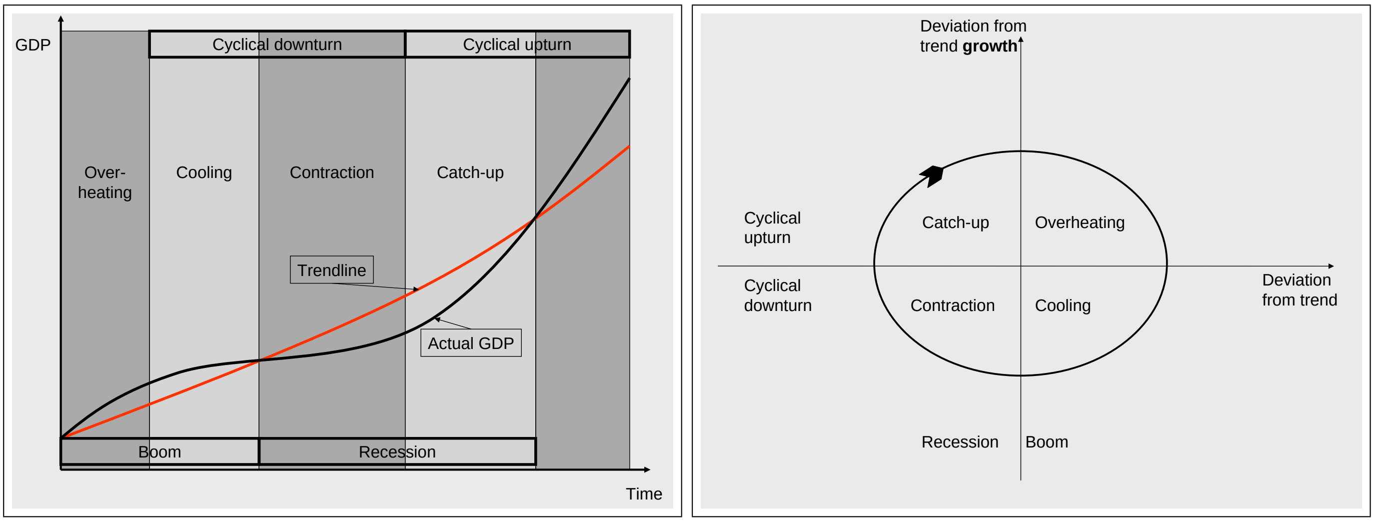
Figure 3.1 (a) shows, in a stylized way, the development of an economy, measured by its GDP, over time. The most distinctive attribute is that the GDP grows as time passes. This is an empirical fact for developed economies, but its cause will not be discussed here.[2] The *trend* in the GDP is shown in red and it is the deviation from this trend that is used as measures of the business cycles. The trend can be interpreted as an estimate on the activity in the economy when all resources are fully exploited.[3]

The definitions of the different stages in the business cycles are found in figures 3.1(a-b): When the economy is *above (below)* its trend, we talk about a *boom (recession)* and

[1] "Bruttonasjonalprodukt (BNP)" in Norwegian.

[2] Growth theory today owes much to the work of Solow and Swan in the 1950s, see e.g. Heijdra & van der Ploeg (2002) for a graduate textbook approach on growth theory.

[3] Some macroeconomic models, included the one used by Statistics Norway (SSB), called KVARTS, assumes that a *steady* growth path will exploit the resources better and therefore also give a higher growth over time.



(a) (b)

Figure 3.1: (a): Stylized path of growth defining terminology used in this thesis. (b): An abstract phase diagram of a business cycle.

when the growth in the actual GDP is *smaller (higher)* than the trend growth, this is called a *cyclical downturn (upturn)*. Also, when the deviation between the actual GDP and the trend (within one cycle) is at its *maximum (minimum)*, this is called a *peak (trough)*. At these extreme points the actual growth equals trend growth.

Figure 3.1 (b) shows the phase diagram where the state of the economy is classified with respect to its deviation from trend and deviation from trend growth: When there is a boom and the economy is still growing faster than the trend growth, this is called *overheating*, but as soon as the the actual growth becomes smaller than the trend growth, the economy is said to enter a period with *cooling*. When the actual GDP drops below its corresponding trend value, the economy is said to enter a *contraction* and this will continue until the growth in actual GDP becomes larger than the trend growth; then the economy is *catching up* before the same cycle starts again.

It is important to emphasize that the business cycles are not deterministic cycles which means that *how* one chooses to calculate the trend is of great importance to the results. Below the Hodrick-Prescott filter (HP-filter), which is a method for extracting this trend, is introduced and later implemented in chapter 4.2.[4]

3.2 The Hodrick-Prescott Filter

The HP-filter was first published in a working paper in 1981 by Robert J. Hodrick and Edward C. Prescott and later published as Hodrick & Prescott (1997). In short; this filter decomposes a data series, in this case a series of GDP figures, into a trend component and a cyclical component as shown in equation 3.1 below:

$$y_t = g_t + c_t \quad \text{for} \quad t = 1, \ldots, T. \tag{3.1}$$

[4]See for instance Kydland & Prescott (1990) for an application of this filter to the US economy.

In the equation above y_t is the (observed) GDP and g_t is the trend (sometimes referred to as "growth") component (unobserved) while c_t is the cyclical components (unobserved), all on logarithmic form.[5] The filter itself is the solution to the (convex) optimization problem:[6]

$$\min_{\{g_t\}_{t=1}^{T}} \left\{ f(y_t, g_t) = \sum_{t=1}^{T} c_t^2 + \lambda \sum_{t=2}^{T-1} [(g_{t+1} - g_t) - (g_t - g_{t-1})]^2 \right\}. \tag{3.2}$$

Here $(g_{t+1} - g_t) - (g_t - g_{t-1})$ is recognized as $\Delta^2 g_{t+1}$ and $c_t = y_t - g_t$.[7]

3.2.1 How to choose the correct λ?

λ is often referred to as a smoothing parameter and it determines the relative weight that is put on the second difference in the minimization problem (as 1 is the weight put on the cyclical components); as λ becomes larger the estimated trend curve becomes smoother.[8]

In (Hodrick & Prescott 1997, page 4) it is written that

> *"If the cyclical components and the second differences of the growth components were identically and independently distributed, normal variables with means zero and variances σ_1^2 and σ_2^2 (which they are not), the conditional expectation of the g_t, given the observations, would be the solution to equation (3.2) when $\sqrt{\lambda} = \sigma_1^2/\sigma_2^2$". ... Our prior view is that a 5 percent cyclical component is moderately large, as is a one-eight of 1 percent change in the growth rate in a quarter. This led us to select $\sqrt{\lambda} = 5/(1/8) = 40$ or $\lambda = 1600$ as a value for the smoothing parameter."*

They also performed a sensitivity analysis on the λ and found that the results are quite consistent even if the smoothing parameter varies between 400 and 6400, i.e. λ can vary quite a lot while the trend estimate remains stable.

Johansen & Eika (2000) argues that the Norwegian economy is more volatile and therefore needs a larger λ; in (Johansen & Eika 2000, page 27, footnote 4) it is written that

> *"Given the deep and long recession in the Norwegian economy that took root towards the end of the 1980s, a high weight ($\lambda = 40000$) has been applied to the straight line in order to obtain a trend that is reasonably consistent with underlying developments in the supply of resources during the period (capital stock and working-age population). This weight also results in a relationship between recessions/booms in the 1980s and 1990s, which is fairly consistent with our a priori perceptions of the business cycles in this period."*

[5] Here capital letters are reserved for variables that has not been transformed in any way and small letters are denote variables that has undergone a logarithmic transformation ($y_t = \ln(Y_t)$).

[6] This particular formulation is found in Reeves, Blyth, Triggs & Small (2000) and differs from the original article, Hodrick & Prescott (1997), in the four endpoints. If one were to use the original formulation, $\min_{\{g_t\}_{t=1}^{T}} \left\{ \sum_{t=1}^{T} c_t^2 + \lambda \sum_{t=1}^{T} [(g_{t+1} - g_t) - (g_t - g_{t-1})]^2 \right\}$, the resulting vector with trend estimates, $\mathbf{g}$, would be four elements shorter than the input vector, $\mathbf{y}$; i.e. data would be "lost" with the original formulation. With the data used in this thesis the two methods gives approximately the same answer.

[7] Δ^2 being the second difference operator: $\Delta^2 g_t = \Delta(g_t - g_{t-1}) = (g_t - g_{t-1}) - (g_{t-1} - g_{t-2})$.

[8] When $\lambda \to \infty$ all effort is put on making the second difference as small as possible. Since this is a quadratic term the best result the filter can give is zero and this is the same as saying that the derivative should be constant which implies that the trend is estimated as a straight line.

This thesis' implementation of the HP-filter will choose to follow Johansen & Eika (2000) with $\lambda = 40000$. This is because, besides what is said in the quote above, the GDP data used in this thesis also includes the petroleum sector, a sector which is left out in the GDP numbers used by Johansen & Eika (2000). This makes the cyclical components in the unadjusted data used here, even more volatile than Johansen & Eika's (2000).[9] To see the actual effects of different λ's when the HP-filter is applied to the GDP data presented in the next chapter, look at figure A.3 in appendix A.5.

3.2.2 The Solution of the HP-filter

The first order conditions (FOC) for the minimization problem (3.2) are derived by solving $\frac{\partial f(y,g)_t}{\partial g_t} = 0$ for each $t = 1, \ldots, T$:

$$\begin{aligned}
c_1 &= \lambda(g_1 - 2g_2 + g_3) \\
c_2 &= \lambda(-2g_1 + 5g_2 - 4g_3 + g_4) \\
c_t &= \lambda(g_{t-2} - 4g_{t-1} + 6g_t - 4g_{t+1} + g_{t+2}) \quad \text{for } t = 3, 4, 5, \ldots, T-2 \\
c_{T-1} &= \lambda(g_{T-3} - 4g_{T-2} + 5g_t - 2g_t) \\
c_T &= \lambda(g_{T-2} - 2g_{T-1} + g_T).
\end{aligned}$$

This can be written in a more compact way using matrix notation:

$$\mathbf{c} = \lambda F \mathbf{g}, \tag{3.3}$$

where F is a quadratic $T \times T$ matrix,[10]

$$F = \begin{bmatrix}
1 & -2 & 1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
-2 & 5 & -4 & 1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
1 & -4 & 6 & -4 & 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & 1 & -4 & 6 & -4 & 1 & 0 & \cdots & \cdots & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & & \vdots \\
\vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & \vdots \\
0 & \cdots & \cdots & 0 & 1 & -4 & 6 & -4 & 1 & & 0 \\
0 & \cdots & \cdots & \cdots & 0 & 1 & -4 & 6 & -4 & & 1 \\
0 & \cdots & \cdots & \cdots & \cdots & 0 & 1 & -4 & 5 & & -2 \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1 & -2 & & 1
\end{bmatrix}$$

such that

$$\mathbf{y} = (\lambda F + I)\mathbf{g}. \tag{3.4}$$

The estimated growth is then

$$\hat{\mathbf{g}} = (\lambda F + I)^{-1}\mathbf{y}. \tag{3.5}$$

[9] In (Johansen & Eika 2000) the GDP data is adjusted for value added from the petroleum sector and ocean transport and the results called "GDP mainland Norway". Johansen & Eika (2000) defends this adjustment with an argument claiming that the petroleum sector can experience large fluctuations in activity without this resulting in corresponding fluctuations in other sectors. Here this is assumed to be of minor significance and no adjustment is made.

[10] One can note that the sum of each column in the F-matrix is zero, $\sum_{t=1}^{T} c_t = 0$, as it should.

3.3 Extensions of the Business Cycle Analysis

The procedure above is normally just the first step in an analysis of the business cycle. In e.g Johansen & Eika (2000) the main part of the analysis is to find what *drives* the business cycles, in this thesis the intent is merely to *classify* the state of the economy using the same framework.

It is possible to decompose the economy in many different ways. In Johansen & Eika (2000) the business cycles in the Norwegian economy are decomposed in five main factors or impulses (international product markets, money and foreign exchange markets, oil prices and petroleum investment, fiscal policy, inventory investment) plus a residual factor (i.e. impulses that by definition cannot be explained), and it is the impacts from these that are assumed to drive the business cycle. Other possible decompositions are for instance the familiar Keynesian approach: $Y = C + I + G + NX$ where Y is the GDP, C is consumption, I is investment, G is government consumption and NX is net export; see for instance Kydland & Prescott (1990) for an example on this type of decomposition.

3.4 Relevant Literature

The literature on business cycles, or economical fluctuations as they are sometimes called, is large and growing. For a comprehensive overview of recent work one can look into Rebelo (2005) and The Royal Swedish Academy of Sciences (2004, chapter 3); the latter also offers a concise history of the research on real business cycles. For a textbook introduction on the subject one can read Romer (2001), in particular chapter four and five.

Chapter 4

Description of the Data Material

This chapter describes the data which is used in the model presented in the next chapter. As stated in the introduction, this thesis carries out a *case study* which means that the data is not representative for the market as it is collected from just one single portfolio.

The portfolio data and time series on different exchange rates are presented before the data used to extract the business cycles. This latter part also includes the results from this thesis' implementation of the HP-filter introduced in the previous chapter, although this implementation should be regarded as a part of the methodology chosen. The reason for presenting its results in this chapter, is that they are used as data in the statistical model presented in the next chapter.

4.1 Data from DnBNOR

The data from DnBNOR is collected from a portfolio consisting of foreign currency, mainly pound sterling (GBP), US dollars (USD), Swiss franc (CHF), Japanese yen (JPY) and Euros (EUR). This part of data has three dimensions, specified in the upper part of table 4.1. Further, time series with these main currencies are collected, specified in the lower part of table 4.1.[1]

Table 4.1: The data material collected from DnBNOR.

Variable	Unit	Original resolution	Resolution used	Period (DD/MM/YY)[a]
PnL, PnL^b	NOK	daily	daily	20/11/03 - 10/11/04
Number of transactions, N	#	daily	daily	20/11/03 - 10/11/04
Volume traded, V	MNOK	daily	daily	20/11/03 - 10/11/04
NOK/USD	-	daily	daily	01/01/02 - 30/12/05
CHF/USD	-	daily	daily	01/01/02 - 30/12/05
USD/EUR	-	daily	daily	01/01/02 - 30/12/05
USD/GBP	-	daily	daily	01/01/02 - 30/12/05
JPY/USD	-	daily	daily	01/01/02 - 30/12/05

[a]The series are broken in weekends and holidays.
[b]Profit and losses.

4.1.1 The Portfolio Data

Figure 4.1 shows a graphical representation of the data collected from the DnBNOR-portfolio; table 4.2 gives the correlation matrix. From figure 4.1(b) it is quite obvious

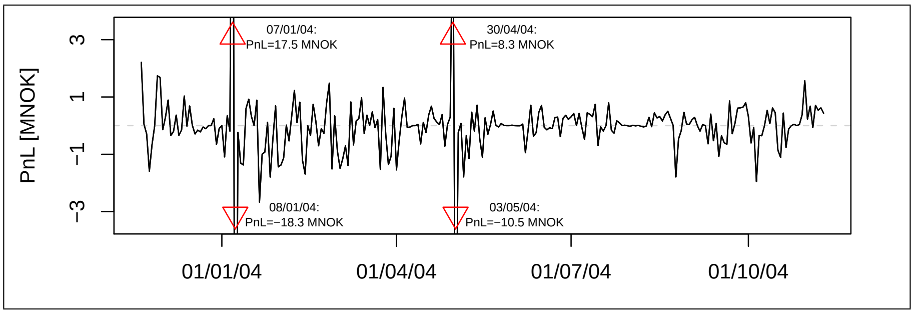
[1]These are also provided by DnBNOR.

that the number and volume of transactions are highly correlated and this is confirmed by the correlation matrix as $\mathrm{Cor}(N, V) = 0.938$. On the other hand, there seems not to be a significant relationship between PnL and N and/or V; a simple regression with PnL as the dependent variable give p-values for N and V to 0.53 and 0.97, respectively.

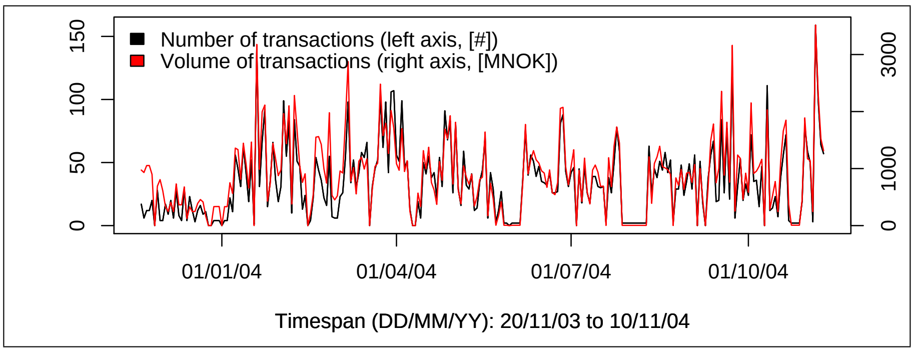
Another thing one should note about figure 4.1(a) is the four outliers on the dates 07/01/04, 08/01/04, 30/04/04 and 03/05/04. The fact that these follow one another two and two, and that they are approximately of the same, but opposite, size, is a bit suspicious as it is not the same volatility in the underlying exchange rates on these dates. So, this can either be a data flaw, or a consequence of a bet made by the portfolio administrator(s). As there is no apparent reason not to trust the data source, the data will not be altered.

Table 4.2: Correlation matrix for the currency portfolio from DnBNOR.

	Number	PnL	Volume
Number	1	-0.096	0.938
PnL	-0.096	1	-0.092
Volume	0.938	-0.092	1



(a)



(b)

Figure 4.1: The portfolio data collected from DnBNOR. (a): Profit and losses. (b): Number and volume of transactions.

4.1.2 The Exchange Rate Data

Figure 4.2 shows the exchange rates in NOK, the log-returns and a normal QQ-plot of these log-returns for each of the exchange rates.[2] The time period is from 01/01/02 to 30/12/05 (DD/MM/YY), including 1044 observations of the five exchange rates, 261 in each year. The correlation matrix for the exchange rates is found in table 4.3, while the correlation matrix for the log-returns on the exchange rates are found in table 4.4. The former shows, for example that when YEN, measured in NOK, becomes more expensive, the same happens with USD. The opposite is true for EUR; when EUR increases its value measured in NOK, USD tends to decrease in value.

The correlation matrix in table 4.4 states that all returns are positive correlated. This is probably due to the fact that it is primarily what happens to the Norwegian krone that determines the returns on all the exchange rates; e.g. if the Norwegian Central Bank decides to increase its interest, this will, ceteris paribus, make NOK more attractive to foreign investors and the NOK becomes more expensive measured in other currencies.[3]

Table 4.3: Correlation matrix for the currencies in figure 4.2(a).

	NOK/USD	NOK/CHF	NOK/EUR	NOK/GBP	NOK/JPY
NOK/USD	1	-0.072	-0.391	0.349	0.713
NOK/CHF	-0.072	1	0.885	0.707	0.448
NOK/EUR	-0.391	0.885	1	0.556	0.173
NOK/GBP	0.349	0.707	0.556	1	0.709
NOK/JPY	0.713	0.448	0.173	0.709	1

Table 4.4: Correlation matrix for the log-returns of the exchange rates.

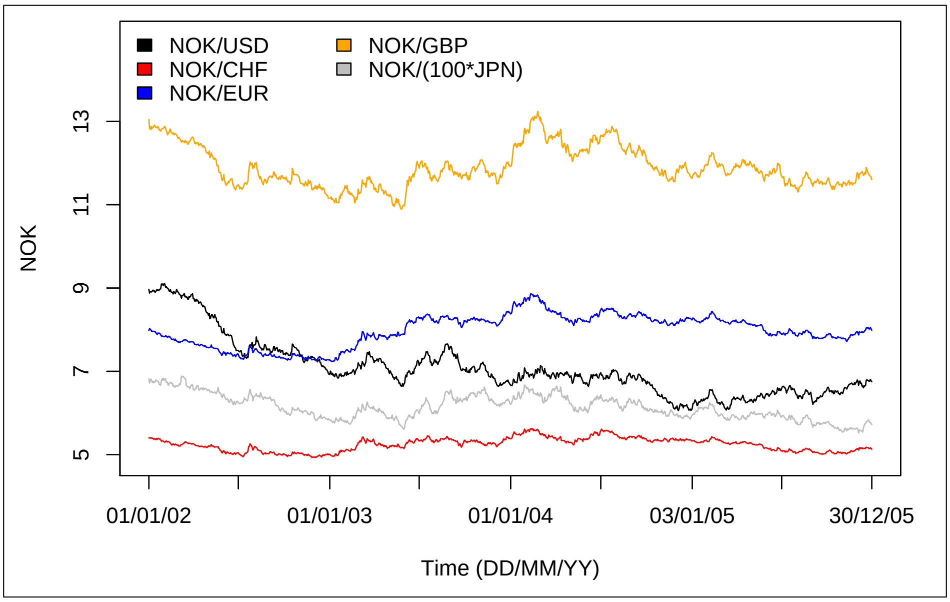
	$\ln\left(\frac{\frac{NOK_t}{USD_t}}{\frac{NOK_{t-1}}{USD_{t-1}}}\right)$	$\ln\left(\frac{\frac{NOK_t}{CHF_t}}{\frac{NOK_{t-1}}{CHF_{t-1}}}\right)$	$\ln\left(\frac{\frac{NOK_t}{EUR_t}}{\frac{NOK_{t-1}}{EUR_{t-1}}}\right)$	$\ln\left(\frac{\frac{NOK_t}{GBP_t}}{\frac{NOK_{t-1}}{GBP_{t-1}}}\right)$	$\ln\left(\frac{\frac{NOK_t}{JPY_t}}{\frac{NOK_{t-1}}{JPY_{t-1}}}\right)$
$\ln\left(\frac{\frac{NOK_t}{USD_t}}{\frac{NOK_{t-1}}{USD_{t-1}}}\right)$	1	0.313	0.459	0.636	0.614
$\ln\left(\frac{\frac{NOK_t}{CHF_t}}{\frac{NOK_{t-1}}{CHF_{t-1}}}\right)$	0.313	1	0.861	0.550	0.442
$\ln\left(\frac{\frac{NOK_t}{EUR_t}}{\frac{NOK_{t-1}}{EUR_{t-1}}}\right)$	0.459	0.861	1	0.634	0.497
$\ln\left(\frac{\frac{NOK_t}{GBP_t}}{\frac{NOK_{t-1}}{GBP_{t-1}}}\right)$	0.636	0.550	0.634	1	0.566
$\ln\left(\frac{\frac{NOK_t}{JPY_t}}{\frac{NOK_{t-1}}{JPY_{t-1}}}\right)$	0.614	0.442	0.497	0.566	1

4.2 Data on Business Cycles

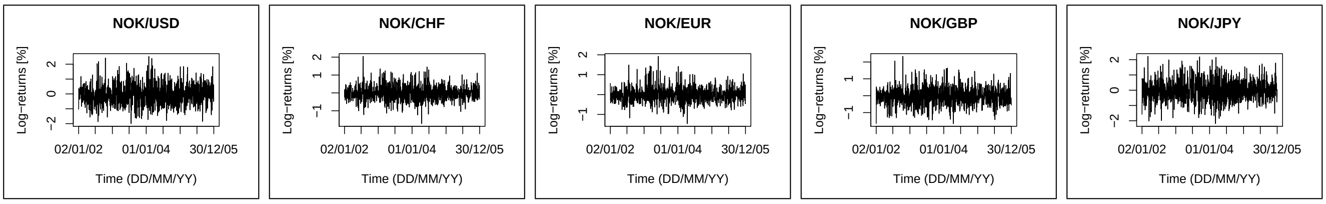
To extract the business cycles, data on GDP and the consumer price index (CPI) are collected from SSB, described in table 4.5, and the CPI is used to measure the GDP in fixed (2005) prices. This resultant GDP is then seasonal adjusted before it undergoes

[2] The log-return at time t on for instance USD measured in NOK is defined as $\ln\left(\frac{\frac{NOK_t}{USD_t}}{\frac{NOK_{t-1}}{USD_{t-1}}}\right)$.

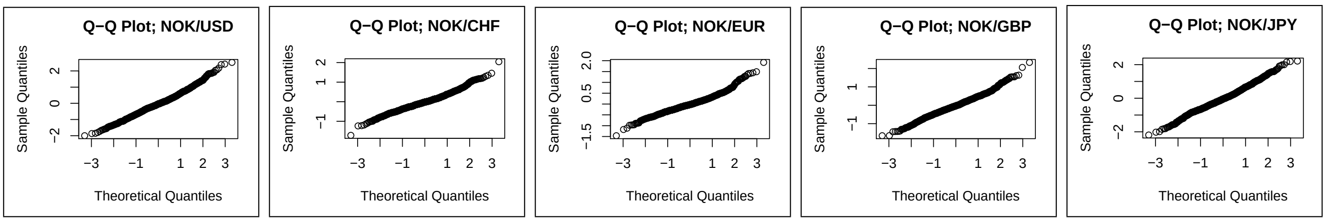
[3] An introduction to the exchange market and its attributes is given by e.g. Rødseth (2000).



(a) The exchange rates measured in NOK



(b) (c) (d) (e) (f)



(g) (h) (i) (j) (k)

Figure 4.2: Exchange rate data collected by DnBNOR. (a): Foreign currency against NOK. (b-f): Log-returns on each exchange rate. (g-k): Normal QQ-plots for the log-returns on each exchange rate.

a logarithmic transformation. Then the HP-filter constructed in the previous chapter is applied to this "logarithmic transformed seasonal adjusted GDP measured in fixed prices". Procedure 4.1 explains this in detail.

Table 4.5: Resolution and description of the data material collected from SSB.

Variable	Unit	Original resolution	Resolution used	Period (quarter/year)
GDP and its components	MNOK	quarterly	quarterly	Q1/1978-Q3/2005
CPI[a]	index	monthly	quarterly	M1/1979-M12/2005

[a]CPI for 1978 is (indirectly) available using the "price calculator" on SSB's webpages; see `http://ssb.no/emner/08/02/10/kpi/`.

Procedure 4.1: This is the procedure used to extract the business cycles in the Norwegian economy.

1. Collect data on GDP and CPI from SSB.[4]
2. Deflate the GDP figures with the CPI such as they are measured in 2005-prices.[5]
3. Adjust for seasonal patterns using a classical decomposition model; $Y_t = T_t + S_t + e_t$ where Y_t is the original GDP data, T_t is the trend, S_t is the seasonal component and I_t is the irregular component. The model prescribes that $E(I_t) = 0$, $S_{t+d} = S_t$ and $\sum_{i=1}^{d} S_i = 0$ where $d = 4$ is the period.[6] The implemented code can be found in appendix A.3.
4. Do a logarithmic transformation of the GDP data.
5. Apply the Hodrick-Prescott filter with $\lambda = 40000$ and extract the cycles.
6. Smooth the results from the HP-filter using a centred MA(5)-process with weights $(\frac{1}{8}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{8})$.[7]

Figure 4.3(a) illustrates the procedure 4.1 graphically. From here one can note that the GDP Mainland Norway, used by Johansen & Eika (2000), shows the same seasonal pattern as the regular GDP, but also that the petroleum sector was more volatile in recent years making the GDP series more volatile than GDP Mainland Norway as argued in chapter 3.2.1 above.

Figure 4.3(b) gives the results after the HP-filter described in section 3.2 has been applied and the results have been adjusted using a centred MA(5)-process as procedure 4.1 prescribes. This particular smoothing is inspired by Johansen & Eika (2000, page 27, footnote 3) as they say that

> "*Quarterly data in figures and tables are smoothed with a five-quarter moving, weighted average in order to eliminate short-term random effects and provide a clearer visual picture.*"

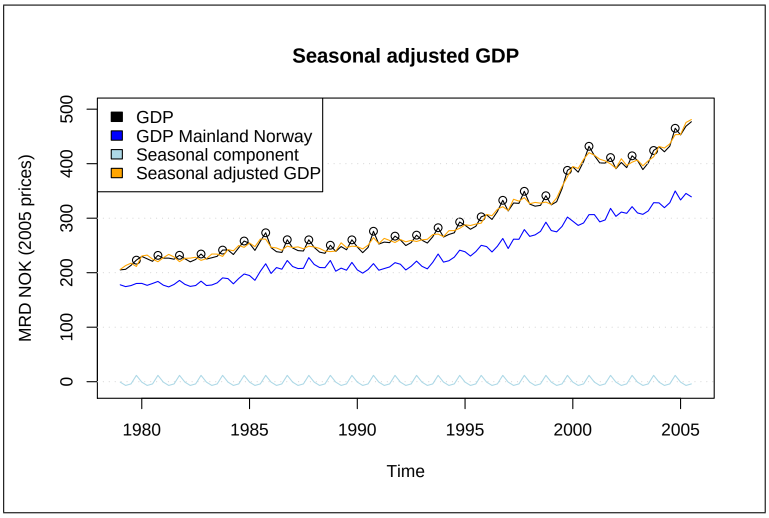
[4]This series is found in the "StatBank" of SSB (`http://statbank.ssb.no/`); then choose: "Subject 09: National economy and external trade Table 03519: Gross domestic product and value added, by industry".

[5]The CPI data is also from SSB; "Subject 08: Prices, price indices and economic indicators Table 03013: Consumer Price Index (1998=100)". The base year is corrected from 1998 to 2005 (fourth quarter).

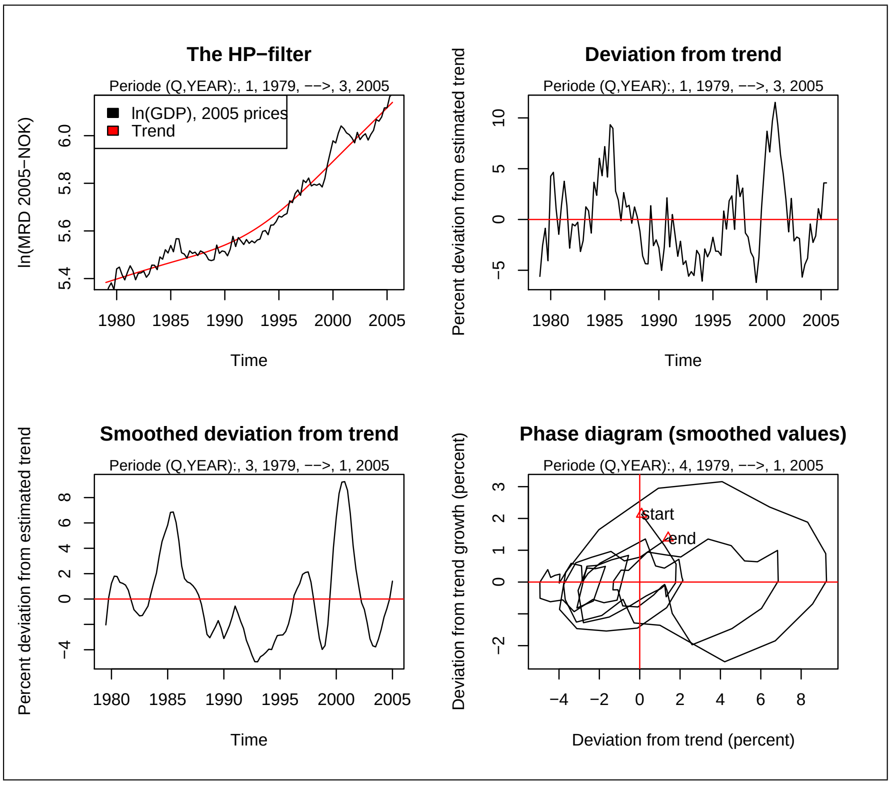
[6]This model is presented in Brockwell & Davis (2002, chapter 1).

[7]The weights were provided in an email from one of the authors, T. Eika.

The upper left and bottom right plots in figure 4.3(b) are the actual realizations of their stylized counterparts in figure 3.1. It is (part of) the values from the bottom right plot that are used as quantifications of the business cycles in the full model which is presented in the next chapter.



(a)



(b)

Figure 4.3: Graphical presentation of procedure 4.1. (a): The cirles in the GDP series points out the fourth quarter in every year. (b): Application of the HP-filter ($\lambda = 40000$) to the logarithmic transformed seasonal adjusted GDP.

Chapter 5

The Model

The previous three chapters have provided the theory and data that the model will be built upon. As written in the introduction, this model is an exemplification of a generic methodology which should be able to answer whether the risk appetite is dependent upon the macroeconomic environment.

As shown in the last chapter, the portfolio and exchange rate data had a daily time resolution while the business cycles were classified every quarter. This will motivate the use of Bayesian statistics with two different states; one macroeconomic and one microeconomic. These two states are presented before a short introduction to Bayesian theory is given. Then the covariates in the micro-state are chosen before the model is fully specified and implemented in `WinBUGS`.

5.1 Building the Model

As stated above, the model will be built using a Bayesian approach. One of the main reasons for using Bayesian statistics is that the data material has different time scales; the business cycles are measured every quarter, while the financial time series have a daily time resolution. This could, of course, be solved by aggregating the financial data to a quarterly basis, but this would also reduce the information in the model. Instead the business cycles, measured as the deviation from trend and the deviation from trend growth, are thought of as a *macroeconomic state* for the economy in the relevant quarters. This "macro-state" is specified in equation (5.1);

$$\mu_T = \beta_0 + \beta_1 dTrend_T + \beta_2 dTrendGwt_T, \tag{5.1}$$

where $dTrend_T$ is the deviation from trend and $dTrendGwt_T$ is the deviation from trend growth in quarter T. As one can see, this looks like a regression type of model, and that is indeed the intention: μ_T represents the macro-state and the coefficients, β_1 and β_2, will give the effects of $dTrend_T$ and $dTrendGwt_T$, respectively. T counts the quarters, $T \in (1, \ldots, N)$, where N is the number of quarters in the set of data.

The idea is now that for each μ_T there are several micro-states, labelled μ_t (note that the macro-state is specified with a capital T while the micro-state has a small t as its subscript). In all of these microstates one wants to quantify, if any, the impact from μ_T on the risk appetite held by the portfolio administrator.

Regardless of the macroeconomic state of the economy, it is also important to control the risk appetite, or realized risk, for other risk factors.[1] And this is what the micro-state should do; a generic specification is

$$\mu_t = \alpha_0 + \alpha_1 x_{1,t} + \alpha_2 x_{2,t} + \ldots + \alpha_q x_{q,t}. \tag{5.2}$$

This is, as one can see, also a regression-like relation where the α's will tell in which way and to what extent the risk factors (the covariates, x_i) has effect on the risk appetite. The covariates are specified in chapter 5.3 below.

The two states will together control for the exogenous factors one believe affect the risk appetite in the data; μ_T and μ_t are connected as shown in equation (5.3) below.

$$\mu_{S,t} = \gamma_1 \mu_T + \gamma_2 \mu_t \text{ for } t \in (K(T), \ldots, K(T+1)) \text{ where } T \in (1, \ldots, N). \tag{5.3}$$

where $(K(T+1) - K(T))$ is the number of time periods, or micro-states, in quarter T. $\mu_{S,t}$ can now be interpreted as *the expected level of risk* and this is what the model, described in detail in chapter 5.4 below, assumes.

Before specifying the model, a short introduction to Bayesian statistics will be given.

5.2 Bayesian Statistics; the Theory Behind

Bayesian statistics is different from the classical approach in two important aspects: It turns what is thought of as parameters in classical statistics into random variables and it assumes that there exists some prior knowledge of these variables, knowledge which is incorporated in what is called prior distributions, $\pi(\boldsymbol{\theta})$. The α's and β's in equations (5.2) and (5.1) above can serve as examples of such parameters that become random variables in the Bayesian paradigm.

A short intro to Bayesian statistics, with instructive examples, is given by Coles, Roberts & Jarner (2002) and Green (2001). For a thorough presentation of Bayesian theory, see Bernardo & Smith (1994).

5.2.1 Bayes Theorem

The cornerstone in Bayesian statistics is, naturally, Bayes Theorem. A representation of this can be found in e.g. Bernardo & Smith (1994). Here D is data, H is hypotheses, $P(H_j)$ is called *the prior probabilities*, $P(D|H_i)$ *the likelihoods* and $P(H_i|D)$ *the posteriori probabilities*.

Theorem 5.1 For any finite partition $\{H_j,\, j \in J\}$ of the certain event Ω and $D > \emptyset$,

$$P(H_i|D) = \frac{P(D|H_i)P(H_i)}{\sum_{j \in J} P(D|H_j)P(H_j)}. \tag{5.4}$$

Proof: Given the well known theorem of conditional probability,

$$P(H_i|D) = \frac{P(H_i \cap D)}{P(D)} = \frac{P(D|H_i)P(H_i)}{P(D)},$$

the result follows when applied to $D = \cup_j (D \cap H_j)$. □

[1] Another way to put this is that the moving variance of the profit and losses (see equation 5.9) becomes a measure for the risk appetite *after* it is controlled for the risk in the exchange rate market.

This result is directly transferable to the continuous case; the probabilities P are shifted with distributions π and $\{H_j, j \in J\}$ becomes (random) variables. So, let the random variables in the model be represented with $\boldsymbol{\theta}$ and call the data $\boldsymbol{y}$. Bayes theorem can now be written

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{\pi(\boldsymbol{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int \pi(\boldsymbol{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}} \tag{5.5}$$

where the integral in the denominator will be a constant. This means that

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) \propto \pi(\boldsymbol{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}), \tag{5.6a}$$

$$\text{i.e.;} \qquad \text{posterior} \propto \text{likelihood} \times \text{prior}. \tag{5.6b}$$

It is quite obvious that the distribution of interest is the (joint) posterior distribution as this is the one that gives estimates on the random variables of interest, *given* the collected data material. So, inference is based upon the joint posterior and one much used method is known as Markov chain Monte Carlo (MCMC).

5.2.2 MCMC

The idea behind MCMC is simple: Suppose the posterior distribution of interest is $\pi(\boldsymbol{\theta}|\boldsymbol{y})$ for $\boldsymbol{\theta} \in \Theta$, but that one cannot sample from $\pi(\boldsymbol{\theta}|\boldsymbol{y})$ directly. However, assume that a Markov chain with state space Θ and stationary distribution $\pi(\boldsymbol{\theta}|\boldsymbol{y})$ can be constructed and that this is easy to simulate from. If the chain is run for a long enough time, simulated values from this can be regarded as samples from the posterior distribution of interest and inference on these realizations will provide the requested information from $\pi(\boldsymbol{\theta}|\boldsymbol{y})$.

A central concept in the use of MCMC methods is the concept known as *conditional conjugacy*. This concept is based on the fact that each of the full conditional distributions in a Bayesian model often have “nice and simple” forms, even in cases when the posterior distribution does not. So, instead of sampling directly from the joint posterior, which may be an arbitrarily high dimensional distribution, MCMC uses the full conditionals to simulate a Markov chain whose stationary distribution is the posterior of interest. The concept of conditional conjugacy is illustrated in example 5.1.[2]

Example 5.1 Let $\pi(Y_i|\mu, \sigma^2) \sim \mathsf{N}(\mu, \sigma^2)$ for $i \in (1, \ldots, n)$ with prior distributions

$$\pi(\mu) \sim \mathsf{N}(\mu_0, \sigma_0^2) \tag{5.7a}$$

$$\pi(\sigma^{-2}) \sim \mathsf{Gamma}(\alpha_0, \beta_0) \tag{5.7b}$$

where μ and σ is considered to be a priori independent and μ_0, σ_0^2, α_0 and β_0 is known parameters. Further, let $\tau = \sigma^{-2}$ and $\tau_0 = \sigma_0^{-2}$. The posterior distribution is then

$$\pi(\mu, \tau|\boldsymbol{y}) \propto e^{-\frac{\tau_0}{2}(\mu-\mu_0)^2} \tau^{\alpha_0+\frac{n}{2}-1} e^{-\beta_0\tau} \prod_{i=1}^{n} e^{-\frac{\tau}{2}(y_i-\mu)^2}. \tag{5.8}$$

So even though the priors are tractable, common distributions, the posterior is not. But, the full conditional distributions on the other hand, will be quite easy to handle:

[2]This example is based on example 3.1 by Coles et al. (2002).

$$\begin{aligned}\pi(\mu|\tau,\boldsymbol{y}) &\propto e^{-\frac{\tau_0}{2}(\mu-\mu_0)^2}\prod_{i=1}^{n} e^{-\frac{\tau}{2}(y_i-\mu)^2} \\ &\sim \mathsf{N}\left(\frac{\tau\sum_{i=1}^{n} y_i+\mu_0\tau_0}{n\tau+\tau_0}, \frac{1}{n\tau+\tau_0}\right)\end{aligned}$$

and

$$\pi(\tau|\mu,\boldsymbol{y}) \sim \mathsf{Gamma}\left(\alpha_0+\frac{n}{2}, \beta_0+\frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2}\right).$$

This result, where the full conditional distributions have nice and simple forms whereas the joint posterior is intractable, is called conditional conjugacy.

The ideas from the example above are basis for one of the most common versions of MCMC; the Gibbs sampler.

5.2.3 The Gibbs Sampler

The Gibbs sampler samples from the (generic) multivariate distribution $\pi(\boldsymbol{\theta})$ where $\boldsymbol{\theta}$ is a vector with d components; $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d)$. The Gibbs sampler will now, successively and repeatedly, simulate from the full conditional distribution *of each component given the other components.* Algorithm 5.1 gives pseudocode for the Gibbs sampler.

Algorithm 5.1 The Gibbs Sampler

Initialize $\boldsymbol{\theta}^{(0)} = (\theta_1^{(0)}, \ldots, \theta_d^{(0)})$
for $i = 1, \ldots, m$ **do**
  Simulate $\theta_1^{(i)}$ from $\pi(\theta_1|\theta_2^{(i-1)}, \theta_3^{(i-1)}, \ldots, \theta_d^{(i-1)})$
  Simulate $\theta_2^{(i)}$ from $\pi(\theta_2|\theta_1^{(i)}, \theta_3^{(i-1)}, \ldots, \theta_d^{(i-1)})$
  $\vdots$
  Simulate $\theta_d^{(i)}$ from $\pi(\theta_d|\theta_1^{(i)}, \theta_2^{(i)}, \ldots, \theta_{d-1}^{(i)})$
end for
return $\boldsymbol{\theta} = \boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \ldots, \boldsymbol{\theta}^{(m)}$

The iteration procedure in algorithm 5.1 does not converge to the stationary distribution immediately, so one has to throw away some, say k, of the first samples (this is referred to as the burn-in period). Then it can be shown that, as $m \to \infty$, $\boldsymbol{\theta}^{(k+1)}, \ldots, \boldsymbol{\theta}^{(m)}$ can be regarded as realizations from $\pi(\boldsymbol{\theta})$.[3]

5.3 Choosing Variables in the Model

Recall that one part of the model is already specified; the macro-state and its covariates were given in equation (5.1) above. This leaves the micro-state, equation (5.2). After a short discussion on different risk concepts, including risk appetite, the covariates in the micro-state are presented.

[3]See e.g. Geman & Geman (1984, chapter XII).

5.3.1 Risk Appetite, Risk Aversion and Risk Premium

The key (dependent) variable in the model is the *risk appetite*. In the literature this is normally defined as *the willingness of investors to bear risk*, but the concept of risk appetite is hard to disentangle from the concepts of *risk aversion* and *risk premium* and the literature is not consistent everywhere. This thesis will give most attention to the definition used by Gai & Vause (2006) who allege that the risk appetite depends on both the *degree* to which investors dislike uncertainty surrounding their future consumption possibilities implied by their assets holdings, and the *level* of that uncertainty. The degree of uncertainty is determined by the investor's risk aversion and the level of uncertainty is derived from the macroeconomic state or environment.

Further, Gai & Vause (2006) claims that the risk aversion, the curvature of individuals' utility functions, is a parameter that is unlikely to change markedly, or frequently, over time. The risk appetite on the other hand, *is* likely to shift periodically as investors respond to episodes of financial distress and macroeconomic uncertainty.[4] This is because investors will require higher excess expected return to hold each unit of risk in adverse circumstances, implying that risk appetite will be low. Together with the quantity of risk in a particular asset, risk appetite will determine the risk premium, i.e. the expected return required to compensate investors for holding that asset. The relationship between these concepts is shown in figure 5.1.[5] Here the dotted box represents the risk appetite measure used in this thesis which must be controlled for the market risk ("Riskiness of asset" in Gai & Vause's (2006) terminology as they are concerned with a single asset). This is done by the micro-state, equation (5.2), above.

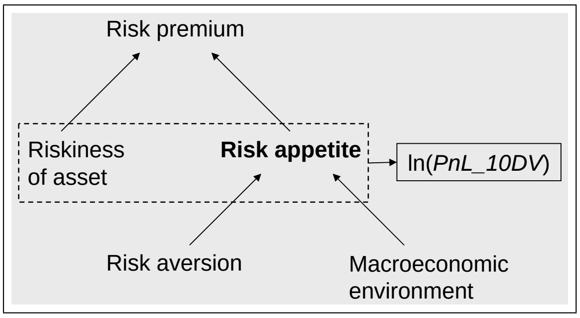


Figure 5.1: Relationship between risk concepts. $\ln(PnL\_10DV_t)$ is explained in equation (5.9).

To measure the risk appetite, Gai & Vause (2006) use the thoughts and notation from Cochrane's (2001) treatment of asset pricing theory and show that a measure of risk appetite may be derived by computing the variation in the ratio of risk-neutral to subjective probabilities used by investors in evaluating the expected payoff of an asset.

This thesis will use a simplified approach for measuring the risk appetite; a case study will be carried out with the data described in chapter 4.1.1 and *the realized risk* in the portfolio will be used as a proxy for the risk appetite and measured as the moving variance in the portfolio's profit and losses (PnL). Intuitively this seems reasonable; if the portfolio's administrator takes on greater risk, this will mean that the profits and losses must fluctuate more heavily, hence, the variance must become larger.[6]

[4] See for instance Varian (1999, chapter 12.5) for an introductory approach on risk aversion.
[5] This is based on Gai & Vause (2006, figure 1).
[6] Note that the variance in profits (PnL) must equal the variance in losses (L) as profit$= -$losses.

The length, labelled l, of this moving variance band can be varied. The preferred specification of the model chooses $l = 10$, which corresponds to two working weeks. The risk appetite, labelled PnL_10DV_t, is given by the equation below.[7]

$$PnL\_l\mathrm{DV}_t = \frac{1}{l-1} \sum_{i=t-(l-1)}^{t} (PnL_i - \overline{PnL_t})^2 \tag{5.9a}$$

$$PnL\_10DV_t = \frac{1}{9} \sum_{i=t-9}^{t} (PnL_i - \overline{PnL_t})^2, \tag{5.9b}$$

where $\overline{PnL_t} = \frac{1}{l} \sum_{j=t-(l-1)}^{t} PnL_j$. As in figure 4.1(a), PnL_10DV_t is measured in MNOK and because the model assumes that its data, i.e. the risk appetite, is normally distributed, PnL_10DV_t is log-transformed before it goes into the model; i.e. the risk appetite is measured as $\ln(PnL\_10DV_t)$.

Figure 5.2(a) shows the risk appetite measured with the log-transformed values from equation (5.9) against time, while figures 5.2(b) and (c) show the QQ-plot and histogram with the same values together with the normal distribution, respectively. From the latter two figures one can see that the normality assumption is a pretty rough approximation as the tails in the empirical distribution are heavier than in the normal distribution, but because this assumption makes the model more stable and easier to solve, this remains the preferred specification.[8]

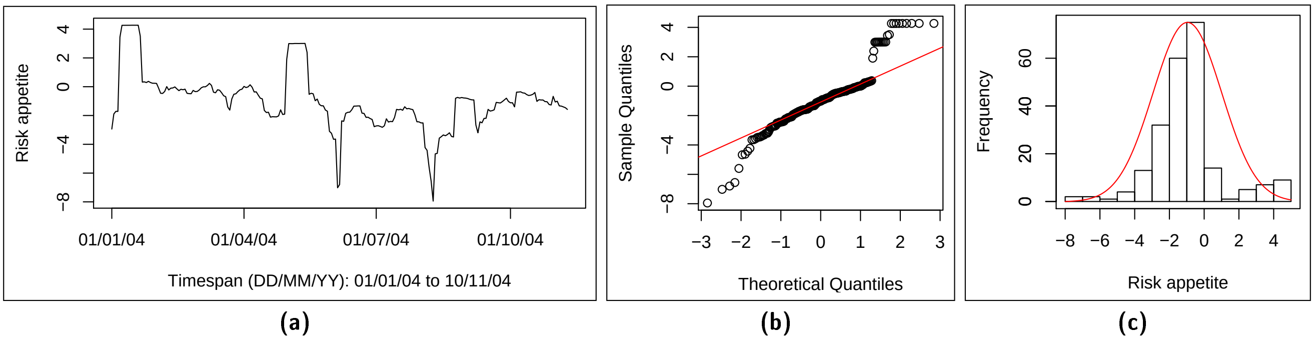


Figure 5.2: Risk appetite measured as the log-transformed values from equation (5.9). (a): Plot against time. (b): QQ-plot. (c): Histogram together with the normal distribution (red line).

5.3.2 The Covariates in the Micro-State

Equation (5.2) above gave a generic specification of the micro-state and now it is time to choose the covariates in this. First and foremost; to answer whether the risk appetite is dependent on the business cycles, it must be controlled for the underlying, *actual* risk in

[7] One should note that it is after the moving variance has been controlled for the market risk is becomes the risk appetite in Gai & Vause's (2006) terminology; this is what the dotted box in figure 5.1 indicates.

[8] One can also note that the four outliers in the profit and losses, pointed out in figure 4.1(a), are the reason behind the two peaks in figure 5.2(a) and therefore also responsible for much of the discrepancy from the normality assumption.

the exchange rate market in the relevant time period. It is very possible that the profits and losses from the portfolio become more volatile due to a higher risk in the underlying assets and not because of a change in the business cycle (or higher risk appetite). The procedure for constructing this "benchmark risk measure" is given below.

Benchmarking the Underlying Risk

To control for the underlying risk in the exchange rate market an Expected Shortfall "benchmark" measure is created, based on the exchange rate data described in chapter 4.1.2.

Usually, when one calculates a VaR or a ES measure, one does so because one needs to estimate the risk exposure for a time period in the future. For instance, if there where 100 trading days with data on the risk factor changes, $\boldsymbol{X}_{t-99}, \boldsymbol{X}_{t-98}, \ldots, \boldsymbol{X}_t$, all these would be used to make a risk estimate for day $t+1$. Here the intent is to control for the *actual or realized* risk in the market at time t and this imply that data on risk factor changes that *includes* day t must be used in the benchmark risk measure. I.e. when the risk appetite is controlled for the market risk at day t, the market risk measure is based on risk factor changes that includes day t.

With use of the terminology from chapter 2.2.1 the log-transformed exchange rates, e.g. $\ln\left(\frac{\text{NOK}}{\text{USD}}\right)$, are chosen as the risk factors $\boldsymbol{Z}$. This means that the change in risk factors, $\boldsymbol{X}_t = \boldsymbol{Z}_t - \boldsymbol{Z}_{t-1}$ is given by the matrix

$$\boldsymbol{X} = \begin{bmatrix} \ln\left(\frac{\frac{\text{NOK}_{t=2}}{\text{USD}_{t=2}}}{\frac{\text{NOK}_{t=1}}{\text{USD}_{t=1}}}\right) & \ln\left(\frac{\frac{\text{NOK}_{t=2}}{\text{CHF}_{t=2}}}{\frac{\text{NOK}_{t=1}}{\text{CHF}_{t=1}}}\right) & \ln\left(\frac{\frac{\text{NOK}_{t=2}}{\text{EUR}_{t=2}}}{\frac{\text{NOK}_{t=1}}{\text{EUR}_{t=1}}}\right) & \ln\left(\frac{\frac{\text{NOK}_{t=2}}{\text{GBP}_{t=2}}}{\frac{\text{NOK}_{t=1}}{\text{GBP}_{t=1}}}\right) & \ln\left(\frac{\frac{\text{NOK}_{t=2}}{\text{JPY}_{t=2}}}{\frac{\text{NOK}_{t=1}}{\text{JPY}_{t=1}}}\right) \\ \ln\left(\frac{\frac{\text{NOK}_{t=3}}{\text{USD}_{t=3}}}{\frac{\text{NOK}_{t=2}}{\text{USD}_{t=2}}}\right) & \ln\left(\frac{\frac{\text{NOK}_{t=3}}{\text{CHF}_{t=3}}}{\frac{\text{NOK}_{t=2}}{\text{CHF}_{t=2}}}\right) & \ln\left(\frac{\frac{\text{NOK}_{t=3}}{\text{EUR}_{t=3}}}{\frac{\text{NOK}_{t=2}}{\text{EUR}_{t=2}}}\right) & \ln\left(\frac{\frac{\text{NOK}_{t=3}}{\text{GBP}_{t=3}}}{\frac{\text{NOK}_{t=2}}{\text{GBP}_{t=2}}}\right) & \ln\left(\frac{\frac{\text{NOK}_{t=3}}{\text{JPY}_{t=3}}}{\frac{\text{NOK}_{t=2}}{\text{JPY}_{t=2}}}\right) \\ \ln\left(\frac{\frac{\text{NOK}_{t=4}}{\text{USD}_{t=4}}}{\frac{\text{NOK}_{t=3}}{\text{USD}_{t=3}}}\right) & \ln\left(\frac{\frac{\text{NOK}_{t=4}}{\text{CHF}_{t=4}}}{\frac{\text{NOK}_{t=3}}{\text{CHF}_{t=3}}}\right) & \ln\left(\frac{\frac{\text{NOK}_{t=4}}{\text{EUR}_{t=4}}}{\frac{\text{NOK}_{t=3}}{\text{EUR}_{t=3}}}\right) & \ln\left(\frac{\frac{\text{NOK}_{t=4}}{\text{GBP}_{t=4}}}{\frac{\text{NOK}_{t=3}}{\text{GBP}_{t=3}}}\right) & \ln\left(\frac{\frac{\text{NOK}_{t=4}}{\text{JPY}_{t=4}}}{\frac{\text{NOK}_{t=3}}{\text{JPY}_{t=3}}}\right) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \ln\left(\frac{\frac{\text{NOK}_{t=T-1}}{\text{USD}_{t=T-1}}}{\frac{\text{NOK}_{t=T-2}}{\text{USD}_{t=T-2}}}\right) & \ln\left(\frac{\frac{\text{NOK}_{t=T-1}}{\text{CHF}_{t=T-1}}}{\frac{\text{NOK}_{t=T-2}}{\text{CHF}_{t=T-2}}}\right) & \ln\left(\frac{\frac{\text{NOK}_{t=T-1}}{\text{EUR}_{t=T-1}}}{\frac{\text{NOK}_{t=T-2}}{\text{EUR}_{t=T-2}}}\right) & \ln\left(\frac{\frac{\text{NOK}_{t=T-1}}{\text{GBP}_{t=T-1}}}{\frac{\text{NOK}_{t=T-2}}{\text{GBP}_{t=T-2}}}\right) & \ln\left(\frac{\frac{\text{NOK}_{t=T-1}}{\text{JPY}_{t=T-1}}}{\frac{\text{NOK}_{t=T-2}}{\text{JPY}_{t=T-2}}}\right) \\ \ln\left(\frac{\frac{\text{NOK}_{t=T}}{\text{USD}_{t=T}}}{\frac{\text{NOK}_{t=T-1}}{\text{USD}_{t=T-1}}}\right) & \ln\left(\frac{\frac{\text{NOK}_{t=T}}{\text{CHF}_{t=T}}}{\frac{\text{NOK}_{t=T-1}}{\text{CHF}_{t=T-1}}}\right) & \ln\left(\frac{\frac{\text{NOK}_{t=T}}{\text{EUR}_{t=T}}}{\frac{\text{NOK}_{t=T-1}}{\text{EUR}_{t=T-1}}}\right) & \ln\left(\frac{\frac{\text{NOK}_{t=T}}{\text{GBP}_{t=T}}}{\frac{\text{NOK}_{t=T-1}}{\text{GBP}_{t=T-1}}}\right) & \ln\left(\frac{\frac{\text{NOK}_{t=T}}{\text{JPY}_{t=T}}}{\frac{\text{NOK}_{t=T-1}}{\text{JPY}_{t=T-1}}}\right) \end{bmatrix}$$

which has $T - 1 = n$ number of rows where $T = 1044$ is the number of rows in the original exchange rate series which stretch from 01/01/02 to 30/12/05.

In the calculation of the benchmark risk measure some assumptions must be made. Because data on the weights of each exchange rate in the portfolio is unavailable, this thesis choose to assume that the weights are equal in all assets through the whole period; since there are five possible placements ($d = 5$) the weight of each must be 0.2. Even though this is a very crude assumption, it still is reasonable as the intention is to say something about the risk in the exchange rate *market*, not in just this one portfolio.

Further, the thesis will use the variance-covariance method from chapter 2.3.1 to calculate the benchmark risk estimate. This implies an assumption that the linearization in equation (2.11) holds and that the changes in the risk factors, i.e. the log-returns on the currencies, follow a multivariate normal distribution such that equation (2.9) applies. The fact that the value of the different currencies does not depend on calendar time together with a short time horizon $\Delta = 1$ day, makes the linearization of the loss distribution a decent approximation. The multivariate normality assumption also seems reasonable at first sight as the QQ-plots in figures 4.2(g-k) are pretty linear, but normal marginal distributions do not necessarily produce a multinormal distribution. So this assumption needs to be tested.[9]

It can be shown that if $\boldsymbol{X} \sim \mathrm{N}_d(\boldsymbol{\mu}, \Sigma)$ with Σ positive definite then[10]

$$(\boldsymbol{X} - \boldsymbol{\mu})' \, \Sigma^{-1} \, (\boldsymbol{X} - \boldsymbol{\mu}) \sim \chi^2_d \tag{5.10}$$

where χ^2_d denotes the chi-square distribution with d degrees of freedom. To see whether the risk factor changes are multivariate normally distributed, one can construct data

$$D_i^2 = \left(\boldsymbol{X}_i - \bar{\boldsymbol{X}}\right)' S^{-1} \left(\boldsymbol{X}_i - \bar{\boldsymbol{X}}\right), \tag{5.11}$$

where $\bar{\boldsymbol{X}} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{X}_i$ and $S = \frac{1}{n}\sum_{i=1}^{n} \left(\boldsymbol{X}_i - \bar{\boldsymbol{X}}\right)\left(\boldsymbol{X}_i - \bar{\boldsymbol{X}}\right)'$, which should behave like an iid sample from a chi-square distribution with d degrees of freedom.[11] Figure 5.3 presents a QQ-plot of the ordered D_i^2-data against the $\chi^2_{\text{df}=5}$-quantile. From here it is obvious that the empirical distribution has heavier tails than its theoretical counterpart which is a common result in financial time series with a high time resolution. This means that the joint normality assumption is indeed a crude approximation. However, as this assumption simplifies the further analysis, it still will be used, but one should note that this imply that large observations, positive or negative, of $\ln(PnL\_10DV_t)$ is less likely to be explained by the benchmark risk measure. This in turn will possibly overestimate the impact from the other microeconomic covariates and also the macro-state.

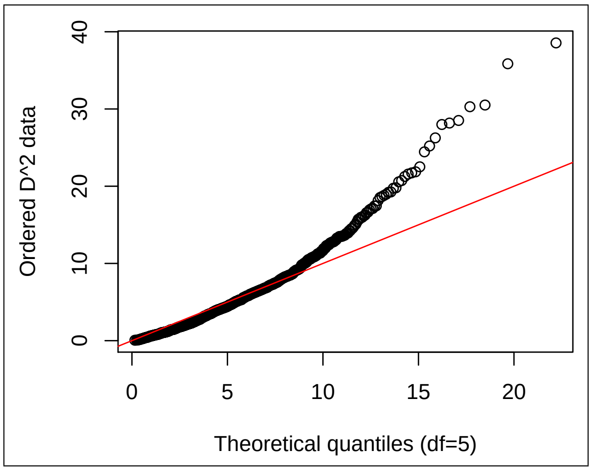


Figure 5.3: QQ-plot of the ordered D_i^2-data from equation (5.11) against the χ^2_5-quantile. The plot is constructed with three steps; first sort the data from (5.11); construct probability levels using $p_i = \left(i - \frac{1}{2}\right)/n$ and then use these probability levels to calculate quantiles from the chi-squared distribution with 5 degrees of freedom.

Figure 5.4 shows a collection of different risk measures calculated with equations (2.7) and (2.9) with different α-levels. Here it becomes apparent that the more data (the size of

[9] See McNeil et al. (2005, chapter 3.1.3) for a short introduction to the multivariate normal distribution or Johnson & Wichern (2002, chapter 4) for a more extensive presentation on the subject.

[10] See for instance Johnson & Wichern (2002, Result 4.7) or McNeil et al. (2005, equation (3.14)).

[11] This is an approximation which depends on the size of data, n; if n is large this is a good approximation.

n) the risk estimate is based on, the less variable it becomes. It also shows, as expected, that $ES_\alpha > VaR_\alpha$ for $\alpha < 1$. The code implemented to extract these estimates can be found in appendix A.6.

Activity as a Risk Measure

In chapter 5.3.3 below it is shown that the variation in the benchmark risk measure does not count for very much of the variation in the risk appetite. This is partly due to the fact that the benchmark risk measure is only calculated one time each day (i.e. at the end of day) which may not be sufficient because the exchange rate market has large intra-day volatility. The idea behind the other risk factor chosen in equation (5.2), called the *activity* (A), is that the portfolio administrator will execute more and bigger transactions on days where this intra-day volatility is high, meaning that *a higher activity is an indication on higher (intra-day) risk in the exchange rate market.*

This motivates the activity as another risk factor in equation (5.2). It is measured as the product of the number (N) and volume (V) of transactions and shown in figure 5.5.[12]

Now equation (5.2) can be specified:

$$\mu_t = \alpha_0 + \alpha_1 A_t + \alpha_2 ES^{n=10}_{\alpha=0.99,t}. \tag{5.12}$$

where A_t is the activity as explained above and $ES^{n=10}_{\alpha=0.99,t}$ is the estimated shortfall based on $\boldsymbol{X}_{t-9}, \ldots, \boldsymbol{X}_t$. The reason for choosing $n = 10$ in the ES estimates is simply that this corresponds with the moving variance band in the risk appetite measure, PnL_10DV_t, from equation (5.9).

5.3.3 A Preliminary Analysis of the Micro-state

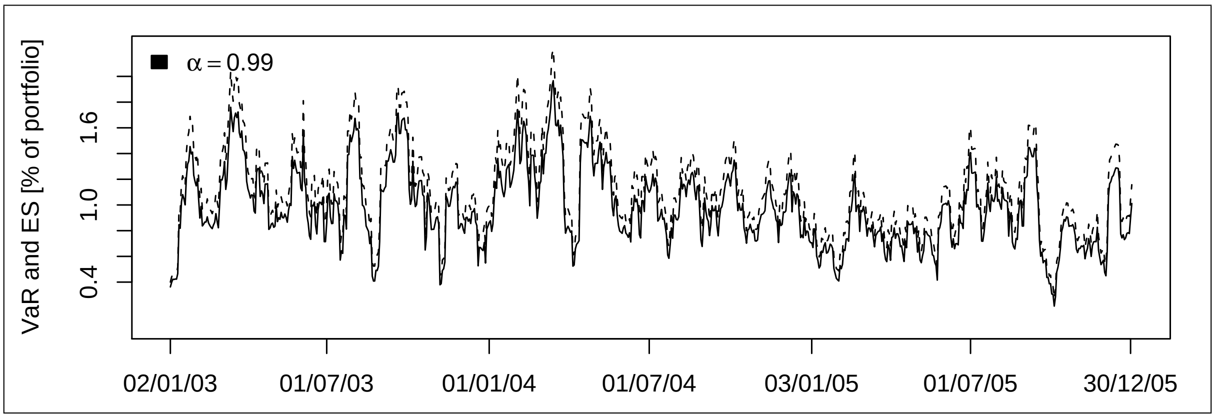
To see what can be expect of dependence between the variables in the micro-state, the dependent variable, $\ln(PnL\_10DV_t)$, is plotted against the two explanatory variables in equation (5.12), shown in figure 5.6. From this figure there is apparently no strong dependency between the variables.

To look further into the relationship between the variables in the micro-state a standard regression analysis is performed:

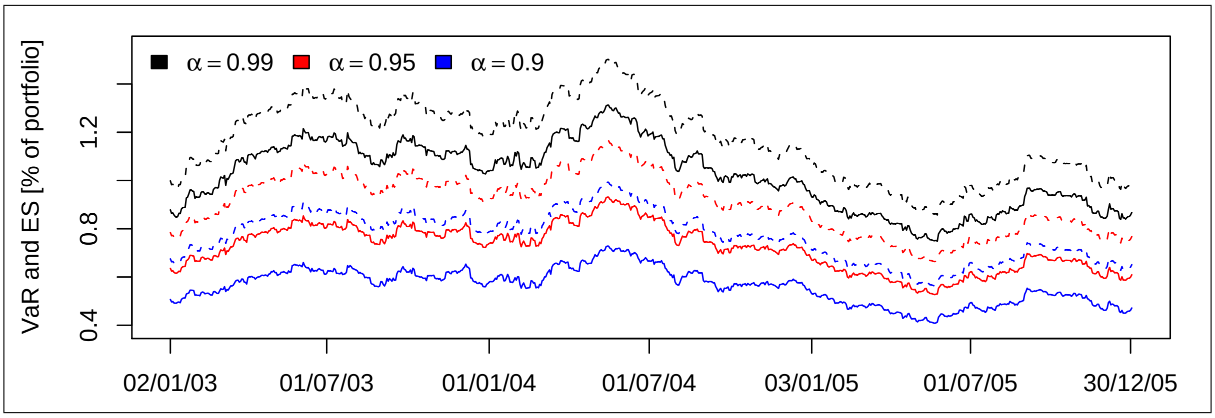
$$\ln(PnL\_10DV_t) = \alpha_0 + \alpha_1 A_t + \alpha_2 ES^{n=10}_{\alpha=0.99,t} \tag{5.13}$$

for $t \in (01/01/04, \ldots, 10/11/04)$ (225 observations) which resulted in the estimates shown in table 5.1. The table shows, as one would expect, that the coefficient belonging to the ES estimate is significantly different from zero and that risk appetite is positively dependent upon the ES estimate. The activity A has no demonstrable effect at a five percent significance level ($\alpha = 0.05$), but should be considered significant at a ten percent significance level as its p-value is 0.0684. The activity is also positively correlated with the risk appetite ($\alpha_1 > 0$). But, these estimates are not very robust; when the

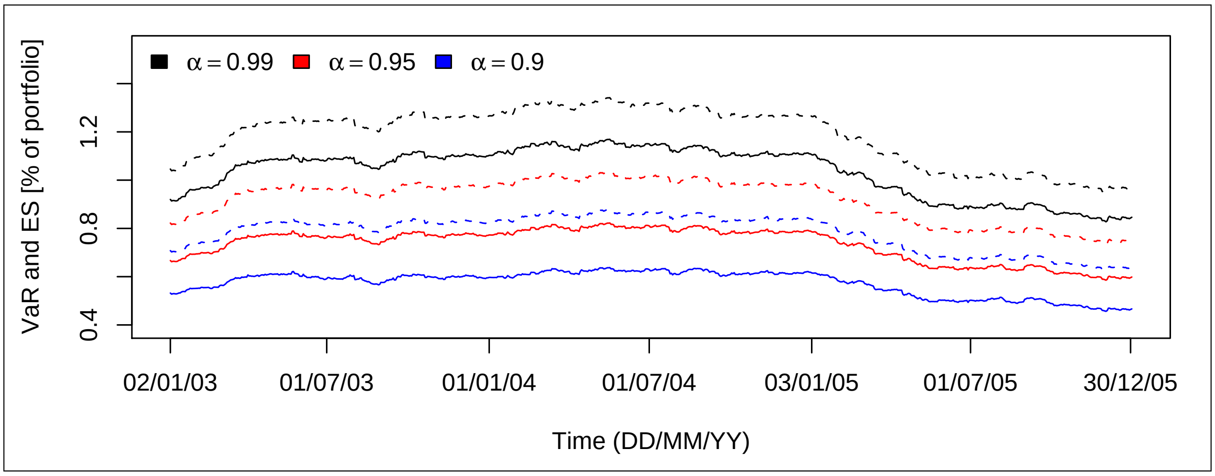
[12] The reason for not just include both the number and volume of transactions in the micro-state, is the problem with multicollinearity; the correlation matrix in table 4.2 showed that the two had a correlation coefficient above 90 % which means that the information carried in one is also more or less present in the other. This would not be a very big problem if the only intention was to predict the risk appetite, but here we are also interested in the α's in equation (5.2) and their standard deviation will be affected (larger) in case of multicollinearity. For an introduction to econometrics and the problem of multicollinearity see e.g. Wooldridge (2003, chapter 3).



(a)



(b)



(c)

Figure 5.4: Different benchmark risk measures. The different colours correspond to different α-levels and the VaR measures are shown with full lines while the ES measures are dotted. (a): Uses $\boldsymbol{X}_{t-9}, \ldots, \boldsymbol{X}_t$ ($n = 10$) to construct VaR_α and ES_α for time t. (b): Uses $\boldsymbol{X}_{t-99}, \ldots, \boldsymbol{X}_t$ ($n = 100$) to construct VaR_α and ES_α for time t. (c): Uses $\boldsymbol{X}_{t-261}, \ldots, \boldsymbol{X}_t$ ($n = 262$; one year) to construct VaR_α and ES_α for time t.

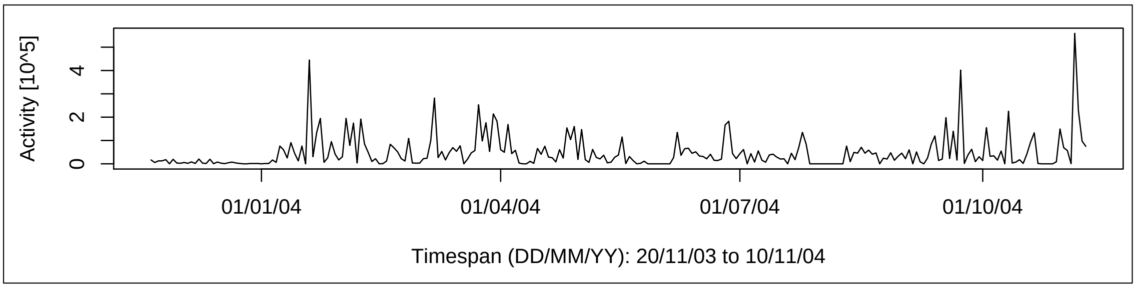


Figure 5.5: The activity (A) in the portfolio measured as the product of the numbers (N) and volumes (V) of transactions, $N \cdot V = A$.

same regression is done with half of the values, both the coefficient estimates and their significance are substantially altered. And with an adjusted R^2 less than eight percent, this analysis does not provide anything conclusive.

Table 5.1: Results from the regression analysis in equation (5.13), using the `lm`-function in `R`. There are 225 observations in the time period from 01/01/04 to 10/11/04. Adjusted $R^2 = 0.07757$.

	Estimate	Std. Error	t-value	p-value
α_0	-3.135	0.5142	-6.097	4.74E-09
α_1	3.154E-06	1.723E-06	1.831	0.0684
α_2	1.586	0.39361	4.028	7.72E-05

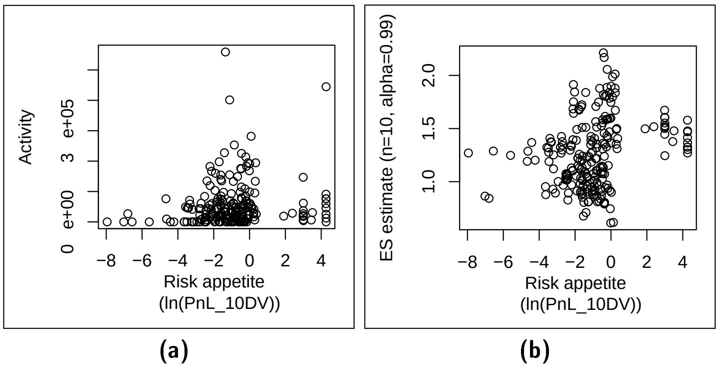


Figure 5.6: Risk appetite plotted against the (a): activity A and (b): $ES^{n=10}_{\alpha=0.99}$.

5.4 The Full Model

Now that both the micro- and macro-state is specified, the full model can be constructed. As shown in equation (5.6), the likelihood and prior distributions need to be specified to find the posterior distribution.

The logarithmic transformed risk appetite is, as written in chapter 5.3.1 above, assumed to be normal distributed given the other variables in the model (i.e. the α's, β's, γ's and τ). This gives the likelihood:

$$\pi(\ln(PnL\_10DV_t)\,|\alpha_0,\alpha_1,\alpha_2,\beta_0,\beta_1,\beta_2,\gamma_1,\gamma_2,\tau) \sim \mathrm{N}\left(\mu_{S,t},\tau^{-1}\right) \tag{5.14}$$

where the variability is represented by the precision τ instead of the variance σ^2 and where[13]

$$\mu_{S,t} = \gamma_1(\beta_0 + \beta_1 dTrend_T + \beta_2 dTrendGwt_T) + \gamma_2(\alpha_0 + \alpha_1 A_t + \alpha_2 ES^{n=10}_{\alpha=0.99,t}) \tag{5.15}$$

which is interpreted as *the expected risk appetite*. The covariates $dTrend$, $dTrendGwt$, A and $ES^{n=10}_{\alpha=0.99,t}$ are deterministic and the (uninformative) prior distributions in the model are given as

$$\pi(\alpha_i) = \mathrm{N}(0,(10^{-6})^{-1}) \text{ for } i \in (0,1,2) \tag{5.16a}$$
$$\pi(\beta_j) = \mathrm{N}(0,(10^{-6})^{-1}) \text{ for } j \in (0,1,2) \tag{5.16b}$$
$$\pi(\gamma_k) = \mathrm{N}(0,(10^{-6})^{-1}) \text{ for } k \in (1,2) \tag{5.16c}$$
$$\pi(\tau) = \mathrm{Gamma}(a = 10^{-3}, \frac{1}{s} = 10^{-3}). \tag{5.16d}$$

where a and s are the shape and scale parameters such that $E(\tau) = 1$. The conditional distributions for $\alpha_i, \beta_j, \gamma_k$ and τ for $i \in (0,1,2)$, $j \in (0,1,2)$ and $k \in (1,2)$ are shown in table 5.2. In this table the notation $\pi(\alpha_0|-\alpha_0)$ means "the full conditional distribution of α_0, given all the other variables in the model"; i.e. $\pi(\alpha_0|-\alpha_0) = \pi(\alpha_0|\alpha_1,\alpha_2,\beta_0,\beta_1,\beta_2,\gamma_1,\gamma_2,\tau)$, $\pi(\alpha_1|-\alpha_1) = \pi(\alpha_1|\alpha_0,\alpha_2,\beta_0,\beta_1,\beta_2,\gamma_1,\gamma_2,\tau)$ and so on. Detailed calculations can be found in appendix A.7.

[13] One have that $\sigma^2 = 1/\tau$.

Table 5.2: Conditional distributions for the variables in the model. Note that $x_{1,t} = A_t$ and $x_{2,t} = ES^{n=10}_{\alpha=0.99,t}$ in the preferred spesification.

$$\pi(\alpha_0|-\alpha_0) \sim \mathrm{N}\left(\frac{\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} \gamma_2(\ln(PnL\_10DV)_t - \gamma_1\beta_0 - \gamma_1\beta_1 dTrend_T - \gamma_1\beta_2 dTrendGwt_T - \gamma_2\alpha_1 x_{1,t} - \gamma_2\alpha_2 x_{2,t})}{\gamma_2^2 K_{N+1}}, \frac{1}{\tau\gamma_2^2 K_{N+1}}\right)$$

$$\pi(\alpha_1|-\alpha_1) \sim \mathrm{N}\left(\frac{\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} \gamma_2 x_{1,t}(\ln(PnL\_10DV)_t - \gamma_1\beta_0 - \gamma_1\beta_1 dTrend_T - \gamma_1\beta_2 dTrendGwt_T - \gamma_2\alpha_0 - \gamma_2\alpha_2 x_{2,t})}{\gamma_2^2 \sum_{t=K_1+1}^{K_{N+1}} x_{1,t}^2}, \frac{1}{\tau\gamma_2^2 \sum_{t=K_1+1}^{K_{N+1}} x_{1,t}^2}\right)$$

$$\pi(\alpha_2|-\alpha_2) \sim \mathrm{N}\left(\frac{\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} \gamma_2 x_{2,t}(\ln(PnL\_10DV)_t - \gamma_1\beta_0 - \gamma_1\beta_1 dTrend_T - \gamma_1\beta_2 dTrendGwt_T - \gamma_2\alpha_0 - \gamma_2\alpha_1 x_{1,t}}{\gamma_2^2 \sum_{t=K_1+1}^{K_{N+1}} x_{2,t}^2}, \frac{1}{\tau\gamma_2^2 \sum_{t=K_1+1}^{K_{N+1}} x_{2,t}^2}\right)$$

$$\pi(\beta_0|-\beta_0) \sim \mathrm{N}\left(\frac{\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} \gamma_1(\ln(PnL\_10DV)_t - \gamma_1\beta_1 dTrend_T - \gamma_1\beta_2 dTrendGwt_T - \gamma_2\alpha_0 - \gamma_2\alpha_1 x_{1,t} - \gamma_2\alpha_2 x_{2,t})}{\gamma_1^2 K_{N+1}}, \frac{1}{\tau\gamma_1^2 K_{N+1}}\right)$$

$$\pi(\beta_1|-\beta_1) \sim \mathrm{N}\left(\frac{\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} \gamma_1 dTrend_T(\ln(PnL\_10DV)_t - \gamma_1\beta_0 - \gamma_1\beta_2 dTrendGwt_T - \gamma_2\alpha_0 - \gamma_2\alpha_1 x_{1,t} - \gamma_2\alpha_2 x_{2,t})}{\gamma_1^2 \sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} dTrend_T^2}, \frac{1}{\tau\gamma_1^2 \sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} dTrend_T^2}\right)$$

$$\pi(\beta_2|-\beta_2) \sim \mathrm{N}\left(\frac{\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} \gamma_1 dTrendGwt_T(\ln(PnL\_10DV)_t - \gamma_1\beta_0 - \gamma_1\beta_1 dTrend_T - \gamma_2\alpha_0 - \gamma_2\alpha_1 x_{1,t} - \gamma_2\alpha_2 x_{2,t})}{\gamma_1^2 \sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} dTrendGwt_T^2}, \frac{1}{\tau\gamma_1^2 \sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} dTrendGwt_T^2}\right)$$

$$\pi(\gamma_1|-\gamma_1) \sim \mathrm{N}\left(\frac{\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} \mu_T(\ln(PnL\_10DV)_t - \gamma_2\mu_t)}{\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} \mu_T^2}, \frac{1}{\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} \mu_T^2}\right)$$

$$\pi(\gamma_2|-\gamma_2) \sim \mathrm{N}\left(\frac{\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} \mu_t(\ln(PnL\_10DV)_t - \gamma_1\mu_T)}{\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} \mu_t^2}, \frac{1}{\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} \mu_t^2}\right)$$

$$\pi(\tau|-\tau) \sim \mathrm{Gamma}\left(a^* = a + \frac{K_{N+1}}{2}, \frac{1}{s^*} = \frac{1}{s} + \frac{\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}\left(\ln(PnL\_10DV)_t - \mu_S\right)^2}{2}\right)^a$$

[a] This specific result is also a consequence of conditional conjugacy; the gamma distribution is a conjugate distribution of the normal distribution which means that a gamma prior and a normal likelihood leads to a gamma posterior.

5.5 The Model Implemented in `WinBUGS`

This thesis uses a program package called **`WinBUGS`** to implement the model described above. This is an open source program which uses Gibbs sampling (recall algorithm 5.1) to make statistical inference about the random variables of interest.[14]

Using **`WinBUGS`**, there is in fact no need to calculate the posterior distributions shown in table 5.2; they are included here for completeness. What **`WinBUGS`** *do* need is the likelihood and the prior distributions, i.e. equations (5.14) and (5.16), respectively. The implementation is given in figure 5.7 and illustrated graphically in figure 5.8.

```
model{
   beta0 ~ dnorm(0, 1.0E-6)
   beta1 ~ dnorm(0, 1.0E-6)
   beta2 ~ dnorm(0, 1.0E-6)
   alpha2 ~ dnorm(0, 1.0E-6)
   alpha1 ~ dnorm(0, 1.0E-6)
   alpha0 ~ dnorm(0, 1.0E-6)
   tau ~ dgamma(1.0E-3, 1.0E-3)
   gamma1 ~ dnorm(0, 1.0E-6)
   gamma2 ~ dnorm(0, 1.0E-6)

   for(T in 1 : N) {
      mu_T[T] <- beta0 + beta1 * dTrend[T] + beta2 * dTrendGwt[T]
      for(t in K[T] + 1 : K[T + 1]) {
         PnL_10DV[t] ~ dnorm(mu_S[t], tau)
         mu_t[t] <- alpha0 + alpha1 * x1[t] + alpha2 * x2[t]
         mu_S[t] <- gamma1 * mu_T[T] + gamma2 * mu_t[t]
      }
   }
}
```

Figure 5.7: `WinBUGS`-code for the directed acyclic graph (DAG) in figure 5.8. In the preferred spesification $x_{1,t} = A_t$ and $x_{2,t} = ES^{n=10}_{\alpha=0.99,t}$.

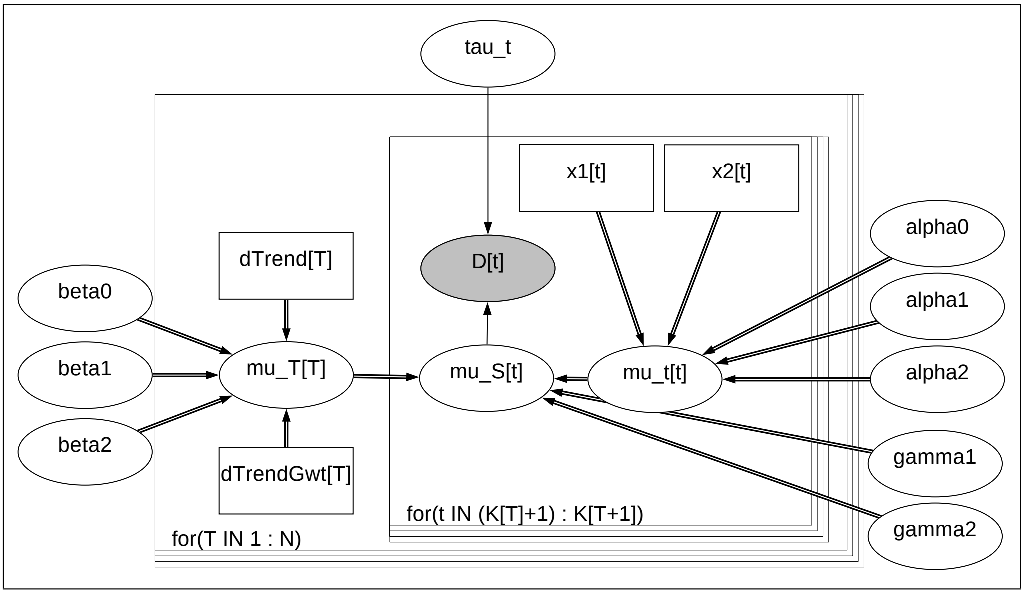


Figure 5.8: The DAG for the model implemented in `WinBUGS`. In the preferred specification the data D_t is the log-transformed ten day variance in the profit and losses in the portfolio, i.e. $D_t = \ln(PnL\_10DV_t)$; $x_{1,t} = A_t$ and $x_{2,t} = ES^{n=10}_{\alpha=0.99,t}$.

[14]Indeed, **`BUGS`** is abbreviation for "Bayesian inference Using Gibbs Sampling". See Spiegelhalter, Thomas, Best & Lunn (2003).

Chapter 6

Results and Interpretations

This chapter presents the results from the model constructed in the previous chapter. The preferred specification and the coefficients from the model is presented first, then a sensitivity analysis of the results is preformed. Finally, the results are interpreted and the conclusions follow in the next chapter.

6.1 Coefficient Estimates

Before the coefficient estimates are presented, a description of the covariates used in the preferred specification of the model is given in table 6.1.

Table 6.1: Overview of the data in the preferred specification.

Variable	Description	Time resolution	Time horizon
$dTrend_T$	Deviation from estimated trend [%]	quarterly	Q1/04 - Q4/04
$dTrendGwt_T$	Deviation from estimated trend growth [%]	quarterly	Q1/04 - Q4/04
A_t	Activity ($A = N \cdot V$, [$\# \cdot$ MNOK])	daily	01/01/04 - 10/11/04
$ES^{n=10}_{\alpha=0.99,t}$	Expected shortfall [% of portfolio]	daily	01/01/04 - 10/11/04
$\ln(PnL\_10DV)$	Logarithmic transformed 10 days sample variance in the portfolio's PnL [MNOK]	daily	01/01/04 - 10/11/04

Table 6.2 gives the parameter estimates when the model, given by equations (5.14) and (5.16), is iterated 150000 times (i.e. $m = 150000$ in algorithm 5.1) with $\boldsymbol{K} = (0, 65, 130, 196, 225)'$, $N = 4$, and the data described in table 6.1. The first 10000 estimates are removed due to burn-in.

Table 6.2 shows that none of the coefficients for variables from the micro-state are significantly different from zero with a 95% confidence interval, but that the coefficients belonging to the variables from the macro-state *are* significantly different from zero. This applies to the γ's as well; γ_1 is significant different from zero while γ_2 is not. One way to put this is that, after controlling for the macroeconomic state of the economy, neither the risk in the foreign exchange market (i.e. the ES measure) nor the activity in the portfolio have significantly effects on the risk appetite held by the portfolio's administrator. The fact that the activity in the portfolio is insignificant is not very surprising, but that riskiness in the portfolio at all times are under influence by the business cycles, not by

Table 6.2: Results from the model with the preferred specification given in equations (5.14) and (5.16) and with the data described in table 6.1. All initial values were set to zero except $\tau_{t=1} = 0.001$.

Node	Mean	Sd	MC error	2.5%	Median	97.5%	Start	Sample
α_0	0.238	1006.0	7.579	-1962.0	-0.7526	1971.0	10001	140000
α_1	-0.8114	267.6	1.262	-573.2	-1.712E-4	562.2	10001	140000
α_2	-1.429	967.1	4.05	-1917.0	0.202	1924.0	10001	140000
β_0	906.2	527.1	23.54	126.3	843.0	2145.0	10001	140000
β_1	221.7	132.6	5.732	30.54	203.1	539.1	10001	140000
β_2	-687.4	400.0	17.77	-1628.0	-636.7	-96.62	10001	140000
γ_1	-0.008618	0.009512	4.402E-4	-0.0369	-0.005551	-0.002069	10001	140000
γ_2	1.55E-5	5.409E-4	1.533E-5	-5.64E-4	-1.518E-10	6.537E-4	10001	140000
τ	0.3341	0.03178	8.696E-5	0.2745	0.3331	0.3992	10001	140000

the variability in the prices of the underlyings, *is* quite surprising. But now one must remember that this is a case study and that the time period is not very long; there is only four data points (quarters) in the macro-state. The signs of the coefficients, i.e. the directions of the dependencies, are discussed in chapter 6.3 below.

The figures 6.1 and 6.2 give trace and density plots, respectively, of the coefficient estimates. From figures 6.1(d-h) it becomes apparent that the mixing in the model is quite slow, but when the model is iterated 150000 times, the MC-error becomes less than 5 % of the sample standard deviation ("Sd" in table 6.2) which is, according to Spiegelhalter et al. (2003), adequate to conclude that the chain has converged.[1]

6.2 Sensitivity Analysis

There are two of the covariates in the model that could be altered without a change in interpretation; PnL_10DV_t and $ES^{n=10}_{\alpha=0.99,t}$. Both of these variables are constructed with the use of two working weeks of data, but there are no particular reason for using exactly two weeks. When the construction of the risk appetite and the benchmark risk measure are altered, i.e. the length of the moving variance band in equation (5.9) and n and α in the ES measure, the results remains robust; meaning that γ_1 and the β-coefficients are significantly different from zero, and that the others are not.

As there are only four observations on each of $dTrend$ and $dTrendGwt$, one cannot perform a test on the robustness of the coefficient estimates.[2]

6.3 Interpretations

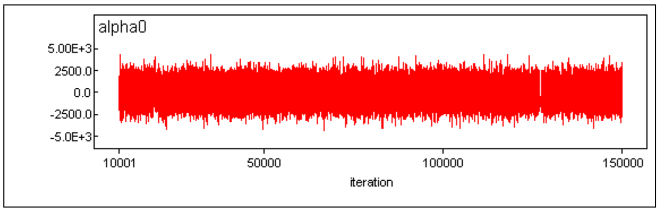
To attempt an interpretation of the results, one must look at the signs and quantities of the coefficient estimates. The partial effects, assuming everything else equal, can be found by partial differentiation:

$$\frac{\partial \ln (PnL\_10DV_t)}{\partial dTrend_T} = \gamma_1 \cdot \beta_1 = -0.008618 \cdot 221.7 = -1.911 < 0 \tag{6.1}$$

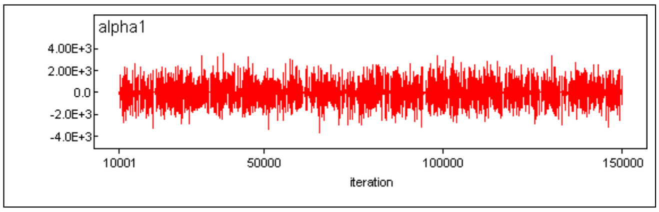
and

[1]The "MC-error" is an abbreviation for the "Monte Carlo error" and is an estimate of the difference between the mean of the sampled values (which `WinBUGS` is using as the estimate of the posterior mean for each parameter) and the true posterior mean.

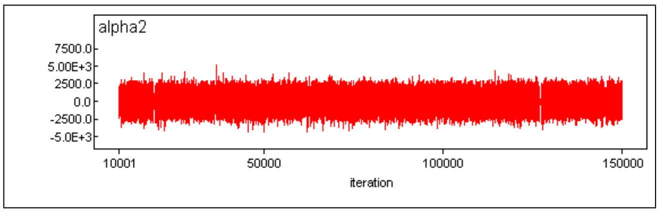
[2]For example use only one half of the data to make the coefficient estimates and then use this estimated model to predict the other half.



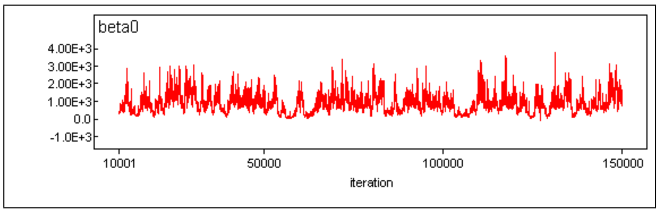
(a) α_0



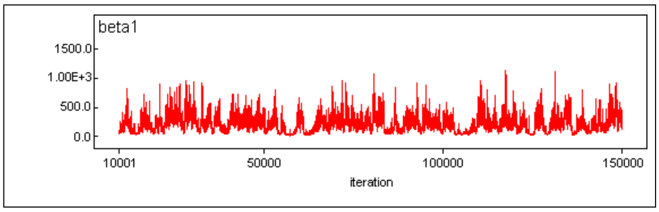
(b) α_1



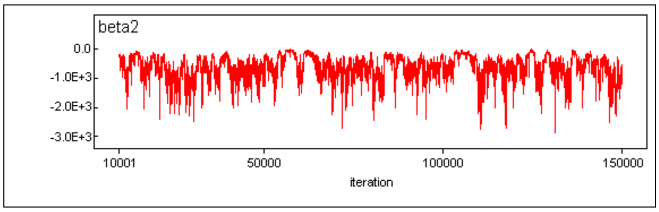
(c) α_2



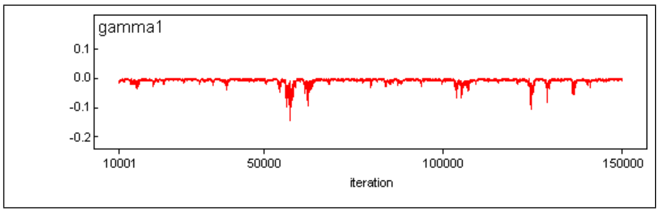
(d) β_0



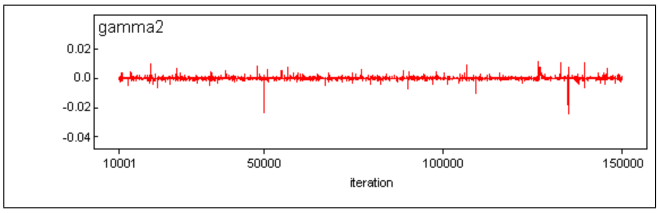
(e) β_1



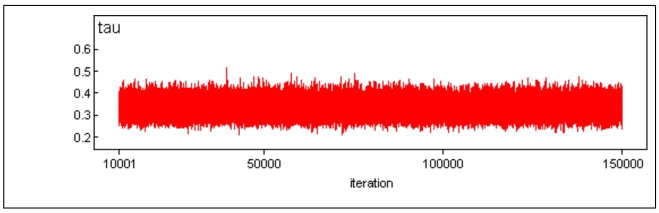
(f) β_2



(g) γ_1

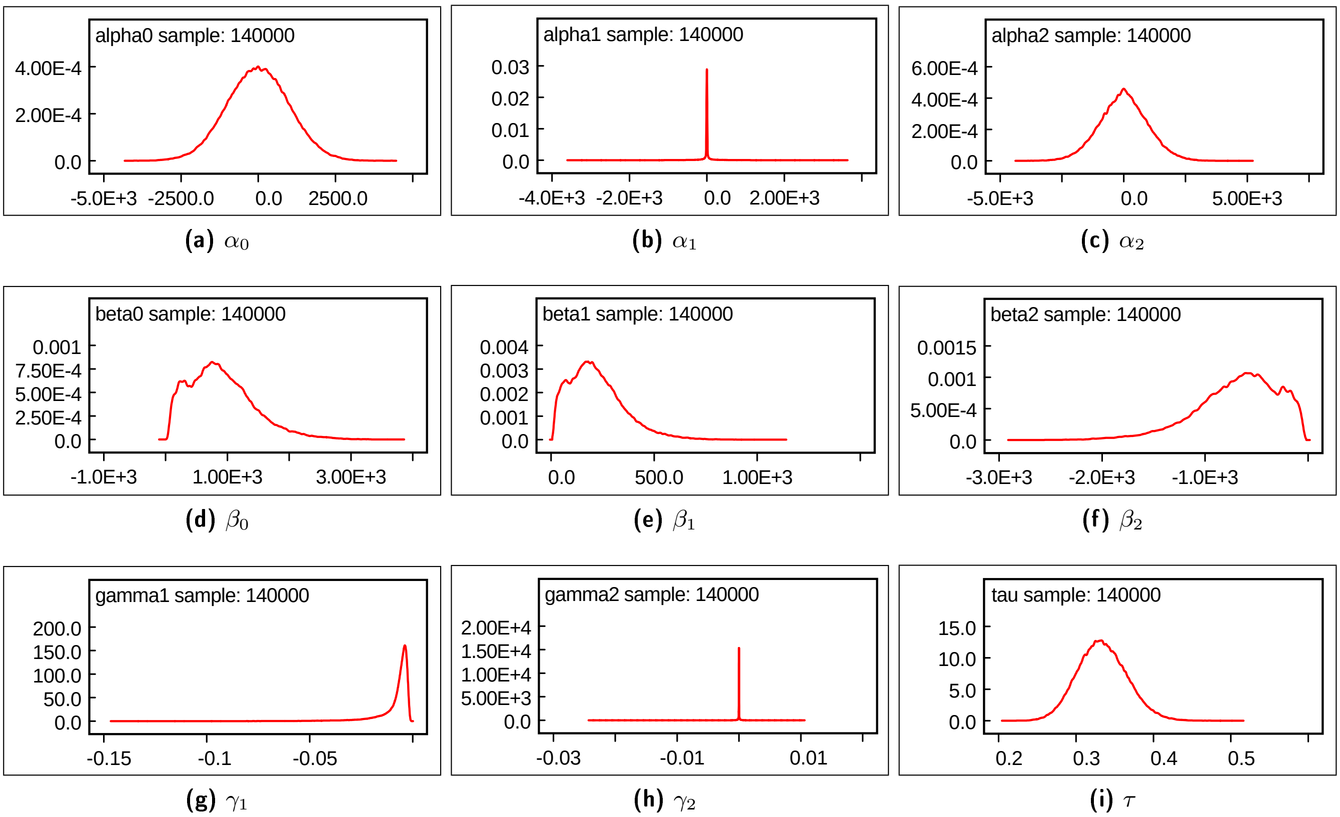


(h) γ_2



(i) τ

Figure 6.1: Trace plots of the variables in the model with the preferred specification, see table 6.1. The model is ran with 150000 iterations where the first 10000 are removed due to burn-in.



(a) α_0 (b) α_1 (c) α_2

(d) β_0 (e) β_1 (f) β_2

(g) γ_1 (h) γ_2 (i) τ

Figure 6.2: Density estimates of the variables in the model with the preferred specification (see table 6.1). The model is ran with 150000 iterations where the first 10000 are removed due to burn-in.

$$\frac{\partial \ln (PnL\_10DV_t)}{\partial dTrendGwt_T} = \gamma_1 \cdot \beta_2 = -0.008618 \cdot -687.4 = 5.924 > 0. \tag{6.2}$$

This means that, ceteris paribus, an increase in the deviation from trend in a quarter, results in a *decreased* risk appetite in the same quarter, while an increase in the deviation from trend growth *increases* the risk appetite.

6.3.1 The Risk Appetite in the Business Cycle

To give a more precise interpretation one can read these results into the phase diagram presented in figure 3.1(b): When the economy is in the part of the business cycle called catch-up, i.e. when $dTrend_T < 0$ and $dTrendGwt_T > 0$, the risk appetite will be growing. The opposite will be true in the state of cooling, while in periods with overheating and contraction equations (6.1) and (6.2) say that the risk appetite could be either increasing or decreasing dependent on the relative size between $dTrend_T$ and $dTrendGwt_T$. This is illustrated in figure 6.3.

So if one accepts that the macroeconomic environment is duly characterized by the business cycle defined as the deviation from trend and deviation from trend growth, one may conclude from table 6.2 that the risk appetite *is* dependent on the macroeconomic environment.

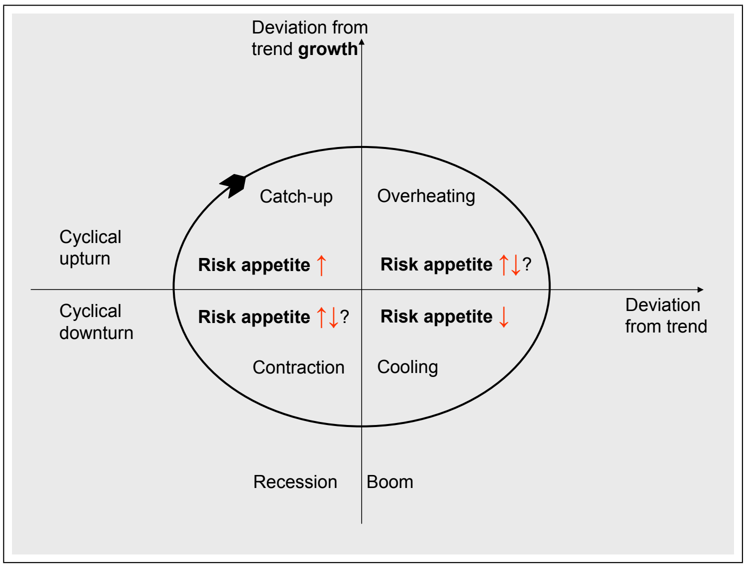


Figure 6.3: The risk appetite in the business cycle.

6.3.2 A Further Discussion

If one looks further into the reasons *why* the risk appetite should be dependent on the business cycles, the results from table 6.2 may be given an additional interpretation.

The logic behind the interdependence between the risk appetite and the macroeconomic environment is quite easy to explain in the case of a stock portfolio. In that case it is reasonable to predict that a “higher than normal” growth rate, i.e. a cyclical upturn, indicates high return forecasts and therefore induces a high willingness to bear risk, while an above normal resource exploitation, i.e. a boom, is an indication on an impending recession which implies lower expected future returns and therefore a low risk appetite.[3] So this argument is in accordance with the results from table 6.2 interpreted trough equations (6.1) and (6.2).

But there are some fundamental differences between the stock market and the foreign exchange market. In the former, the dependence between the risk appetite and the macroeconomic environment is a consequence of the fact that the *return forecasts* are high at the bottom of recessions and inverted at the top of a boom. The foreign exchange market on the other hand, is a zero-sum game and it is not very intuitive why the returns in this market, and therefore the risk appetite, should *directly* depend on the business cycles in the Norwegian economy. Thus the important question is:

Does the business cycle, measured as the deviation from trend and deviation from trend growth, encapsulate any form of uncertainty about the future value of a portfolio consisting of foreign currency, *that is not taken into account by the risk factors in the micro-state* (recall equation (5.12))?

If the answer is yes, then the results in table 6.2 are in correspondence with the theory set out by Gai & Vause (2006) and others. If the answer is no, i.e. that there are no direct or indirect ways the business cycle carries any more information about the risk in the exchange market than what is already given by the micro-state,[4] then the assumption set out by Gai & Vause (2006) saying that the risk aversion is *"unlikely to change markedly, or frequently, over time."* is questionable.[5]

This statement is a consequence of Gai & Vause's (2006) definition of risk appetite given in chapter 5.3.1 and the argument goes like this: If the level of uncertainty surrounding the future value of this currency portfolio is unaffected by the macroeconomic environment, but that the macroeconomic environment still has a significant impact on the risk appetite, this could mean that the macroeconomic environment has effect on the risk *aversion* which then again affects the risk appetite. This implies that the risk aversion *does* experience significant changes over time as visualized in figure 6.4. Here the new red arrow represents the impact from the macroeconomic environment on the risk aversion (and then on the risk appetite).

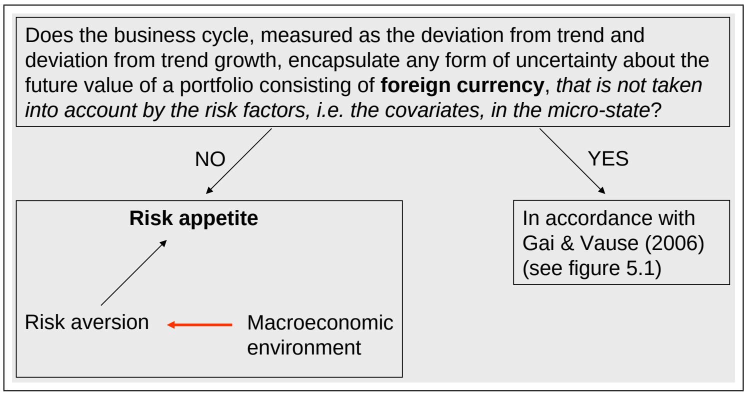


Figure 6.4: Alternative relationship between risk concepts saying that the risk aversion also is dependent on the macroeconomic environment.

[3]See for instance Cochrane (2001, chapter 20) for a survey on empirical results on expected returns in time series.

[4]Recall that the risk appetite is controlled for the benchmark risk measure before the effects from the macroeconomic environment is quantified; this is done by the micro-state in equation 5.12.

[5]Gai & Vause (2006, page 168).

If one believes that the answer to the question posed above is "no", then the results in table 6.2 may be interpreted as a psychological effect; a cyclical upturn gives the portfolio administrator the impression that he or she can endure greater risk ($\frac{\partial \ln\left(PnL\_10DV_t\right)}{\partial dTrendGwt_T} > 0$), while the expectation of a impending cyclical downturn makes the portfolio administrator more afraid of taking on risk ($\frac{\partial \ln\left(PnL\_10DV_t\right)}{\partial dTrend_T} < 0$).

There are also reasons why the answer to the question above may be "yes": Even though the level on expected *returns* in the currency portfolio at hand does not depend on the macroeconomic environment, it is possible that the macroeconomic environment carries information about the *volatility* on the returns and therefore on the volatility of the future value of the portfolio. If this is the case, the interpretation of the results in table 6.2 would be that an increased deviation from trend, ceteris paribus, induces a higher level of uncertainty and hence a lower risk appetite, while an increased deviation from trend growth, ceteris paribus, induces a lower level of uncertainty on future volatilities and hence a higher risk appetite.

So much said, this thesis will not and cannot give an unambiguous conclusion as the data foundation is small. To be confident about the results one needs to estimate the coefficients through a full cycle, preferably. This last discussion is merely meant to question the distinction between risk appetite and risk aversion.

6.3.3 Quantifying the Results

The quantification of these effects is somewhat demanding to interpret as the risk appetite is measured as a logarithmic transformed moving variance and not in some relative measure (because the data does not tell the size of the portfolio); the unit of the risk appetite measure is ln(MNOK^2).

From equation (6.1) one can say that a one percentage increase in $dTrend_T$ in a quarter gives a -1.911 decrease in $\ln\left(PnL\_10DV_t\right)$; equation (6.2) states that a one percentage increase in $dTrendGwt_T$ in a quarter gives a 5.924 increase in $\ln\left(PnL\_10DV_t\right)$ for all micro-states in the same quarter.

To put this into perspective one can note that the standard deviation for the risk aversion during the whole period is 1.98. This means that an one percentage point increase in $dTrend_T$ will reduce the risk aversion with approximately one standard deviation in all micro-states in the relevant quarter, while one percentage point increase in $dTrendGwt_T$ in a quarter will increase the risk aversion with approximately three standard deviations in all micro-states in the relevant quarter.

Chapter 7

Conclusions

This chapter will first give a quick walk-through of the model with all its building stones, then the assumptions and limitations underlying the model will be addressed before the conclusions are stated.

7.1 A Quick Walk-through

Before the conclusions is drawn it is perhaps helpful to recapitulate the building of the model and its main parts.[1]

The intention in this thesis was to establish a methodology which could answer whether the risk appetite is dependent on the macroeconomic state in an economy. The first step was to look into the literature on market risk and how this should be measured. This was done in chapter 2 with emphasis on Value at Risk and Expected Shortfall. The next step was to find a way to measure the macroeconomic state of the economy. In chapter 3 it was argued that this could be done using a phase diagram which classified the economy in terms of its deviation from the estimated trend and its deviation from the estimated trend growth. To carry out this classification, a Hodrick-Prescott-filter was constructed and applied.

The model was built around the collected data which was presented in chapter 4. The data consisted of the profit and losses in a currency portfolio together with the number and volume of transactions each day and time series with different exchange rates. This chapter also presented the results from the HP-filter. Here the portfolio and exchange rate data had a daily time resolution, while the business cycles where classified every quarter.

The main model was presented in chapter 5. To overcome the different time resolutions of the data, the model was constructed with a Bayesian approach. A short introduction to this statistical method was given and a hierarchic approach was chosen as several microeconomic states were contained in one macroeconomic state. The risk appetite was measured as a moving variance in the portfolio's profit and losses and two (microeconomic) risk factors were chosen: The *activity* A and $ES^{n=10}_{\alpha=0.99,t}$. Further, the model assigned uninformative prior distributions to all the variables (which also is, except τ, referred to as "coefficients") and it was assumed that the risk appetite was normal distributed given the other variables in the model.

[1] Recall figure 1.1 as this might help to visualize this walk-through.

7.2 Critical Assumptions and Reservations

The most critical element in the model is perhaps not any of the assumptions, but the data used. As stressed several times through this thesis, the data which the risk appetite is derived from is collected from just one portfolio and this is of course not likely to be representative for the (Norwegian) financial market as such. Nevertheless, this is a quite large portfolio administrated by one of the largest Norwegian banks, DnBNOR, so it does carry important information and serves as a highly relevant case study.

The fact that the model finds that the risk appetite in this one portfolio is dependent on the business cycles does *not* mean that the risk appetite in the *market* is dependent on the business cycles, nor that the risk appetite in this specific portfolio is independent of other risk factors that are not considered here.

Another considerable drawback concerning the data is its short time horizon. Because there is not more than a year with portfolio data, this means that the model can use only four data points (quarters) classifying the business cycles which makes the coefficient estimates very sensitive towards perturbations in the data. The lack of data also makes this sensitivity hard to quantify and further analyses with more data would be needed for firm conclusions.

There are also two assumptions made in the model that one should recall before the conclusions are stated. One of them is the normality assumption regarding the risk appetite (see equation (5.14)). As figures 5.2(b-c) shows, this is a crude assumption as the empirical distribution has heavier tails than its theoretical counterpart which means that extreme outcomes are more likely than this assumption predicts. Another rough approximation is the multinormal assumption implied in the variance-covariance method used when the benchmark risk measure was constructed in chapter 5.3.2. Figure 5.3 shows that this approximation is inaccurate.

7.3 Conclusion

This thesis has attempted to establish a methodology to measure the dependency between risk appetite and the macroeconomic environment with use of a hierarchical Bayesian approach. The results from table 6.2 shows that the risk appetite, measured as the logarithmic transformed ten days variance in the profit and losses in the selected portfolio, *is* dependent on the business cycle, measured as the deviation from trend and deviation from trend growth. The risk appetite depends negatively on the deviation from trend and positively on the deviation from trend growth. The causality behind this dependency, however, is still unclear and alternative interpretations may be considered as discussed in chapter 6.3.2.

7.4 Possible Extensions

A model such as the one built and used in this thesis has a range of possible extensions. A first extension (and improvement) would be to consider more explanatory variables in the micro-state given in equation (5.2) than the two chosen (i.e. the activity A and $ES^{n=10}_{\alpha=0.99}$) although one normally would think that the expected shortfall measure would incorporate the relevant risk factors (i.e. the log-returns).

Further, one could also look into whether lagged observations of the macroeconomic covariates in the model have significant effects. The logic behind would be that the risk appetite is formed on the basis of expectations about the future state of the economy which may be argued to be dependent on previous macroeconomic states.

It is also possible to specify the macroeconomic state in alternative ways. One could for instance decompose the business cycles as mentioned in chapter 3.3 as this perhaps would allow for a more specific interpretation of the dependencies between the risk appetite and the macroeconomic environment.

Finally, one should note that the modelling approach used in this thesis is quite generic and can easily be applied to other types of portfolios and classes of assets.

Bibliography

Artzner, P., Delbaen, F., Eber, J.-M. & Heath, D. (1999), 'Coherent Measures of Risk', *Mathematical Finance* (9), 203–228.

Bernardo, J. M. & Smith, A. F. M. (1994), *Bayesian Theory*, John Wiley & Sons.

Brockwell, P. J. & Davis, R. A. (2002), *Introduction to Time Series and Forecasting*, Springer.

Brooks, C. (2002), *Introductory Econometrics for Finance*, Cambridge University Press.

Cochrane, J. H. (2001), *Asset Pricing*, Princeton University Press.

Coles, S., Roberts, G. & Jarner, S. (2002), Computer Intensive Methods. Not published. Internet: `http://www.actuar.aegean.gr/Undergraduate/Ypologistikh%20statistikh.ps`.

Cont, R. & Tankov, P. (2004), *Financial Modelling With Jump Processes*, Chapman & Hall/CRC.

Crouchy, M., Galai, D. & Mark, R. (2001), *Risk Management*, McGraw-Hill.

Föllmer, H. & Schied, A. (2002), 'Convex measures of risk and trading constraints', *Finance and Stochastics* **6**(4), 429–447.

Gai, P. & Vause, N. (2006), 'Measuring Investors' Risk Appetite', *International Journal of Central Banking* **2**(1), 167–188.

Geman, S. & Geman, D. (1984), 'Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images', *IEEE Transactions on pattern analysis and machine intelligence* **6**, 721–741.

Glasserman, P. (2004), *Monte Carlo Methods in Financial Engineering*, Springer.

Global Derivatives Study Group (1993), 'Derivatives: Practices and Principles'. Published by The Group of Thirty, Washington, DC.

Green, P. J. (2001), A Primer on Markov chain Monte Carlo, *in* 'Complex Stochastic Systems', Chapman and Hall, pp. 1–62.

Heijdra, B. J. & van der Ploeg, F. (2002), *The Foundations of Modern Macroeconomics*, Oxford.

Hodrick, R. J. & Prescott, E. C. (1997), 'Postwar U.S. Business Cycles: An Empirical Investigation', *Journal of Money, Credit and Banking* **29**(1), 1–16.

Holton, G. A. (2003), *Value-at-Risk*, Academic Press.

Johansen, P. R. & Eika, T. (2000), 'Driving forces behind business cycles in the 1990s', *Economic Survey* (3), 25–38.

Johnson, R. A. & Wichern, D. W. (2002), *Applied Multivariate Statistical Analysis*, Prentice Hall. International edition.

Jorion, P. (2001), *Value at Risk*, McGraw-Hill.

Øksendal, B. (2003), *Stochastic Differential Eqations: An Introduction with Applications*, Springer.

Kydland, F. E. & Prescott, E. C. (1990), 'Business cycles: Real facts and a monetary myth', *Federal Reserve Bank of Minneapolis Quarterly Review* **14**(2), 3–18.

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

McNeil, A. J., Frey, R. & Embrechts, P. (2005), *Quantitative Risk Management: Concepts, Techniques and Tools*, Princeton University Press.

Misina, M. (2003), What Does the Risk-Appetite Index Measure? Bank of Canada Working Paper 2003-23. Internet: `http://www.bankofcanada.ca/en/res/wp/2003/wp03-23.pdf`.

Price, K., Price, B. & Nantell, T. J. (1982), 'Variance and Lower Partial Moment Measures of Systematic Risk: Some Analytical and Empirical Results', *The Journal of Finance* **37**(3), 843–855.

Rødseth, A. (2000), *Open Economy Macroeconomics*, Cambridge University Press.

Rebelo, S. (2005), 'Real business cycle models: Past, present and future', *NBER working paper series* . Working paper nr. 11401, Internet: `http://www.nber.org/papers/w11401`.

Reeves, J. J., Blyth, C. A., Triggs, C. M. & Small, J. P. (2000), 'The Hodrick-Prescott Filter, a Generalization, and a New Procedure for Extracting an Empirical Cycle from a Series', *Studies in Nonlinear Dynamics and Econometrics* **4**(1), 1–16.

Romer, D. (2001), *Advanced Macroeconomics*, McGraw-Hill.

Spiegelhalter, D., Thomas, A., Best, N. & Lunn, D. (2003), *WinBUGS User Manual.* Internet: `http://www.mrc-bsu.cam.ac.uk/bugs`.

The Royal Swedish Academy of Sciences (2004), Finn Kydland and Edward Prescott´s Contribution to Dynamic Macroeconomics: The Time Consistency of Economic Policy and the Driving Forces Behind Business Cycles. Internet: `http://www.kva.se/KVA_Root/files/newspics/DOC_20041011123429_28842881920_ecoadv04.pdf`.

Unser, M. (2000), 'Lower partial moments as measures of perceived risk: An experimental study', *Journal of Economic Psychology* (21), 253–280.

Varian, H. R. (1999), *Intermediate Microeconomics: A Modern Approach*, W.W. Norton & Company.

Wilmott, P. (2001), *Paul Wilmott Introduces Quantitative Finance*, John Wiley & Sons.

Wooldridge, J. M. (2003), *Introductory Econometrics : A Modern Approach*, Thomson South-Western.

Appendix A

Appendix

A.1 The Probability Space $(\Omega, \mathcal{F}, P)$

In more advanced parts of the statistical literature one uses the probability space $(\Omega, \mathcal{F}, P)$ and not just the probability distribution. This distinction is perhaps not very intuitive and Cont & Tankov (2004, chapter 2.2.2) offers an excellent explanation on why one sometimes need the hole space and not just the distribution. Øksendal (2003) gives a concise definition of the terms in the space $(\Omega, \mathcal{F}, P)$ and the properties needed:[1]

Definition A.1 If Ω is a given set, a σ-algebra $\mathcal{F}$ on Ω is a family F of Ω with the following properties:

$i)$ $\emptyset \in \mathcal{F}$
$ii)$ $F \in \mathcal{F} \Rightarrow F^C \in \mathcal{F}$ where $F^C = \Omega \backslash F$ is the complement of F in Ω
$iii)$ $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

The pair $(\Omega, \mathcal{F})$ is called a measurable space. A probability measure P on a measurable space $(\Omega, \mathcal{F})$ is a function $P : \mathcal{F} \to [0, 1]$ such that

a) $P(\emptyset) = 0, \; P(\Omega) = 1$
b) If $A_1, A_2, \ldots \in \mathcal{F}$ and $\{A_i\}_{i=1}^{\infty}$ is disjoint (i.e. $A_i \cap A_j = \emptyset$ if $i \neq j$) then
$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$

The triple $(\Omega, \mathcal{F}, P)$ is called a probability space.

A.2 The Choice and Scaling of the Time Horizon

VaR_α and ES_α is risk measures that tells you something about how much of your portfolio you can loose, with some probability, *over a pre-specified time horizon.* A natural question to ask would be: *"If I have calculated VaR_α with $\Delta = 1$ day, how can I use this to find VaR_α for one week?"*. One simlpe answer would be to *"aggregate the daily risk factor change data to a weekly resolution and calculate VaR_α once again"*. Or, one could think that the daily estimate could be scaled in some way to obtain a weekly VaR_α measure. Generally one seeks a procedure to transform a one-period risk measure into a h-period risk measure for $h > 1$. Let $L_{t+h}^{(h)}$ denote the loss from time t over the next h periods. Similar to equation (2.4) one can write

[1] This is a shortened version of (Øksendal 2003, Definition 2.1.1).

$$\begin{aligned} L_{t+h}^{(h)} &= -(V_{t+h} - V_t) \\ &= -(f(t+h, \boldsymbol{Z}_{t+h}) - f(t, +\boldsymbol{Z}_t)) \\ &= -(f(\boldsymbol{Z}_t + \boldsymbol{X}_{t+1} + \ldots + \boldsymbol{X}_{t+h}) - f(t, \boldsymbol{Z}_t)) \\ &= -(f(\boldsymbol{Z}_t + \sum_{i=1}^{h} \boldsymbol{X}_{t+i}) - f(t, \boldsymbol{Z}_t)) \end{aligned}$$

such that $\Delta = h$. Now the question becomes how risk measures applied to the distribution of $L_{t+h}^{(h)}$ scale with h which does not have a simple answer except in special cases. An exemplification of an important special case is found in example A.1.

Example A.1 Assume that the risk factor change vectors, $\boldsymbol{X}_t$, are independent and identically distributed $N(\boldsymbol{0}, \Sigma)$. Then $\sum_{i=1}^{h} \boldsymbol{X}_{t+i} \sim N(\boldsymbol{0}, h\Sigma)$. Further, assume that the underlying assets of the portfolio in question do not depend on calendar time[2] and that the loss is linearly dependent on the risk factors, i.e. that $L_{t+h}^{h} = \sum_{i=1}^{h} \boldsymbol{b}_t' \boldsymbol{X}_{t+i}$ where $\boldsymbol{b}_t$ is a known vector at time t.[3] Then it follows that $L_{t+h}^{(h)} \sim N(\boldsymbol{0}, h\boldsymbol{b}_t'\Sigma\boldsymbol{b}_t)$ and that both the VaR and ES_α measure will scale according to the square root of time ($\sqrt{h}$) such that

$$VaR_\alpha^{(h)} = \sqrt{h}\sigma\Phi^{-1}(\alpha) = \sqrt{h}VaR_\alpha^{(1)} \tag{A.2}$$

$$ES_\alpha^{(h)} = \sqrt{h}\sigma\frac{\phi(\Phi^{-1}(\alpha))}{1-\alpha} = \sqrt{h}ES_\alpha^{(1)} \tag{A.3}$$

where $VaR_\alpha^{(1)}$ and $ES_\alpha^{(1)}$ are found in equations (2.7) and (2.9), respectively.

A.3 Seasonal Adjustment; the Classical Decomposition Model

Figure A.1 offers an implementation of the classical decomposition model done in `R`. The implementation is based on Brockwell & Davis (2002). This algorithm first checks whether the period is odd or even, then it extracts the trend component T using a moving average filter chosen to eliminate the seasonal component and dampen the noise. The filter depends then on whether the period is odd or even. Then the algorithm finds the difference between the original series and the trend and then calculates the mean for each time interval within the full period. A estimate on the seasonal element, *ses.temp*, is then found using the mean described above adjusted with the mean for all those means such that the sum of all the seasonal adjustments over all elements within the period is zero. The name *ses.temp* is due to the fact that the indices of this estimates is displaced compared to its actual values; this is accounted for in the if-sentence that follows. The algorithm replicates the estimates of the seasonal components such that the returned vector is equally long as the original series (the input).

A.4 `R`-code for the HP-filter

The code in figure A.2 applies the HP-filter to a timeseries used to extract the business cycles in the Norwegian economy. Essentially this is an implementation of equation (3.5) with the use of the `solve`-function in `R`.

[2]This excludes for instance derivatives.

[3]E.g. the weight of each asset in the portfolio.

```
season <- function(x,d){
        ######################
        # The classical decomposition model.
        # The function takes in a vector with data, x, and
        # the period d. It returns the seasonal components.
        ######################

        # is the period d even or odd?
        q <- 0
        indicator <- 0
        if(d %% 2 == 0){
                q <- d/2
                indicator <- 0 # period is even
        }
        else{
                q <- (d-1)/2
                indicator <- 1 # period is odd
        }

        # estimate the trend vector T
        T <- 0
        T[1:q] <- NA
        T[(q+1):length(x)] <- NA
        if(indicator == 0){
                for(i in (q+1):(length(x)-q)){
                        T[i] <- (1/d)*(.5*x[i-q] + sum(x[(i-q+1):(i+q-1)]) + .5*x[i+q])
                }
        }
        else{
                for(i in (q+1):(length(x)-q)){
                        T[i] <- mean(x[i-q:i+q])
                }
        }

        # calculate the difference between the original data and the trend
        w <- 0
        for(k in 1:d){
                indices <- seq(k+q,length(T)-q-((length(T)-q) %% d),d)
                w[k] <- mean(x[indices]-T[indices]) # nr k is displaced with q
        }

        # estimate the seasonal components (which shall add to 0)
        ses.temp <- 0
        ses <- 0
        for(l in 1:d){
                ses.temp[l] <- w[l]- mean(w)
                if(indicator == 0)
                        if(l <= q){
                                ses[q+l] <- ses.temp[l]
                        }
                        else{
                                ses[l-q] <- ses.temp[l]
                        }
                else{
                        if(l <= (q+1)){
                                ses[q+l] <- ses.temp[l]
                        }
                        else{
                                ses[l-q+1] <- ses.temp[l]
                        }
                }
        }

        ses.vector <- 0
        for(p in 1:length(x)){
                if(p<d){
                        ses.vector[p] <- ses[p]
                }
                if(p %% d == 0){
                        ses.vector[p] <- ses[d]
                }
                else{
                        ses.vector[p] <- ses[p %% d]
                }
        }

        return(ses.vector)

}# end function
```

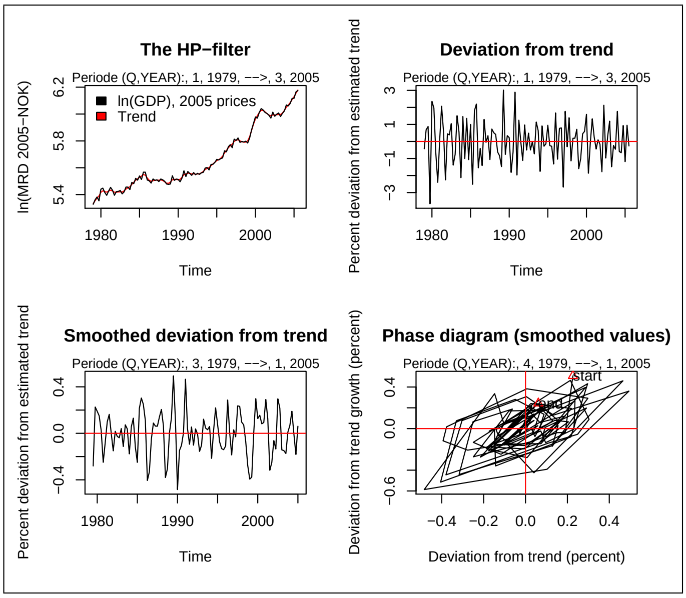
Figure A.1: R-code for the seasonal adjustment of the GDP series used in the HP-filter.

```
hp<-function(x,lambda,year.start,qrt.start,year.end,qrt.end){

        ######################
        # The traditional Hodrick-Prescott filter.
        # hp is a function that takes in six arguments; the
        # dataseries (log-transformed) that
        # is to be decomposed in a cyclic and a growth component,
        # the weight parameter lambda and
        # the start and end year and quarter.
        # It returns a matrix with the trend (growth component).
        ######################

        # the length of the series that is to be decomposed
        l <- length(x)

        # the input vector is ensured to be a column vector
        y <- matrix(x, nrow=l, ncol=1)

        # intitialize the filter F
        F <- diag(6, nrow=l, ncol=l)

        # filling in "the end values"
        F[2,1] <- -2; F[1,2] <- -2
        F[l-1,l] <- -2; F[l,l-1] <- -2
        F[1,1] <- 1; F[l,l] <- 1
        F[2,2] <- 5; F[l-1,l-1] <- 5

        # filling in the other diagonals
        for(i in 1:(l-2)){
                F[i,2+i] <- 1
                F[2+i,i] <- 1
        }
        for(k in 1:(l-3)){
                F[1+k,2+k] <- -4
                F[2+k,1+k] <- -4
        }

        # make an identity matrix
        I <- diag(1, nrow=l, ncol=l)
        intermediate <- ((lambda*F)+I)

        # apply the filter on the dataseries y and return the trend estimate
        # together with the original series and the deviation from trend
        g <- solve(a=intermediate, b=y)
        res<-matrix(NA,nrow=l,ncol=3)
        res[,1] <- y
        res[,2] <- g
        res[,3] <- y-g

        # the smoothed series is 2 periodes entries than
        # the original in both ends; 4 entries together
        smooth.dev.trend<-0
        for(j in 3:(l-2)){
                smooth.dev.trend[j-2] <- 1/8*res[j-2,3]+
                        1/4*(res[j-1,3]+res[j,3]+res[j+1,3])+1/8*res[j+2,3]
        }

        # the vector containing the deviation from trend is one
        # entry shorter than the original data series
        dev.trend.growth <- (res[c(-1),1]-res[1:(l-1),1])-
                        (res[c(-1),2]-res[1:(l-1),2])

        # the smoothed deviation from trend growth is five
        # entries shorter than the original data series;
        # four is 'lost' due to the smoothing process,
        # one entry due to the growth estimation
        smooth.dev.trend.growth <- 0
        for(j in 3:(l-3)){
                smooth.dev.trend.growth[j-2] <- 1/8*dev.trend.growth[j-2]+
                        1/4*(dev.trend.growth[j-1]+dev.trend.growth[j]+dev.trend.growth[j+1])+
                        1/8*dev.trend.growth[j+2]
        }

        # make plots (code not included here);
        # -the trend vs the original series
        # -the deviation from trend
        # - ----------"----------- smoothed with a centered MA(5)-process
        # - the phase diagram (with and without the smoothed values)

        return(res)
}
```

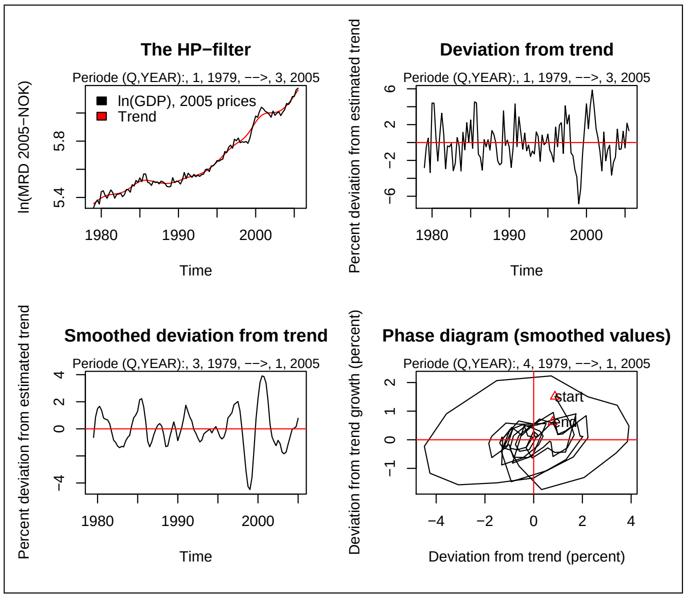
Figure A.2: R-code for the HP-filter.

A.5 A Sensitivity Analyses of λ in the HP-filter

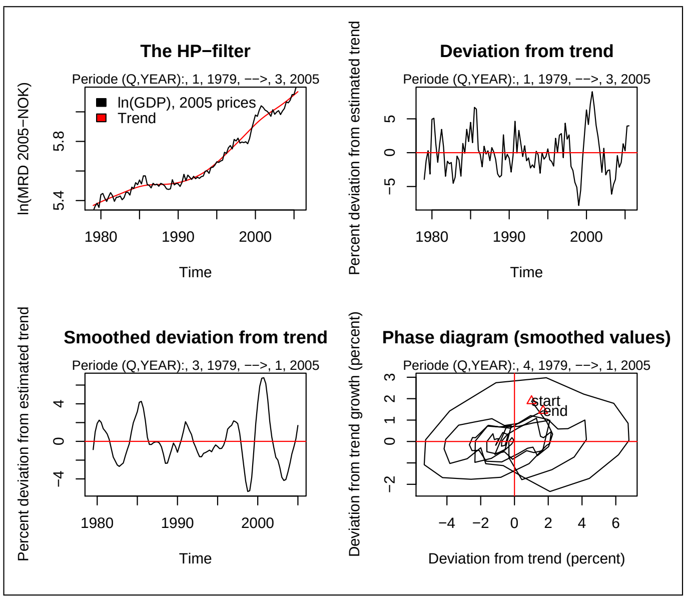
Figure A.3 shows how different values of λ affects the results of the filter; for small values the estimated trend is very close to the original GDP series, while a large value on λ produces an almost constant trend. Note that $\lambda = 1600$ is the size recommended by Hodrick & Prescott (1997) for the US economy.



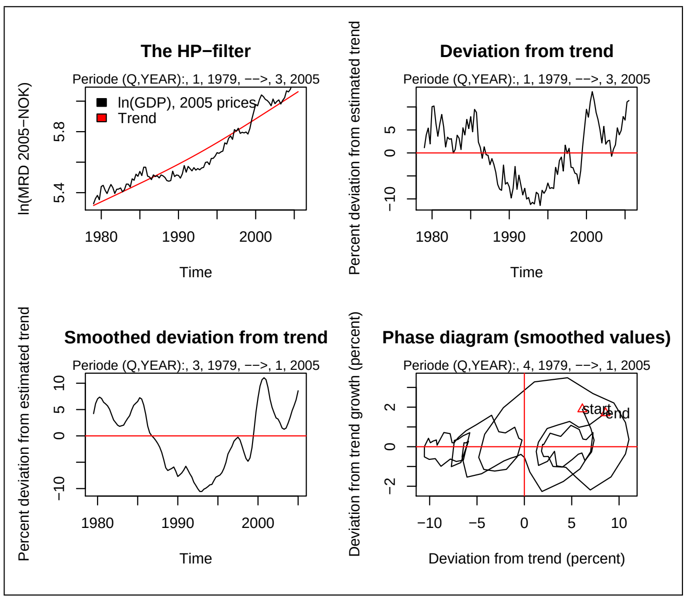
(a) $\lambda = 1$



(b) $\lambda = 100$



(c) $\lambda = 1600$



(d) $\lambda = 1000000$

Figure A.3: The HP-filter with different values of λ; (a): $\lambda = 1$, (b): $\lambda = 100$,(c): $\lambda = 1600$ and (d): $\lambda = 1000000$.

A.6 R-code for the ES and VaR Estimates

The code in figures A.4 and A.5 constructs the ES and VaR estimates used in the micro-state in equation 5.2.

```
ES <- function(return.data, est.length, alpha){

        n.row <- dim(return.data)[1]
        n.col <- dim(return.data)[2]
        variance <- 0
        mean.return <- 0
        ES.res <- 0
        mu <- matrix(NA, nrow=(n.row-est.length) ,ncol=n.col)
        cutoff <- 0

        # initializing the weights
        w <- matrix(rep(1/n.col,n.col),nrow=n.col, ncol=1)

        for(t in 1:(n.row-est.length+1)){
                mean.return[t] <- mean(return.data[t:(t+est.length-1),])
                variance[t] <- t(w)%*%cov(return.data[t:(t+est.length-1),1:n.col])%*%w
                ES.res[t] <- -mean.return[t]+sqrt(variance[t])*dnorm(qnorm(alpha))/(1-alpha)
        }
        return(ES.res)
}
```

Figure A.4: R-code for the ES estimates in figure 5.4.

```
VaR <- function(return.data, est.length, alpha){

        n.row <- dim(return.data)[1]
        n.col <- dim(return.data)[2]
        variance <- 0
        mean.return <- 0
        VaR.res <- 0
        mu <- matrix(NA, nrow=(n.row-est.length) ,ncol=n.col)
        cutoff <- 0

        # initializing the weights
        w <- matrix(rep(1/n.col,n.col),nrow=n.col, ncol=1)

        for(t in 1:(n.row-est.length+1)){
                mean.return[t] <- mean(return.data[t:(t+est.length-1),])
                variance[t] <- t(w)%*%cov(return.data[t:(t+est.length-1),1:n.col])%*%w
                VaR.res[t] <- -mean.return[t]+sqrt(variance[t])*qnorm(alpha)
        }
        return(VaR.res)
}
```

Figure A.5: R-code for the VaR estimates in figure 5.4.

A.7 Calculations of the Conditional Distributions

Below one can find the calculations for the conditional distributions used in the Gibbs sampler in section 5.4. All calculations are based upon Bayes Theorem from equation (5.5). To save space, R_PnL_t is used instead of $\ln(PnL\_10DV)_t$ from the preferred spesification.

In the beginning the procedure is the same for the α's, β's, γ's and τ and it is convenient to introduce the general parameter space $\phi \in \{\alpha_i, \beta_j, \gamma_k\}$ for $i \in (0, 1, \ldots, q)$, $j \in (0, 1, 2, 3)$ and $k \in (1, 2)$.

$$
\begin{aligned}
\pi(\phi| - \phi) &\propto \pi(\phi) \prod_{T=1}^{N} \prod_{t=K_T+1}^{K_{T+1}} \pi(R\_PnL_t | \alpha_0, \alpha_1, \ldots, \alpha_q, \beta_0, \beta_1, \beta_2, \gamma_1, \gamma_2, \tau) \\
&= \frac{1}{\sqrt{2\pi}\sigma_\phi} \exp\left\{ -\frac{1}{2\sigma^2_{\phi,p}} (\phi - \mu_{\phi,p})^2 \right\} \prod_{T=1}^{N} \prod_{t=K_T+1}^{K_{T+1}} \sqrt{\frac{\tau}{2\pi}} \times \\
&\quad \exp\Big\{ -\frac{\tau}{2} [R\_PnL_t - \gamma_1(\beta_0 + \beta_1 dTrend_T + \beta_2 dTrendGwt_T) - \\
&\quad \gamma_2(\alpha_0 + \alpha_1 x_{1,t} + \alpha_2 x_{2,t} + \ldots + \alpha_q x_{q,t})]^2 \Big\} \\
&= \frac{1}{\sqrt{2\pi}\sigma_\phi} \exp\left\{ -\frac{1}{2\sigma^2_{\phi,p}} (\phi - \mu_{\phi,p})^2 \right\} \left(\frac{\tau}{2\pi}\right)^{\left(\frac{K_{N+1}}{2}\right)} \prod_{t=K_1+1}^{K_2} \exp\Big\{ -\frac{\tau}{2} [R\_PnL_t - \\
&\quad \gamma_1(\beta_0 + \beta_1 dTrend_{T=1} + \beta_2 dTrendGwt_{T=1}) - \\
&\quad \gamma_2(\alpha_0 + \alpha_1 x_{1,t} + \alpha_2 x_{2,t} + \ldots + \alpha_q x_{q,t})]^2 \Big\} \prod_{t=K_2+1}^{K_3} \exp\Big\{ -\frac{\tau}{2} [R\_PnL_t - \\
&\quad \gamma_1(\beta_0 + \beta_1 dTrend_{T=2} + \beta_2 dTrendGwt_{T=2}) - \\
&\quad \gamma_2(\alpha_0 + \alpha_1 x_{1,t} + \alpha_2 x_{2,t} + \ldots + \alpha_q x_{q,t})]^2 \Big\} \times \cdots \times \\
&\quad \prod_{t=K_N+1}^{K_{N+1}} \exp\Big\{ -\frac{\tau}{2} [R\_PnL_t - \\
&\quad \gamma_1(\beta_0 + \beta_1 dTrend_{T=N} + \beta_2 dTrendGwt_{T=N}) - \\
&\quad \gamma_2(\alpha_0 + \alpha_1 x_{1,t} + \alpha_2 x_{2,t} + \ldots + \alpha_q x_{q,t})]^2 \Big\} \\
&= \frac{1}{\sqrt{2\pi}\sigma_\phi} \left(\frac{\tau}{2\pi}\right)^{\left(\frac{K_{N+1}}{2}\right)} \exp\Big\{ -\frac{1}{2} \Big[\frac{1}{\sigma^2_{\phi,p}} (\phi - \mu_{\phi,p})^2 + \\
&\quad \tau \sum_{t=K_1+1}^{K_2} [R\_PnL_t - \gamma_1(\beta_0 + \beta_1 dTrend_{T=1} + \beta_2 dTrendGwt_{T=1}) - \\
&\quad \gamma_2(\alpha_0 + \alpha_1 x_{1,t} + \alpha_2 x_{2,t} + \ldots + \alpha_q x_{q,t})]^2 + \\
&\quad \tau \sum_{t=K_2+1}^{K_3} [R\_PnL_t - \gamma_1(\beta_0 + \beta_1 dTrend_{T=2} + \beta_2 dTrendGwt_{T=2}) - \\
&\quad \gamma_2(\alpha_0 + \alpha_1 x_{1,t} + \alpha_2 x_{2,t} + \ldots + \alpha_q x_{q,t})]^2 + \cdots + \\
&\quad \tau \sum_{t=K_N+1}^{K_{N+1}} [R\_PnL_t - \gamma_1(\beta_0 + \beta_1 dTrend_{T=N} + \beta_2 dTrendGwt_{T=N}) - \\
&\quad \gamma_2(\alpha_0 + \alpha_1 x_{1,t} + \alpha_2 x_{2,t} + \ldots + \alpha_q x_{q,t})]^2 \Big] \Big\}
\end{aligned}
$$

$$\begin{aligned}
\pi(\phi | -\phi) &\propto \frac{1}{\sqrt{2\pi}\sigma_\phi}\left(\frac{\tau}{2\pi}\right)^{\left(\frac{K_{N+1}}{2}\right)} \exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma^2_{\phi,p}}(\phi-\mu_{\phi,p})^2 + \right.\right. \\
&\quad \tau\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}[R_PnL_t - \gamma_1(\beta_0 + \beta_1 dTrend_T + \beta_2 dTrendGwt_T) - \\
&\quad \left.\left. \gamma_2(\alpha_0 + \alpha_1 x_{1,t} + \alpha_2 x_{2,t} + \ldots + \alpha_q x_{q,t})]^2\right]\right\},
\end{aligned}$$

where subscript p means that it belongs to the prior distribution. This is as far one can go without specifying ϕ. But before that is done, please note that an alternative way to write the normal distribution is given by

$$\begin{aligned}
\Phi \sim \mathrm{N}(\mu_\phi, \sigma^2_\phi) \Rightarrow f(\phi) &= \frac{1}{\sqrt{2\pi}\sigma_\phi}\exp\left\{-\frac{1}{2\sigma^2_\phi}(\phi-\mu_\phi)^2\right\} \\
&= \frac{1}{\sqrt{2\pi}\sigma_\phi}\exp\left\{-\frac{1}{2\sigma^2_\phi}\left(\phi^2 - 2\phi\mu_\phi + \mu^2_\phi\right)\right\} \\
&\propto \exp\left\{a\phi^2 + b\phi + C\right\}
\end{aligned}$$

where $a = -1/(2\sigma^2_\phi)$, $b = \mu_\phi/\sigma^2_\phi$ and $C = \mu_\phi/(2\sigma^2_\phi)$ is a constant. The two equations above can now be used to find the specific conditional distributions.

A.7.1 Conditional Distributions for the α's

First the full conditional distributions for α_i, $i = (0, 1, 2)$:

$$\begin{aligned}
\pi(\alpha_0 | -\alpha_0) &\propto \exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma^2_{\alpha_{0,p}}}(\alpha_0 - \mu_{\alpha_{0,p}})^2 + \right.\right. \\
&\quad \tau\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}[R_PnL_t - \gamma_1(\beta_0 + \beta_1 dTrend_T + \beta_2 dTrendGwt_T) - \\
&\quad \left.\left. \gamma_2(\alpha_0 + \alpha_1 x_{1,t} + \alpha_2 x_{2,t})]^2\right]\right\} \\
&= \exp\left\{-\frac{\alpha_0^2}{2\sigma^2_{\alpha_{0,p}}} - \frac{\tau K_{N+1}\gamma_2^2\alpha_0^2}{2} + \frac{\alpha_0\tau}{2}\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} 2\gamma_2(R_PnL_t - \gamma_1\beta_0 - \right. \\
&\quad \left. \gamma_1\beta_1 dTrend_T - \gamma_1\beta_2 dTrendGwt_T - \gamma_2\alpha_1 x_{1,t} - \gamma_2\alpha_2 x_{2,t}) + C\right\} \\
&= \exp\left\{\left[-\frac{1}{2}\left(\frac{1}{\sigma^2_{\alpha_{0,p}}} + \gamma_2^2\tau K_{N+1}\right)\right]\alpha_0^2 + \left[\tau\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}\gamma_2(\cdots)\right]\alpha_0 + C\right\}
\end{aligned}$$

which assumes $q = 2$. $\mu_{\alpha_{0,p}} = 0$ from the prior distribution given in equation (A.1) and $\sigma^2_{\alpha_{0,p}}$ is the variance from the same prior. This means that

$$\frac{1}{\sigma^2_{\alpha_0}} = \frac{1}{\sigma^2_{\alpha_{0,p}}} + \tau\gamma_2^2 K_{N+1} \Rightarrow \sigma^2_{\alpha_0} = \left(\frac{1}{\sigma^2_{\alpha_{0,p}}} + \tau\gamma_2^2 K_{N+1}\right)^{-1}$$

and that

$$\frac{\mu_{\alpha_0}}{\sigma^2_{\alpha_0}} = \tau \sum_{T=1}^{N} \sum_{t=K_T+1}^{K_{T+1}} \gamma_2(\cdots) \Rightarrow \mu_{\alpha_0} = \left(\frac{1}{\sigma^2_{\alpha_{0,p}}} + \tau\gamma_2^2 K_{N+1} \right)^{-1} \tau \sum_{T=1}^{N} \sum_{t=K_T+1}^{K_{T+1}} \gamma_2(\cdots)$$

such that

$$\pi(\alpha_0|-\alpha_0) \sim \mathrm{N}\left(\left(\frac{1}{\sigma^2_{\alpha_{0,p}}} + \tau\gamma_2^2 K_{N+1} \right)^{-1} \tau \sum_{T=1}^{N} \sum_{t=K_T+1}^{K_{T+1}} \gamma_2(\cdots), \left(\frac{1}{\sigma^2_{\alpha_{0,p}}} + \tau\gamma_2^2 K_{N+1} \right)^{-1} \right),$$

where $(\cdots) = (R\_PnL_t - \gamma_1\beta_0 - \gamma_1\beta_1 dTrend_T - \gamma_1\beta_2 dTrendGwt_T - \gamma_2\alpha_1 x_{1,t} - \gamma_2\alpha_2 x_{2,t})$. Further, if one chooses the variance in the prior distribution sufficiently large, $\sigma^2_{\alpha_{0,p}} \to \infty$, one have that

$$\pi(\alpha_0| - \alpha_0) \sim \mathrm{N}\left(\frac{1}{\gamma_2^2 K_{N+1}} \sum_{T=1}^{N} \sum_{t=K_T+1}^{K_{T+1}} \gamma_2(\cdots), \frac{1}{\tau\gamma_2^2 K_{N+1}} \right).$$

For $\pi(\alpha_1| - \alpha_1)$ the calculations are as follows:

$$\begin{aligned}
\pi(\alpha_1| - \alpha_1) \quad &\propto \quad \exp\left\{ -\frac{1}{2} \left[\frac{1}{\sigma^2_{\alpha_{1,p}}} (\alpha_1 - \mu_{\alpha_{1,p}})^2 + \right.\right. \\
&\quad \tau \sum_{T=1}^{N} \sum_{t=K_T+1}^{K_{T+1}} [R_PnL_t - \gamma_1(\beta_0 + \beta_1 dTrend_T + \beta_2 dTrendGwt_T) - \\
&\quad \left.\left. \gamma_2(\alpha_0 + \alpha_1 x_{1,t} + \alpha_2 x_{2,t})]^2 \right] \right\} \\
&= \quad \exp\left\{ -\frac{\alpha_1^2}{2\sigma^2_{\alpha_{1,p}}} - \frac{\tau\gamma_2^2\alpha_1^2 \sum_{t=K_1+1}^{K_{N+1}} x_{1,t}^2}{2} + \frac{\alpha_1\tau}{2} \sum_{T=1}^{N} \sum_{t=K_T+1}^{K_{T+1}} 2\gamma_2 x_{1,t}(R_PnL_t - \right. \\
&\quad \left. \gamma_1\beta_0 - \gamma_1\beta_1 dTrend_T - \gamma_1\beta_2 dTrendGwt_T - \gamma_2\alpha_0 - \gamma_2\alpha_2 x_{2,t}) + C \right\} \\
&= \quad \exp\left\{ \left[-\frac{1}{2} \left(\frac{1}{\sigma^2_{\alpha_{1,p}}} + \tau\gamma_2^2 \sum_{t=K_1+1}^{K_{N+1}} x_{1,t}^2 \right) \right] \alpha_1^2 + \right. \\
&\quad \left. \left[\tau \sum_{T=1}^{N} \sum_{t=K_T+1}^{K_{T+1}} \gamma_2 x_{1,t}(\cdots) \right] \alpha_1 + C \right\}
\end{aligned}$$

which gives, with the same assumptions as in the case with α_0 above,

$$\pi(\alpha_1| - \alpha_1) \sim \mathrm{N}\left(\frac{\sum_{T=1}^{N} \sum_{t=K_T+1}^{K_{T+1}} \gamma_2 x_{1,t}(\cdots)}{\gamma_2^2 \sum_{t=K_1+1}^{K_{N+1}} x_{1,t}^2}, \frac{1}{\tau\gamma_2^2 \sum_{t=K_1+1}^{K_{N+1}} x_{1,t}^2} \right).$$

An identical approach gives the conditional distribution $\pi(\alpha_2| - \alpha_2)$:

$$\pi(\alpha_2| - \alpha_2) \sim \mathrm{N}\left(\frac{\sum_{T=1}^{N} \sum_{t=K_T+1}^{K_{T+1}} \gamma_2 x_{2,t}(\cdots)}{\gamma_2^2 \sum_{t=K_1+1}^{K_{N+1}} x_{2,t}^2}, \frac{1}{\tau\gamma_2^2 \sum_{t=K_1+1}^{K_{N+1}} x_{2,t}^2} \right).$$

A.7.2 Conditional Distributions for the β's

The calculations are quite similar for the β's:

$$\begin{aligned}
\pi(\beta_0|-\beta_0) \quad &\propto \quad \exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma^2_{\beta_{0,p}}}(\beta_0-\mu_{\beta_{0,p}})^2+\right.\right.\\
&\tau\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}[R_PnL_t-\gamma_1(\beta_0+\beta_1 dTrend_T+\beta_2 dTrendGwt_T)-\\
&\left.\left.\gamma_2(\alpha_0+\alpha_1x_{1,t}+\alpha_2x_{2,t})]^2\right]\right\}\\
&= \quad \exp\left\{-\frac{\beta_0^2}{2\sigma^2_{\beta_{0,p}}}-\frac{\tau K_{N+1}\gamma_1^2\beta_0^2}{2}+\frac{\beta_0\tau}{2}\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}2\gamma_1(R_PnL_t-\gamma_1\beta_1 dTrend_T-\right.\\
&\gamma_1\beta_2 dTrendGwt_T-\gamma_2\alpha_0-\gamma_2\alpha_1x_{1,t}-\gamma_2\alpha_2x_{2,t})+C\}\\
&= \quad \exp\left\{\left[-\frac{1}{2}\left(\frac{1}{\sigma^2_{\beta_{0,p}}}+\tau\gamma_1^2K_{N+1}\right)\right]\beta_0^2+\left[\tau\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}\gamma_1(\cdots)\right]\beta_0+C\right\};
\end{aligned}$$

$$\begin{aligned}
\pi(\beta_1|-\beta_1) \quad &\propto \quad \exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma^2_{\beta_{1,p}}}(\beta_1-\mu_{\beta_{1,p}})^2+\right.\right.\\
&\tau\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}[R_PnL_t-\gamma_1(\beta_0+\beta_1 dTrend_T+\beta_2 dTrendGwt_T)-\\
&\left.\left.\gamma_2(\alpha_0+\alpha_1x_{1,t}+\alpha_2x_{2,t})]^2\right]\right\}\\
&= \quad \exp\left\{-\frac{\beta_1^2}{2\sigma^2_{\beta_{1,p}}}-\frac{\tau\gamma_1^2\beta_1^2\sum_{T=1}^{N}\sum_{t=K_1+1}^{K_{N+1}}dTrend_T^2}{2}+\right.\\
&\frac{\beta_1\tau}{2}\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}2\gamma_1 dTrend_T(R_PnL_t-\\
&\gamma_1\beta_0-\gamma_1\beta_2 dTrendGwt_T-\gamma_2\alpha_0-\gamma_2\alpha_1x_{1,t}-\gamma_2\alpha_2x_{2,t})+C\}\\
&= \quad \exp\left\{\left[-\frac{1}{2}\left(\frac{1}{\sigma^2_{\beta_{1,p}}}+\tau\gamma_1^2\sum_{T=1}^{N}\sum_{t=K_1+1}^{K_{N+1}}dTrend_T^2\right)\right]\beta_1^2+\right.\\
&\left.\left[\tau\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}\gamma_1 dTrend_T(\cdots)\right]\beta_1+C\right\};
\end{aligned}$$

$$\begin{aligned}
\pi(\beta_2|-\beta_2) \quad &\propto \quad \exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma^2_{\beta_{2,p}}}(\beta_2-\mu_{\beta_{2,p}})^2+\right.\right.\\
&\tau\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}[R_PnL_t-\gamma_1(\beta_0+\beta_1 dTrend_T+\beta_2 dTrendGwt_T)-\\
&\left.\left.\gamma_2(\alpha_0+\alpha_1x_{1,t}+\alpha_2x_{2,t})]^2\right]\right\}
\end{aligned}$$

$$
\begin{aligned}
\pi(\beta_2|-\beta_2) \quad &\propto \quad \exp\left\{-\frac{\beta_2^2}{2\sigma^2_{\beta_{2,p}}} - \frac{\tau\gamma_1^2\beta_2^2\sum_{T=1}^{N}\sum_{t=K_1+1}^{K_{N+1}} dTrendGwt_T^2}{2} + \right.\\
&\quad \frac{\beta_2\tau}{2}\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} 2\gamma_1 dTrendGwt_T(R\_PnL_t - \\
&\quad \left.\gamma_1\beta_0 - \gamma_1\beta_1 dTrend_T - \gamma_2\alpha_0 - \gamma_2\alpha_1 x_{1,t} - \gamma_2\alpha_2 x_{2,t}) + C\right\}\\
&= \quad \exp\left\{\left[-\frac{1}{2}\left(\frac{1}{\sigma^2_{\beta_{2,p}}} + \tau\gamma_1^2\sum_{T=1}^{N}\sum_{t=K_1+1}^{K_{N+1}} dTrendGwt_T^2\right)\right]\beta_2^2 + \right.\\
&\quad \left.\left[\tau\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}\gamma_1 dTrendGwt_T(\cdots)\right]\beta_2 + C\right\}.
\end{aligned}
$$

This result in the following normal distributions:

$$
\begin{aligned}
\pi(\beta_0|-\beta_0) \quad &\sim \quad \mathrm{N}\left(\frac{1}{\gamma_1^2 K_{N+1}}\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}\gamma_1(\cdots), \frac{1}{\tau\gamma_1^2 K_{N+1}}\right);\\
\pi(\beta_1|-\beta_1) \quad &\sim \quad \mathrm{N}\left(\frac{\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}\gamma_1 dTrend_T(\cdots)}{\gamma_1^2\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} dTrend_T^2}, \frac{1}{\tau\gamma_1^2\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} dTrend_T^2}\right);\\
\pi(\beta_2|-\beta_2) \quad &\sim \quad \mathrm{N}\left(\frac{\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}\gamma_1 dTrendGwt_T(\cdots)}{\gamma_1^2\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} dTrendGwt_T^2}, \frac{1}{\tau\gamma_1^2\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} dTrendGwt_T^2}\right).
\end{aligned}
$$

A.7.3 Conditional Distributions for the γ's

The γ's follows the same pattern;

$$
\begin{aligned}
\pi(\gamma_1|-\gamma_1) \quad &\propto \quad \exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma^2_{\gamma_{1,p}}}(\gamma_1-\mu_{\gamma_{1,p}})^2 + \tau\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}[R\_PnL_t - \gamma_1\mu_T - \gamma_2\mu_t]^2\right]\right\}\\
&= \quad \exp\left\{-\frac{\gamma_1^2}{2\sigma^2_{\gamma_{1,p}}} - \frac{\tau\gamma_1^2}{2}\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}\mu_T^2 + \frac{\tau\gamma_1}{2}\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}} 2\mu_T(R\_PnL_t - \gamma_2\mu_t) + C\right\}\\
&= \quad \exp\left\{\left[-\frac{1}{2}\left(\frac{1}{\sigma^2_{\gamma_{1,p}}} + \tau\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}\mu_T^2\right)\right]\gamma_1^2 + \right.\\
&\quad \left.\left[\tau\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}\mu_T(R\_PnL_t - \gamma_2\mu_t)\right]\gamma_1 + C\right\}.\\
\pi(\gamma_1|-\gamma_1) \quad &\sim \quad \mathrm{N}\left(\frac{\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}\mu_T(R\_PnL_t - \gamma_2\mu_t)}{\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}\mu_T^2}, \frac{1}{\tau\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}\mu_T^2}\right)
\end{aligned}
$$

$$\begin{aligned}
\pi(\gamma_2|-\gamma_2) &\propto \exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma^2_{\gamma_{2,p}}}(\gamma_2-\mu_{\gamma_{2,p}})^2+\tau\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}[R_PnL_t-\gamma_1\mu_T-\gamma_2\mu_t]^2\right]\right\} \\
&= \exp\left\{-\frac{\gamma_2^2}{2\sigma^2_{\gamma_{2,p}}}-\frac{\tau\gamma_2^2}{2}\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}\mu_t^2+\frac{\tau\gamma_2}{2}\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}2\mu_t(R_PnL_t-\gamma_1\mu_T)+C\right\} \\
&= \exp\left\{\left[-\frac{1}{2}\left(\frac{1}{\sigma^2_{\gamma_{2,p}}}+\tau\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}\mu_t^2\right)\right]\gamma_2^2+\right. \\
&\quad \left.\left[\tau\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}\mu_T(R_PnL_t-\gamma_1\mu_T)\right]\gamma_2+C\right\} \\
\pi(\gamma_2|-\gamma_2) &\sim \mathrm{N}\left(\frac{\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}\mu_t(R_PnL_t-\gamma_1\mu_T)}{\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}\mu_t^2},\frac{1}{\tau\sum_{T=1}^{N}\sum_{t=K_T+1}^{K_{T+1}}\mu_t^2}\right).
\end{aligned}$$

A.7.4 Conditional Distributions for τ

τ is a bit different. Although the procedure is the same, τ has a gamma prior distribution with shape and scale parameters a and s, respectively:

$$\begin{aligned}
\pi(\tau|-\tau) &\propto \pi(\tau)\prod_{T=1}^{N}\prod_{t=K_T+1}^{K_{T+1}}\pi(R_PnL_t|\alpha_0,\alpha_1,\ldots,\alpha_q,\beta_0,\beta_1,\beta_2,\gamma_1,\gamma_2,\tau) \\
&= \frac{1}{s^a\Gamma(a)}\tau^{a-1}\exp\left\{-\frac{\tau}{s}\right\}\prod_{T=1}^{N}\prod_{t=K_T+1}^{K_{T+1}}\sqrt{\frac{\tau}{2\pi}}\times \\
&\quad \exp\left\{-\frac{\tau}{2}\left[R_PnL_t-\gamma_1(\beta_0+\beta_1 dTrend_T+\beta_2 dTrendGwt_T)-\right.\right. \\
&\quad \left.\left.\gamma_2(\alpha_0+\alpha_1x_{1,t}+\alpha_2x_{2,t}+\ldots+\alpha_qx_{q,t})\right]^2\right\} \\
&= \frac{1}{s^a\Gamma(a)}\tau^{a-1}\exp\left\{-\frac{\tau}{s}\right\}\left(\frac{\tau}{2\pi}\right)^{\left(\frac{K_{N+1}}{2}\right)}\prod_{t=K_1+1}^{K_2}\exp\left\{-\frac{\tau}{2}\left[R_PnL_t-\right.\right. \\
&\quad \gamma_1(\beta_0+\beta_1 dTrend_{T=1}+\beta_2 dTrendGwt_{T=1})- \\
&\quad \left.\left.\gamma_2(\alpha_0+\alpha_1x_{1,t}+\alpha_2x_{2,t}+\ldots+\alpha_qx_{q,t})\right]^2\right\}\prod_{t=K_2+1}^{K_3}\exp\left\{-\frac{\tau}{2}\left[R_PnL_t-\right.\right. \\
&\quad \gamma_1(\beta_0+\beta_1 dTrend_{T=2}+\beta_2 dTrendGwt_{T=2})- \\
&\quad \left.\left.\gamma_2(\alpha_0+\alpha_1x_{1,t}+\alpha_2x_{2,t}+\ldots+\alpha_qx_{q,t})\right]^2\right\}\times\cdots\times \\
&\quad \prod_{t=K_N+1}^{K_{N+1}}\exp\left\{-\frac{\tau}{2}\left[R_PnL_t-\gamma_1(\beta_0+\beta_1 dTrend_{T=N}+\beta_2 dTrendGwt_{T=N})-\right.\right. \\
&\quad \left.\left.\gamma_2(\alpha_0+\alpha_1x_{1,t}+\alpha_2x_{2,t}+\ldots+\alpha_qx_{q,t})\right]^2\right\}
\end{aligned}$$

$$
\begin{aligned}
\pi(\tau|-\tau) \quad &\propto \quad (2\pi)^{-\left(\frac{K_{N+1}}{2}\right)} \frac{1}{s^a \Gamma(a)} \tau^{(a-1)} \tau^{\left(\frac{K_{N+1}}{2}\right)} \exp\left\{-\frac{\tau}{s}\right\} \times \\
&\quad \exp\left\{ \sum_{t=K_1+1}^{K_2} -\frac{\tau}{2}[R\_PnL_t - \gamma_1(\beta_0 + \beta_1 dTrend_{T=1} + \beta_2 dTrendGwt_{T=1}) - \right. \\
&\quad \left. \gamma_2(\alpha_0 + \alpha_1 x_{1,t} + \alpha_2 x_{2,t} + \ldots + \alpha_q x_{q,t})]^2 \right\} \times \cdots \times \\
&\quad \exp\left\{ \sum_{t=K_N+1}^{K_{N+1}} -\frac{\tau}{2}[R\_PnL_t - \gamma_1(\beta_0 + \beta_1 dTrend_{T=N} + \beta_2 dTrendGwt_{T=N}) - \right. \\
&\quad \left. \gamma_2(\alpha_0 + \alpha_1 x_{1,t} + \alpha_2 x_{2,t} + \ldots + \alpha_q x_{q,t})]^2 \right\} \\
&\propto \quad \tau^{\left(a-1+\frac{K_{N+1}}{2}\right)} \exp\left\{-\frac{\tau}{s}\right\} \exp\left\{-\frac{\tau}{2} \times \right. \\
&\quad \sum_{T=1}^{N} \sum_{t=K_T+1}^{K_{T+1}} [R\_PnL_t - \gamma_1(\beta_0 + \beta_1 dTrend_T + \beta_2 dTrendGwt_T) - \\
&\quad \left. \gamma_2(\alpha_0 + \alpha_1 x_{1,t} + \alpha_2 x_{2,t} + \ldots + \alpha_q x_{q,t})]^2 \right\} \\
&= \quad \tau^{\left(a-1+\frac{K_{N+1}}{2}\right)} \exp\left\{ -\tau \left(\frac{1}{s} + \frac{1}{2} \sum_{T=1}^{N} \sum_{t=K_T+1}^{K_{T+1}} [R\_PnL_t - \mu_{Tt}]^2 \right) \right\} \\
\pi(\tau|-\tau) \quad &\propto \quad \text{Gamma}\left(a^* = a + \frac{K_{N+1}}{2}, \frac{1}{s^*} = \frac{1}{s} + \frac{\sum_{T=1}^{N} \sum_{t=K_T+1}^{K_{T+1}} (PnL\_10DV_t - \mu_S)^2}{2} \right).
\end{aligned}
$$