Time-Varying Exposure in Hedge Funds

Thesis for Master of Banking and Finance (MBF)
HEC Lausanne

October 2005

Author: Florian Albrecht[1]

[1]Florian Albrecht, Master of Banking and Finance (MBF), HEC Lausanne. I would like to thank my supervisor Prof. Eric Jondeau, HEC Lausanne. This thesis was written during an internship at RMF in Pfäffikon/SZ, Switzerland. I am grateful to Dr. Jean-François Bacmann of RMF for his guidance, support and insightful comments. The views expressed in this thesis are those of the author and do not necessarily reflect those of RMF.

Abstract

Despite the frequent definition of hedge funds as absolute return investment vehicles, research has shown that they are exposed to traditional as well as alternative risk factors. Option-like components in hedge fund models indicate that hedge fund exposure is not static, but time-varying. Intuitively, this can be explained by assuming that fund managers change portfolio-weights using their market-timing skills.

Since it is known that static regression methods are limited in describing hedge fund returns and can therefore misleadingly attribute parts of the return from market exposure to manager skill or vice versa, we examine the hypothesis that modelling hedge fund returns with models that allow time-variation in exposure will lead to more accurate results and better forecasts of exposure. We use various models and recursive filtering for estimation.

Results show that in the case of long/short equity hedge funds indices these extended models perform indeed better than ordinary regressions, both in-sample as well as out-of-sample. In addition, there is evidence that part of the time-variation comes from active alteration of portfolio-weights by the manager, indicating that besides asset selection, market-timing is an important part of hedge fund managers' skills.

Keywords: Hedge fund, time-varying exposure, beta forecast, Kalman filter

JEL Classification: C32, G11, G12

Executive Summary

Over the past years, hedge fund research has largely rejected the original assertion that hedge funds are uncorrelated to market risk. Consequently, the opinion that hedge funds offer positive absolute returns (also called alpha) irrespective of market conditions cannot be maintained. In addition, it seems that the low exposure (also called beta) of hedge funds to market risk measured with traditional regression methods is just an average value of a dynamic, time-varying exposure of the portfolios. Various multivariate regression models have been proposed in order to better describe hedge fund returns, and the fact that they include option-like components as risk factors is an additional indication for time-variation in the exposure of hedge funds.

From a top view, there are two explanations for this time-variation: First, hedge fund managers are allowed to change positions, shortsell assets and use leverage to adapt their portfolios to changing market environments, consequently they adjust portfolio weights and exposure. Second, market premiums or betas of the underlying assets are time-variant and cause the change in portfolio exposure, although portfolio weights themselves remain constant. Both effects may be present in parallel.

If there is time-varying exposure in hedge fund returns, we would expect dynamic estimation models to perform better than static ones. If this is not the case, we can conclude that exposure is indeed static and investors could get the same exposure by investing into a mutual or index-tracking fund while paying lower fees. If on the other hand the dynamic models find time-varying exposure, it is important to know whether this is due to the first or second of above-mentioned explications. Since the first represents a manager skill, it justifies higher fees, while the second signals no additional value generated by the manager.

We will assume that managers do actively manage the exposure of their portfolios, and it is therefore our hypothesis that models that are built to capture this behaviour will give equal or better results than other models. For the verification of the hypothesis, we select different state space models which allow for time-varying alphas and betas:

- Random Coefficient Model (RCM)
- Random Walk Model (RWM)
- Mean Reverting Model (MRM)
- Model by Mamaysky et al. (MSZM)

The latter is presented in Mamaysky, Spiegel, and Zhang (2003) and assumes explicitely that the change in portfolio exposure comes from an active adjustment of the portfolio weights by the manager, whereas RCM, RWM and MRM assume that time-varying betas of the underlying assets create a change in exposure of the portfolio. The models are estimated by using Kalman filter[1]and the estimations are then rated in terms of their ability to forecast alpha and beta. As benchmark, we use the forecasts of a moving window OLS regression.

A univariate regression of HFRI and CSFB/Tremont long/short equity indices from 2001 to 2005 shows that there is no unambiguous winning model. However, except RCM all dynamic models provide better forecasts than traditional moving window regression. MSZM, which is the most complex model, provides on average also the best results. Also MRM and even the relatively simple RWM outperform OLS in most cases. The good performance of MSZM hints that first exposure is dynamic and second the dynamics comes from an active management of exposure by the fund manager. However, for the average of funds and the analysed time period, the model parameter measuring timing skills is negative, indicating an average negative timing skill which might stress the importance of careful assessment of individual funds.

With respect to exposure, we find a strong reduction after the crash in year 2000 when analysing long/short equity style against S&P 500. Only after 2002, hedge fund exposure was increased back again. The exposure of equity market neutral style to S&P 500 on the other hand is found to be close to zero with all models.

Finally, forecasting exposure with the multivariate four factor model introduced by Carhart (1997) produces smaller forecast errors for all models[2]. In some cases, dynamic models perform still better than OLS, but differences are becoming small. This suggests that the advantage of time-varying models decrease with the number of regressors included.

Summarised we can thus conclude that time-varying models provide better results than simple OLS in forecasting exposure of long/short equity hedge fund indices and

[1]The nonlinear MSZM requires the application of Extended or Unscented Kalman filter.

[2]Including moving window regression. Because of the big number of parameters to be estimated, MSZM is not computed in the multivariate case.

offer more insight into the dynamic evolution of exposure and thus the behaviour of hedge fund managers. Moreover, results indicate that time-variation is at least partly caused by an active changement of portfolio weights by the manager. This confirms that the value generated by hedge fund managers comes not only from their asset selection skills, but also from their dynamic management of the portfolio, and that time-varying models can help to correctly identify these skills.

Contents

1 Introduction

Hedge funds are frequently defined as investment vehicles that offer positive absolute returns. This signifies that due to the skills of their managers, they should deliver positive returns (net of fees) no matter whether markets are rising or falling. In contrast, mutual or index tracking funds try to achieve positive relative returns by beating their benchmark-index. The part of fund returns which is contributed by absolute returns is called alpha, while the part contributed by the exposure to market risk is called beta. While mutual funds are thus highly correlated with market risk, hedge funds should in theory be uncorrelated with traditional and/or alternative market risks. Practical evidence however shows that this is not the case and hedge funds are exposed to market risks to a certain extent.

Nevertheless, standard regression models which work well for mutual fund returns have low explanatory power for hedge fund returns. The desire to better describe hedge fund behaviour and to increase the significance of the regression models has led to the introduction of extended models that include alternative risk factors with option-like payoffs. The importance of these option-like risk factors and further hedge fund research suggest that exposure of hedge funds to risk factors is not static, but dynamically evolving through time. This is presumably also the reason for the poor performance of traditional models with respect to hedge funds.

The most intuitive and straightforward explication for time-varying exposure of hedge funds lies in their nature: Hedge fund managers are allowed to quickly change exposures, take short positions and use leverage to adapt their portfolios to changing market environments, hence portfolio weights are actively changed. Another explanation approach is that while portfolio weights are kept constant, the change in portfolio exposure is caused by time-variant market premiums or betas of the underlying assets. Eventually, both effects might be present simultaneously. For these reasons, it can be expected that extended models that take into account time-variation of portfolio alpha and beta have an increased significance compared to standard regression methods and are thus better able to describe the returns of hedge funds. In parallel, the time-varying exposures obtained with these models will provide new insights about the activities and skills of hedge fund managers.

If the exposure to a certain risk factor is found to be static, we can conclude that there is no dynamic management of the hedge fund exposure with regard to this risk factor, thus investors could get the same exposure by investing into a mutual or index-tracking fund, but at lower fees. If the exposure is time-varying, investors would like to know whether this is because of an active change of the portfolio weights by the manager or just because of the time-variation in the assets of the portfolio. In the first case, investors could assume that the manager is using some market-timing skills to adapt the portfolio and this would justify the charge of fees. In the second case however, there is no action by the manager, hence investors should pay no fees. In a nutshell, a dynamic portfolio beta could either represent a timing skill of the fund manager or not.

Although the analysis of time-varying exposure is thus an important part of the performance analysis of hedge funds, it has not been explored in depth yet. In particular the application of extended models and filtering methods may provide new insights (see Géhin and Vaissié (2005)). In this thesis, we therefore analyse the existence of time-varying exposure in long/short equity hedge funds. Taking static regression methods as benchmark, we use models that allow for time-variation in portfolio alpha and beta. While some of the models assume that the development of the underlying stocks is time-varying, one model assumes a dynamic change of fund-exposure by a change in the portfolio-weights. Assuming that fund managers are generating added value by active management of the portfolio weights, we expect the latter model to perform best or at least as good as the others. Since we are interested to rate the models by their power to forecast alphas and betas, we use out of sample rating criterias instead of in sample comparisons for the valuation of the models.

The rest of this thesis is organised as follows: in chapter 2, we give a literature overview of models proposed to analyse and describe hedge fund returns with a special focus on literature considering time-varying exposure. The time-varying models used for the tests are presented in chapter 3. Since the estimation of these models requires the use of Kalman filter, chapter 4 introduces Kalman filter and the implementation of the models within its framework. Tests and results are shown in chapter 5, and the conclusions of chapter 6 complete the thesis.

2 Literature Review

While theory claims that hedge funds provide absolute returns and are uncorrelated to markets, practical evidence confirms this only in parts. Using the asset class factor model for mutual funds by Sharpe (1992), Fung and Hsieh (1997) find that mutual funds are highly correlated to standard asset classes. But although lower than in mutual funds, correlation does still exist between hedge fund returns and returns of standard asset classes. In the case of long/short equity hedge funds, Fung and Hsieh (2003) show that there is mainly an exposure to equity market (proxied by S&P 500) and the spread between small cap and large cap stock (proxied by SMB). Studies by Agarwal and Naik (2004) and Capocci (2002) confirm that hedge funds are significantly exposed to standard asset classes, although the exposure is small compared to the one found with mutual funds. In a recent study, Géhin and Vaissié (2005) evaluate the exposure of various hedge fund styles to systematic risk factors[1]. Results show that most styles are exposed to traditional market risk.

One possible explanation for low correlation could be that hedge fund returns react asynchronously to standard asset class returns. Krail, Asness, and Liew (2001) regress the returns of CSFB/Tremont hedge fund indices against S&P 500. When only the actual S&P 500 returns are used as regressor, the authors confirm the finding of low correlation. Arguing that rarely traded assets or result smoothing by the manager might lead to correlations with lagged S&P 500 returns, they run a multivariate regression containing also lagged returns as regressors. It turns out that correlation increases significantly with the latter method. The authors conclude that hedge funds are much stronger correlated to traditional market risks than is assumed in general. Their results are however questioned by Fung and Hsieh (2004a) who claim that the lagged dependency could be due to the choice of benchmark.

Fung and Hsieh (1997) explain the limited explanatory power of standard asset classes with respect to hedge funds in another way: Contrary to mutual fund managers who

[1] Results are computed by calculating time-varying exposure using Kalman smoother and then calculating the average value.

have relative return targets, hedge fund managers have absolute return targets. While mutual fund managers follow generally buy-and-hold strategies with limited leverage and the only decision being where to invest, hedge fund managers can choose not only the location, but also the trading strategy of their investments. This leads to nonlinear option-like exposures of hedge funds to standard asset classes. In order to increase the explanatory power, the authors extend the model of Sharpe (1992) by three dynamic trading strategies, showing that R^2 of hedge funds is significantly increased with this model. Combined with nine buy-and-hold styles in standard asset classes, they end up with 12 investment styles[2]which build a framework for style analysis of hedge funds as well as for mutual funds.

Starting from the findings of Fung and Hsieh (1997), Agarwal and Naik (2004) introduce a general asset class model containing buy-and-hold strategies (location factors) and passive option-based strategies (trading strategy factors). They find that the option-based strategies can explain a significant proportion of variation in hedge fund returns, resulting in a higher R^2 than with the model of Sharpe (1992). Capocci (2002) compares several models, including the four-factor model by Carhart (1997) and an extension of the model by Agarwal and Naik (2004). The latter model proves to perform best in explaining the variation of hedge fund returns. The results of Agarwal and Naik (2004) and Capocci (2002) indicate that besides exposure to traditional asset classes, which is measured by traditional beta, hedge funds are exposed to additional alternative risk factors. The exposure to this alternative risk factors is measured by alternative beta. The results of Géhin and Vaissié (2005) confirm these findings. The authors observe that certain hedge fund styles have significant exposures to alternative risk factors such as for example change in credit spread.

Another reason why static regression models do not perform well in describing hedge fund returns may be the inability of these models to measure time-varying exposure. Brealey and Kaplanis (2001) investigate in the question whether factor exposures of hedge funds are constant. Using a moving window regression, they find that the null hypothesis of stable coefficients is strongly rejected which indicates that factor exposures of hedge funds are indeed time-varying. McGuire, Remolona, and Tsatsaronis (2005) analyse the time-varying exposure of different hedge fund investment styles (directional, market-

[2]Buy-and-hold asset classes are: MSCI US equities, MSCI non-US equities, IFC emerging market equities, JP Morgan US government bonds, JP Morgan non-US government bonds, 1-month eurodollar deposit, price of gold, Federal Reserve's Trade Weighted Dollar Index and high yield corporate bonds. Dynamic trading strategies are: Systems/Trend Following, Systems/Opportunistic and Global/Macro style.

neutral and equity-focused) to various risk factors using moving window regression. They show that despite the homogeneity of hedge fund strategies, the exposure of the analysed strategies to some common risk factors such as options on S&P 500 index or to the Fama-French SMB factor is changing over time, although similar between the strategies. Exposure to other market risk factors like fixed income is found to be homogeneous between strategies. Unfortunately, both above-mentioned authors give no indication about whether R^2 increases when time-varying exposure is accounted for.

For the analysis of time-varying exposure, simple regression methods may however not be the best tool. Kat and Miffre (2003) estimate abnormal performance by using a conditional asset pricing model first introduced by Ferson and Schadt (1996). In order to generate conditionality, the model uses lagged information variables[3]. The model is tested with three different sets of regressors: The market portfolio, the market portfolio plus SMB and HML, and finally the market portfolio plus five macroeconomic variables. The authors find that conditioning risk and performance measures onto past information results in higher abnormal performance as measured by alpha, which implies that static models underestimate abnormal performance. They find also evidence that risk exposures are time-varying.

Given the predominant evidence in favour of time-varying exposure in literature, one may ask for the reason of this characteristic. The beta of a portfolio changes when either the underlying asset betas change and/or when the portfolio weights are changed. For hedge funds containing mainly stocks (long/short equity and equity market neutral), the presence of the first characteristic can be tested by looking at time-variation in stock betas. Research by Wells (1996) on Swedish stocks and Yao and Gao (2004) on Australian industry portfolios confirm that betas of stocks and stock portfolios are time-varying. Both authors use dynamic models and recursive filtering techniques.

The presence of the second characteristic implies an active fund management. Fung and Hsieh (1997) argue that hedge fund managers reduce the correlation to asset class

[3] The model is defined as follows:

$$r_t = \alpha_0 + \alpha_1 z_{t-1} + \beta_0 r_{m,t} + \beta_1 r_{m,t} z_{t-1} + \epsilon_t$$

α_0 is the conditional version of Jensen's measure of abnormal performance. The lagged information variable z_{t-1} has mean zero, $\alpha_1 z_{t-1}$ measures the time variation in abnormal performance and $\beta_1 z_{t-1}$ the time variation in measures of risk. If there is no information, the equation reduces to the static unconditional one. In the multivariate case, there are k regressors $r_{m,t}$ and l information variables z_{t-1}.

The lagged information variables are: dividend yield on Datastream world equity portfolio, a proxy for the global risk of default on money market instruments, the US term structure of interest rates and the US one-month Treasury bill.

returns by adjusting the portfolio weights and hereby actively changing the exposure of their funds to asset classes. Mamaysky, Spiegel, and Zhang (2003) apply a model that assumes constant asset betas, but time-varying portfolio weights to analyse the dynamics of mutual funds alphas and betas (see also 3.7)[4]. The authors use a generalised version of the model proposed by Ferson and Schadt (1996). Contrary to the latter, Mamaysky, Spiegel, and Zhang (2003) assume that fund managers assign the weights of individual assets within a portfolio according to a lagged information variable which is latent and has no economic explanation. Based on this, they develop a model of the evolution of the portfolio's alpha and beta that requires no knowledge of the alphas and betas of the individual underlying stocks and their weights within the portfolio. The fund's dynamic alpha and beta are estimated using an Extended Kalman Filter in order to identify funds with substantially positive alphas. The authors show that using the model to select successful funds and to build a portfolio out of them leads to results that beat the market benchmark. A comparison with the Ferson-Schadt model shows strong advantages of the Mamaysky model in the univariate case which decrease in a multivariate comparison.

Considering the high fees of hedge funds, investors should not pay fees for static exposure which they can get cheaper by investing in mutual or index-tracking funds. It is therefore essential to know whether hedge funds still deliver absolute returns once the exposure to traditional and alternative risk factors has been neutralised. A major part of earlier studies found evidence that this is the case. Krail, Asness, and Liew (2001) find significantly positive absolute returns on average when only the actual S&P 500 returns are used as regressor. As soon as lagged returns are added as regressors, absolute returns get close to zero or become even negative. Results by Kat and Miffre (2003) indicate that hedge fund managers are able to generate persistent abnormal performance on average, although hedge funds are found to be counter-cyclical, hence more attractive during economic downturns. These results are similar to those of Capocci (2002) who finds persistence in hedge fund performance for several strategies, although performance seems to be unstable over time. Fung and Hsieh (2004a) extract alternative alphas from diversified hedge fund portfolios after hedging away the exposure to S&P 500 and SMB. They show that these alternative alphas exhibit almost no sensitivity neither to the eight traditional asset-class indices[5]nor to the seven hedge fund risk factors, even under extreme market conditions, and are thus portable. The authors conclude that long/short

[4] While the authors use the model in the context of mutual funds, it is also applicable in the hedge fund context.

equity hedge funds show significant absolute returns after taking into account alternative and traditional market risk and that these absolute returns are not only an effect of bull markets.

By using time-variant models and traditional as well as alternative risk factors, Géhin and Vaissié (2005) disentangle hedge fund returns into the following three components : Pure alpha, which describes the security selection skills of the manager, static beta which describes the static exposures of the funds to traditional and alternative risk factors, and dynamic betas which come from the factor timing of the manager. Total alpha, i.e. the manager's skill, is then the sum of pure alpha and dynamic betas. The authors then estimate the contribution of these components by applying a dynamic model with Kalman filtering (although they do not disclose the exact model they use). They find that while on average about half of the variability in returns comes from alpha (with 25% from pure alpha and 24% from dynamic beta), the contribution of static beta to performance is more than 99% on average, with a positive pure alpha of 4% and a negative dynamic beta of -3%. Despite these average values, in most cases the contribution of dynamic betas to total alpha is much more pronounced than the contribution of pure alpha, which indicates that total alpha is rather driven by factor timing than the selection skills and underlines the importance of correctly identifying time-varying betas. This argument however neglects the possibility that dynamic betas might not come from managers' timing skills, but be a result of dynamic betas of the underlying assets.

Taking the above decomposition of alpha and beta, Géhin and Vaissié (2005) conclude that hedge funds cannot simply be defined as absolute return vehicles that always deliver positive results without exposure to risks. They argue that exposures to traditional and alternative risk factors are undervalued compared to pure alpha although they are responsible for an overwhelming part of the hedge fund returns. Consequently, in the authors view hedge funds should be considered as an asset class which in combination with traditional investments offers beta diversification. This oppinion is shared by Jaeger and Wagner (2005). They point out that as it is difficult to decompose hedge fund returns into alpha and beta and as there is no model for describing alpha directly, alpha is the remaining average part of the return when all beta contributions (traditional or alternative) are subtracted. From this point of view, any unaccounted beta will

[5]Standard conventional asset-class indices are US equities, non-US equities, emerging market equities, US bonds, non-US bonds, gold, trade-weighted dollar index and 1-month Eurodollar deposit rate. Hedge fund risk factors (also called asset-based style factors or ABS factors) are defined in Fung and Hsieh (2004b).

erroneously be attributed to alpha. Under- respectively overestimation of beta will lead to an under- respectively overestimation of alpha. The authors conclude that given the described decomposition difficulties it is hard to verify whether hedge funds really deliver absolute returns. Moreover they claim that estimations of beta are more accurate than those of alpha. In their view, investors should start to recognize the diversifying opportunities offered by alternative betas rather than just focus on absolute returns.

3 Modelling Time-Varying Exposure

3.1 Introduction

Hedge fund returns can be expressed as the sum of the exposure to the market, which is measured by beta, and the abnormal returns, which are defined by alpha. Alpha characterises the manager's selection skill. Now a portfolio may exhibit time-varying alpha and beta for three reasons: First, the underlying assets in the portfolio may have changing alphas and betas. Second, due to active trading strategies and thus dynamic change of portfolio weights, alpha and beta of the portfolio may be time-varying. Third, both of these characteristics may be inherent in the portfolio. While the first behaviour is not connected to the fund manager, the second is an indication of market timing skills of the manager.

The assumption of static alphas and betas can cause misleading conclusions. For example, let us assume that the value for static beta is lower than the real beta. Based on this assumption and the realised return of the fund, an outsider may conclude that alpha, i.e. the stock-selection skill of the manager, is higher than it is in reality. Better time-varying estimations of alpha and beta will therefore lead to better estimations of the real skills and behaviour of fund managers. Moreover, considering the high fees of hedge funds, clients do not want to pay these fees for passive returns they could get cheaper by investing in other products. Hence, it is important to know whether beta is static (indicating a passive strategy) or dynamic. If beta is dynamic, the question is whether this is due to changing betas of the underlying assets or to the change of weights of the portfolio through the manager. Only the latter represents an active strategy which justifies the raise of fees.

The models presented below have already been used in the context of measuring time-varying alphas and betas. Wells (1996) uses (amongst others) the Random Walk (RWM), Random Coefficient (RCM) and Mean Reverting (MRM) model to analyse the dynamics in alpha and beta of Swedish stocks. Yao and Gao (2004) use the same models for the analysis of Australian industry-portfolios. Lately, Mamaysky, Spiegel, and Zhang (2003) have proposed a model to simulate the asset allocation of a mutual fund manager and thus the dynamics of alpha and beta in a mutual fund portfolio (MSZM).

According to the above definition, we can assign the models to one of the two main characteristic groups: RWM and RCM belong both to the group assuming dynamic changes of alpha and beta of the underlying assets, while MSZM belongs to the group assuming dynamic changes of portfolio weights. MRM is somehow a hybrid model since it assumes some kind of long-term mean the exposure is reverting to. However it can be rather attributed to the first group than to the second one. Since MSZM assumes that the underlying assets have static alphas and betas, there is no model to describe the simultaneous change of underlyings and portfolio weights. However, we can assume that if both characteristics are present, the model which describes the predominant characteristic gives better results.

Based on the claim that hedge fund exposure is actively managed by the managers, we would accordingly expect that first models with time-varying betas perform better than static OLS in general. Secondly, the MSZM model which assumes the active management of portfolio weights should outperform the models assuming time-variation of only asset betas.

In the rest of this chapter, the models presented above and the economic reasoning behind them will be explained. A common factor to them is that they can be described in a state space environment where the observable variable, i.e. the return of hedge funds, is created from a state variable. The state variable itself evolves as an AR(1) process over time. In most cases, alpha and beta are directly state variables, whereas in MSZM, they are computed from the state variable. For the estimation of the models, the forecast of the state variables and hereby alpha and beta Kalman filter or in the case of MSZM Extended or Unscented Kalman filter are used. These filters are described in chapter 4.

In the following and for the rest of this thesis, returns are assumed to be excess returns.

3.2 The Static Case: OLS

Since it is the benchmark for all dynamic estimation methods, the equation for 'static' OLS is given below as a reminder. α and β are fixed and computed by minimizing the Mean Square Error (MSE).

$$r_t = \alpha + \beta r_{m,t} + v_t,\ v_t \sim \left(0, \sigma^2\right)$$

where α corresponds to the definition of Jensen's alpha.

In order to get somewhat time-adjusted alphas and betas, OLS can be run in a moving window which contains only part of the whole dataset.

3.3 Random Walk Model (RWM)

The Random Walk Model (RWM) assumes that alpha as well as beta evolve according to a random walk. This means that the current exposure to the market is a normally distributed random variable taking as mean the exposure of the last period. The system-noises are normally distributed and not correlated.

$$r_t = \alpha_t + r_{m,t}\beta_t + v_t,\ v_t \sim N\left(0, \sigma^2\right)$$

$$\begin{pmatrix} \alpha_t \\ \beta_t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha_{t-1} \\ \beta_{t-1} \end{pmatrix} + \begin{pmatrix} w_{1t} \\ w_{2t} \end{pmatrix}$$
$$\begin{pmatrix} w_{1t} \\ w_{2t} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_b^2 \end{pmatrix}\right)$$

3.4 Random Coefficient Model (RCM)

The Random Coefficient model (RCM) assumes that the state values alpha and beta vary randomly around a steady state mean $\overline{\alpha}$ and $\overline{\beta}$. Economically, this means that the exposure to the market deviates randomly and for a short time from a long term mean. The system noise is normally distributed with a mean of zero. Since the noise-term of alpha adds up to the observation noise and can not be distinguished from it, alpha is equal to the constant value $\overline{\alpha}$ and the dimension of the system equation is reduced to one (Wells 1996).

$$\begin{aligned} r_t &= \overline{\alpha} + r_{m,t}\beta_t + v_t,\ v_t \sim N\left(0, \sigma^2\right) \\ \beta_t &= \overline{\beta} + n_t,\ n_t \sim N\left(0, \sigma_n^2\right) \end{aligned}$$

3.5 Mean Reverting Model with Fixed Alpha

In this model, we assume that alpha is equal to a steady state value $\overline{\alpha}$, whereas beta is mean-reverting to a steady state value $\overline{\beta}$. In an economic sense, this model may be attractive because it models the behaviour of a manager who has a long-term target exposure, but deviates from it when the markets are either more or less favourable. In contrast to RCM, the actual beta does not randomly fluctuate around the long-term exposure, but approaches it over time.

$$\begin{aligned} r_t &= \overline{\alpha} + r_{m,t}\beta_t + v_t,\ v_t \sim N\left(0, \sigma^2\right) \\ \beta_t - \overline{\beta} &= \phi(\beta_{t-1} - \overline{\beta}) + n_t,\ n_t \sim N\left(0, \sigma_n^2\right) \end{aligned}$$

3.6 Mean Reverting Model with Random Walk Alpha

Instead of having a fixed alpha as in the model before, we introduce an alpha that evolves according to a random walk. Beta is still mean-reverting.

$$r_t = \alpha_t + r_{m,t}\beta_t + v_t,\ v_t \sim N\left(0, \sigma^2\right)$$

$$\begin{pmatrix} \alpha_t \\ \beta_t - \overline{\beta} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & phi \end{pmatrix} \begin{pmatrix} \alpha_{t-1} \\ \beta_{t-1} - \overline{\beta} \end{pmatrix} + \begin{pmatrix} w_{1t} \\ w_{2t} \end{pmatrix}$$
$$\begin{pmatrix} w_{1t} \\ w_{2t} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_b^2 \end{pmatrix}\right)$$

3.7 Mamaysky/Spiegel/Zhang Model (MSZM)

The model introduced by Mamaysky, Spiegel, and Zhang (2003) was originally developed for the estimation of alpha and beta of mutual funds, i.e. portfolios of long assets. While the authors assume that the alphas and betas of the underlying stocks are constant (which is not necessarily the case), they argue that those of mutual fund portfolios are not by nature. Hence, it is assumed that mutual fund portfolio returns follow a linear factor model with time varying-coefficients. If f_{t-1} is a vector of the fractions invested in each asset at time $t-1$, the portfolio return at time t can then be written as

$$\begin{aligned} r_{p,t} &= f_{t-1}^{t}\left(\alpha_t + \beta r_{m,t} + \epsilon_t\right) - k_t \\ &= \alpha_{p,t} + \beta_{p,t} r_{m,t} + \epsilon_{p,t}, \end{aligned} \tag{3.1}$$

where k_t is an estimation of the fund's management fees. According to the model, fund managers allocate the weights of the underlying assets based on a latent signal F_t which evolves as an AR(1) process:

$$F_t = \gamma_F F_{t-1} + \eta_{F,t},\ \eta_{F,t} \sim N(0, Q)$$

The persistence of the signal's value over time is measured by $\gamma_F \in [0, 1)$, and the economical meaning of F_t is not specified. Now define the fraction invested in each asset and the asset's alpha as follows

$$\begin{aligned} f_{i,t} &= \overline{f}_i + l_i F_t \\ \alpha_{i,t} &= \overline{\alpha}_i F_t \end{aligned}$$

Putting everything together and defining $\overline{f}$, l and $\overline{\alpha}$ as vectors consisting of the respective asset parameters, the portfolio's alpha and beta can be expressed in the following way

$$\begin{aligned} \alpha_{p,t} &= \overline{f}^t \overline{\alpha} F_{t-1} + l^t \overline{\alpha} F_{t-1}^2 - k_t \\ &= \overline{\alpha}_p F_{t-1} + b_p F_{t-1}^2 - k_t \\ \beta_{p,t} &= \beta \overline{f} + \beta l F_{t-1} \\ &= \overline{\beta}_p + c_p F_{t-1} \end{aligned}$$

Now, it is possible to estimate the mutual fund's alpha and beta without knowing the individual weights and coefficients of the underlying assets. Inserting alpha and beta into equation 3.1, the main equation to estimate is

$$r_{p,t} = b_p F_{t-1}^2 - k_t + \overline{\beta}_p r_{m,t} + \left(\overline{\alpha}_p + c_p r_{m,t}\right) F_{t-1} + \epsilon_{p,t} \tag{3.2}$$

The linear term $\overline{\alpha}_P$ reflects the contribution of the instantaneous alphas of the underlying stocks to the fund's strategy. However, as F_t itself can be positive or negative, a positive $\overline{\alpha}_P$ does not indicate over- or underperformance. On the other hand, parameter b_p in the quatratic term "measures the degree to which funds choose trading strategies that systematically profit from high frequency variation in security alphas over time" (Mamaysky, Spiegel, and Zhang 2003). It "measures the degree to which a fund is able to systematically go long (short) positive (negative) alpha stocks" and may give rise to occasional outperformance of the fund.

Since we are finally interested in an alpha net of fees and since k_t is an additional parameter to estimate, it is ommitted in our implementation of the model. In addition the authors show that in the univariate case, c_p can be normalised to one.

4 Kalman Filter

4.1 Introduction

The Kalman filter was first presented in 1960 in a paper by R.E. Kalman. The filter provides a recursive solution to a discrete-data linear filtering problem. It estimates a hidden system state in a way that minimizes the mean squared error. Since its introduction and with the increase in computing power, the Kalman filter has become an important tool mainly in engineering. Moreover, the filter has been extended in order to cover also nonlinear filtering problems.

Various books and publications cover the basics and applications of the filter. Probably the most comprehensive of them was written by Harvey (1989). It covers the mathematical background of the filter as well as applications on financial time series in general. Welch and Bishop (2001) give a well understandable overview of the Kalman and Extended Kalman Filter. In their tutorial about Bayesian algorithms for nonlinear/non-Gaussian tracking problems, Arulampalam, Maskell, Gordon, and Clapp (2002) derive the Kalman and Extended Kalman Filter from the more general framework of Bayesian Filter which includes also Particle Filter. Regarding the application of filters in finance, Javaheri, Lautier, and Galli (1980) provide a compact introduction to Kalman Filter which covers also the Unscented Kalman Filter.

In the rest of the chapter, we give a brief overview of the standard Kalman filter and its extensions. We follow the approach of Arulampalam, Maskell, Gordon, and Clapp (2002) by starting with recursive Bayesian filters, of which the Kalman filters are special cases.

4.2 Recursive Bayesian Filter

Filtering has been used for a long time in Control Engineering and Signal Processing, e.g. for tracking the position of an aircraft or robot. The general idea is that an outsider only observes the noisy measurement of an observation equation, which is realized by a hidden state not observable by the outsider. The state itself evolves over time according to a state equation. The goal of the outsider is to estimate the hidden state by using the filter and the observations made. In the Bayesian context, this means that based on all information available until a certain time, the posterior probability density function (pdf) of the state is estimated. In principle, an optimal estimate for the state may be obtained from this pdf, as it contains all statistical information. Often, the estimation of the state is necessary when new measurement information is available. In such a case, it is convenient to use a recursive filter since for every step only a sequential update must be computed, saving computing power and memory. This process is thus iterative and consists of two main steps:

- The a priori estimation of the hidden state or state pdf by using all prior information.
- The update of this estimation in order to get an a posteriori state estimation or state estimation pdf by using the new observation. This step is done by using Bayes theorem.

4.3 Nonlinear Bayesian Tracking

The evolotion of the state sequence is given by the transition equation.

Transition equation

$$\mathbf{x}_k = \mathbf{f}_k \left(\mathbf{x}_{k-1}, \mathbf{w}_k \right) \tag{4.1}$$

where $\mathbf{f}_k$ is a linear or nonlinear function of the state $\mathbf{x}_{k-1}$. The index k indicates that f_k may be time-variant. The process noise $\mathbf{w}_k$ is an i.i.d. sequence.

In order to estimate the state $\mathbf{x}_k$ recursively from the observations, the measurement equation is used.

Measurement equation

$$\mathbf{z}_k = \mathbf{h}_k \left(\mathbf{x}_k, \mathbf{u}_k \right) \tag{4.2}$$

Again, $\mathbf{z}_k$ is a linear or nonlinear function of the state $\mathbf{x}_k$, the index k indicates that $\mathbf{h}_k$ may be time-variant, and the measurement noise $\mathbf{u}_k$ is an i.i.d sequence.

A priori process estimate The process of recursive filtering is now defined as calculating an a priori belief in the state given the observation information $\mathbf{z}_{1:k-1}$.

$$\hat{\mathbf{x}}_k^- = E[\mathbf{x}_k]$$

A posteriori process estimate After the observation $\mathbf{z}_k$ has become available, the a posteriori process estimate is made

$$\hat{\mathbf{x}}_k = E[\mathbf{x}_k|\mathbf{z}_{1:k}] = E[\mathbf{x}_k|\mathbf{z}_k]$$

where the latter equals sign comes from the fact that the process equation is a Markov process.

Prediction step In terms of probability distributions, this means that the prior distribution of the pdf is obtained by

$$\begin{aligned} p\left(\mathbf{x}_k|\mathbf{z}_{1:k-1}\right) &= \int p\left(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_{1:k-1}\right) p\left(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1}\right) d\mathbf{x}_{k-1} \\ &= \int p\left(\mathbf{x}_k|\mathbf{x}_{k-1}\right) p\left(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1}\right) d\mathbf{x}_{k-1} \end{aligned}$$

whereas it is assumed that the initial prior of the state vector $p\left(\mathbf{x}_0|\mathbf{z}_0\right) \equiv p\left(\mathbf{x}_0\right)$ is available.

Measurement update step When the observation $\mathbf{z}_k$ becomes available, the prior is updated via Bayes' rule.

$$\begin{aligned} p\left(\mathbf{x}_k|\mathbf{z}_{1:k}\right) &= \frac{p\left(\mathbf{z}_k|\mathbf{x}_k\right) p\left(\mathbf{x}_k|\mathbf{z}_{1:k-1}\right)}{p\left(\mathbf{z}_k|\mathbf{z}_{1:k-1}\right)} \\ p\left(\mathbf{z}_k|\mathbf{z}_{1:k-1}\right) &= \int p\left(\mathbf{z}_k|\mathbf{x}_k\right) p\left(\mathbf{x}_k|\mathbf{z}_{1:k-1}\right) d\mathbf{x}_k \end{aligned}$$

This way, the posterior density is computed from the prior density by using the last observation. The density $p\left(\mathbf{z}_k|\mathbf{x}_k\right)$ is called the likelihood function.

The above equations desribe the optimal Bayesian solution, which can in general not be calculated analytically. Solutions only exist under certain restrictions, as is the case for the Kalman filter. In addition, Extended Kalman and Particle filters approximate the optimal Bayesian solution when there is no analytical solution.

4.4 Kalman Filter

The assumption of the Kalman filter is that the posterior density is Gaussian at every time step. If $p(x_{k-1}|z_{1:k-1})$ is Gaussian, also $p(x_k|z_{1:k})$ will be Gaussian if

- v_k and u_k are normally distributed
- $\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{w}_k)$ is a linear function of $\mathbf{x}_{k-1}$ and $\mathbf{w}_k$
- $\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{u}_k)$ is a linear function of $\mathbf{x}_k$ and $\mathbf{u}_k$

The Transition Equation (4.1) can then be written as state space model

$$\mathbf{x}_k = \mathbf{A}_k \mathbf{x}_{k-1} + \mathbf{w}_k \tag{4.3}$$

$$\mathbf{w}_k \sim (0, \mathbf{Q}_k)$$

The Observation Equation (4.2) can be written as

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_{k-1} + \mathbf{u}_k \tag{4.4}$$

$$\mathbf{u}_k \sim (0, \mathbf{R}_k)$$

The index k in $\mathbf{A}_k$, $\mathbf{H}_k$, $\mathbf{Q}_k$ and $\mathbf{R}_k$ indicates that these matrices may be time-variant. The Kalman filter consists then of the following recursive steps

Prediction step The prediction of the system state and covariance are defined as

$$\hat{\mathbf{x}}_k^- = \mathbf{A}_k \hat{\mathbf{x}}_{k-1} \tag{4.5}$$

$$\mathbf{P}_k^- = \mathbf{A}_k \mathbf{P}_{k-1} \mathbf{A}_k^t + \mathbf{W}_k \mathbf{Q}_{k-1} \mathbf{W}_k^t \tag{4.6}$$

Innovation Using the prediction of the state, the innovation $\hat{\mathbf{z}}_k^-$ and once the new observation is available also the prediction error v_k can be computed

$$\hat{\mathbf{z}}_k^- = \mathbf{H}_k \hat{\mathbf{x}}_k \tag{4.7}$$

$$v_k = \mathbf{z}_k - \hat{\mathbf{z}}_k^- \tag{4.8}$$

Measurement update step The covariance of the innovation term is now

$$\mathbf{S}_k = \mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^t + \mathbf{U}_k \mathbf{R}_k \mathbf{U}_k^t \tag{4.9}$$

and the Kalman gain

$$\begin{aligned}\mathbf{K}_k &= \mathbf{P}_k^- \mathbf{H}_k^t \left(\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^t + \mathbf{U}_k \mathbf{R}_k \mathbf{U}_k^t\right)^{-1} &(4.10)\\ &= \mathbf{P}_k^- \mathbf{H}_k^t \mathbf{S}_k^{-1} &(4.11)\end{aligned}$$

Finally, the updated system state and covariance can be computed:

$$\begin{aligned}\hat{\mathbf{x}}_k &= \hat{\mathbf{x}}_k^- + \mathbf{K}_k \upsilon_k &(4.12)\\ \mathbf{P}_k &= \left(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k\right) \mathbf{P}_k^- &(4.13)\end{aligned}$$

When the above assumptions hold, the Kalman filter is the optimal solution to the problem. For this reason, the Kalman filter is the minimum mean-square estimator (MMSE).

4.5 Extended Kalman Filter (EKF)

When the strict assumptions of the Kalman filter do not hold, approximate filters must be used. The most simple of this category is the Extended Kalman filter (EKF). As the standard Kalman filter, EKF assumes that the posterior density $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ is normally distributed, i.e. it is approximated by a Gaussian. However, the system and/or observation equation are now no longer linear and must be linearised by computing the Jacobian matrices. After the linearisation, the equations for the standard Kalman filter can be used.

Jacobian Matrices The Jacobian matrices are defined as the derivatives of the nonlinear functions with respect to the corresponding variables in the system and observation equations.

$$\begin{aligned}\mathbf{A}_{ij} &= \partial \mathbf{f}_i / \partial \mathbf{x}_j \left(\hat{\mathbf{x}}_{k-1}, 0\right)\\ \mathbf{W}_{ij} &= \partial \mathbf{f}_i / \partial \mathbf{w}_j \left(\hat{\mathbf{x}}_{k-1}, 0\right)\\ \mathbf{H}_{ij} &= \partial \mathbf{h}_i / \partial \mathbf{x}_j \left(\hat{\mathbf{x}}_k^-, 0\right)\\ \mathbf{U}_{ij} &= \partial \mathbf{h}_i / \partial \mathbf{u}_j \left(\hat{\mathbf{x}}_k^-, 0\right)\end{aligned}$$

Once this matrices are computed, they can be plugged into the standard Kalman filter algorithm.

4.6 Unscented Kalman Filter (UKF)

The Unscented Kalman filter (UKF) is an approach first introduced by Julier and Uhlmann for Kalman filtering in the case of nonlinear equations. It also approximates the posterior density $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ with a Gaussian, but compared to EKF, which uses Jacobians, UKF approximates the distribution of the state variable by using an unscented transformation. UKF will provide a more accurate mean than EKF while giving the same accuracy for covariance. In general, the state space must be augmented by concatenating the noises and will have a new dimension $n_a = n_x + n_w + n_u$. However this is not necessary when the equations are linear in noise, i.e. $\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}) + \mathbf{w}_k$, $\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{u}_k$. An explicit description of UKF can be found in van der Merwe, Doucet, de Freitas, and Wan (2000). For convenience, we describe the main steps of the algorithm below.

Initialisation The algorithm is initialised with the initial weights for the so-called sigma points (which are necessary for the unscented transformation) and as always with an initial state and state covariance. Also, the scaling parameters κ, α and β must be chosen. We use the default settings of van der Merwe et al.: $\kappa = 0$, $\alpha = 1$ and $\beta = 0$.

$$\begin{aligned} W_0^{(m)} &= \frac{\lambda}{n_a + \lambda} \\ W_0^{(c)} &= \frac{\lambda}{n_a + \lambda} + (1 - \alpha^2 + \beta) \\ W_i^{(m)} &= W_i^{(c)} = \frac{1}{2(n_a + \lambda)} \quad i = 1, \ldots, 2n_a \end{aligned}$$

The following steps are now performed for every datapoint.

Augmentation of State Space The original state space is extended by adding the observation and system noise.

$$
\begin{aligned}
\hat{\mathbf{x}}_0 &= E[\mathbf{x}_0] \\
\mathbf{P}_0 &= E[(\mathbf{x}_0 - \hat{\mathbf{x}}_0)(\mathbf{x}_0 - \hat{\mathbf{x}}_0)^t] \\
\mathbf{x}^a_{k-1} &= \begin{pmatrix} \mathbf{x}_{k-1} \\ \mathbf{w}_{k-1} \\ \mathbf{u}_{k-1} \end{pmatrix} \\
\hat{\mathbf{x}}^a_{k-1} &= E[\mathbf{x}^a_{k-1} | z_k] = \begin{pmatrix} \hat{\mathbf{x}}_{k-1} \\ 0 \\ 0 \end{pmatrix}
\end{aligned}
$$

$$
\mathbf{P}^a_{k-1} = \begin{pmatrix} \mathbf{P}_{k-1} & \mathbf{P}_{xw}(k-1|k-1) & 0 \\ \mathbf{P}_{xw}(k-1|k-1) & \mathbf{P}_{ww}(k-1|k-1) & 0 \\ 0 & 0 & \mathbf{P}_{uu}(k-1|k-1) \end{pmatrix}
$$

Calculation of Sigma Points The sigma points, which are needed for the unscented transformation, are computed in the following way:

$$
\begin{aligned}
\chi^a_{k-1}(0) &= \hat{\mathbf{x}}^a_{k-1} \\
\chi^a_{k-1}(i) &= \hat{\mathbf{x}}^a_{k-1} + \left(\sqrt{(n_a + \lambda)\mathbf{P}^a_{k-1}}\right)_i \quad i = 1, \ldots, n_a \\
\chi^a_{k-1}(i) &= \hat{\mathbf{x}}^a_{k-1} - \left(\sqrt{(n_a + \lambda)\mathbf{P}^a_{k-1}}\right)_{i-n_a} \quad i = n_a + 1, \ldots, 2n_a
\end{aligned}
$$

Prediction Step The equations for the prediction of the state value and covariance are

$$
\begin{aligned}
\chi_{k|k-1}(i) &= \mathbf{f}(\chi^x_{k-1}(i), \chi^w_{k-1}(i)) \quad i = 0, \ldots, 2n_a + 1 \\
\hat{\mathbf{x}}^-_k &= \sum_{i=0}^{2n_a} W^{(m)}_i \chi_{k|k-1}(i) \\
\mathbf{P}^-_k &= \sum_{i=0}^{2n_a} W^c_i \left(\chi_{k|k-1}(i) - \hat{\mathbf{x}}^-_k\right) \left(\chi_{k|k-1}(i) - \hat{\mathbf{x}}^-_k\right)^t
\end{aligned}
$$

Innovation By using the state prediction, the innovation and the prediction error v_k are received.

$$\begin{aligned}
\mathbf{z}_{k|k-1}(i) &= \mathbf{h}\left(\chi_{k|k-1}(i), \chi^u_{k-1}(i)\right) \\
\hat{\mathbf{z}}_k^- &= \sum_{i=0}^{2n_a} W_i^{(m)} \mathbf{z}_{k|k-1}(i) \\
v_k &= \mathbf{z}_k - \hat{\mathbf{z}}_k^-
\end{aligned}$$

Measurement Update Step Finally, by computing the predicted covariance, we get the Kalman gain $\mathbf{K}_k$.

$$\begin{aligned}
\mathbf{P}_{z_k z_k} &= \sum_{i=0}^{2n_a} W_i^{(c)} \left(\mathbf{z}_{k|k-1}(i) - \hat{\mathbf{z}}_k^-\right)\left(\mathbf{z}_{k|k-1}(i) - \hat{\mathbf{z}}_k^-\right)^t \\
\mathbf{P}_{x_k z_k} &= \sum_{i=0}^{2n_a} W_i^{(c)} \left(\chi_{k|k-1}(i) - \hat{\mathbf{x}}_k^-\right)\left(\mathbf{z}_{k|k-1}(i) - \hat{\mathbf{z}}_k^-\right)^t \\
\mathbf{K}_k &= \mathbf{P}_{x_k z_k} \mathbf{P}_{z_k z_k}^{-1}
\end{aligned}$$

As before, we can now update the system state and covariance.

$$\begin{aligned}
\hat{\mathbf{x}}_k &= \hat{\mathbf{x}}_k^- + \mathbf{K}_k v_k \\
\mathbf{P}_k &= \mathbf{P}_k^- - \mathbf{K}_k \mathbf{P}_{z_k z_k} \mathbf{K}_k^t
\end{aligned}$$

4.7 Parameters and Maximum Likelihood Estimation (MLE)

The equations that form the Kalman filter imply various parameters. In general, these parameters can be subdivided into two categories

- hyperparameters
- initial parameters

The hyperparameters are all parameters inherent in the system and observation equation, in particular the system and observation covariances as well as possible additional model parameters (see chapter 3). Whereas under ideal circumstances these parameters are known in reality they are most often not.

As the Kalman filter is a recursive filter with regard to its state, it needs some initial state parameters, namely an initial state value and an initial state covariance. These values may be estimated by various approches. For a long enough time-series, errors in the estimation of the initial parameters should become small over the several recursive steps.

Maximum Likelihood Estimation (MLE) As mentioned above, the hyperparamters are normally not known under real-world conditions, hence these parameters must be estimated. In order to do so, a (quasi) maximum likelihood estimation can be used where the function $\prod_{k=1}^{N} p(z_k|z_{1:k-1})$ is maximised. The density $p(z_k|z_{1:k-1})$ is Gaussian, so the loglikelihood to be maximised is

$$L_{1:N} = -\frac{N}{2}\ln(2\pi) - \frac{1}{N}\sum_{k=1}^{N}\ln(S_k) - \frac{1}{N}\sum_{k=1}^{N}\frac{v_k^2}{S_k} \tag{4.14}$$

This method for estimating parameters is applicable for standard Kalman filter as well as for EKF and UKF. The maximisation is normally done by using numerical maximisation methods. In order to decrease the optimisation time and to find a global maximum, appropriate initial values for the hyperparameters should be provided to the maximisation method.

4.8 Filter-Application to Models with Time-Varying Exposure

In order to use Kalman filter for the models introduced in chapter 3, some provisions need to be made. First, the values for the initial state and state covariance must be defined. Wells (1996) and Harvey (1989) give directions for the various models. Moreover, the use of EKF requires the linearisation of the MSZM model.

4.8.1 Initial Values

RWM According to Wells (1996), the state values and covariances can be initialized in the following way:

$$\begin{aligned}
\alpha_0 &= \alpha \text{ of OLS with first ten observations} \\
\beta_0 &= \beta \text{ of OLS with first ten observations} \\
\sigma_{\alpha,0} &= \text{variance of } \alpha \text{ of OLS with first ten observations} \\
\sigma_{\beta,0} &= \text{variance of } \beta \text{ of OLS with first ten observations} \\
P_0 &= \begin{pmatrix} \sigma_{\alpha,0} & 0 \\ 0 & \sigma_{\beta,0} \end{pmatrix}
\end{aligned}$$

RCM Since β_t is independent from β_{t-1} in this model, $\mathbf{A}_t$ in equation 4.3 is equal to zero for all t. Thus, the initial state value and variance are superfluous in the Kalman prediction step (equation 4.6).

MRM with Random Walk Alpha For alpha the initial state value for is set to the OLS estimate of the first ten observations. The initial state variance is set to the variance of the OLS estimate of the first 10 observations. For beta the initial state value is set to zero and the initial state variance is set to a big value (according to Wells):

$$\begin{aligned}
\alpha_0 &= \alpha \text{ of OLS with first ten observations} \\
\beta_0 &= 0 \\
\sigma_{\alpha,0} &= \text{variance of } \alpha \text{ of OLS with first ten observations} \\
\sigma_{\beta,0} &= 100 \\
P_0 &= \begin{pmatrix} \sigma^2_{\alpha,0} & 0 \\ 0 & \sigma^2_{\beta,0} \end{pmatrix}
\end{aligned}$$

MSZM Model The initial state value is set to zero, whereas the initial state variance is set to one. As the estimated variance of the state is mostly below 0.1, the initial state variance is fairly big.

$$
\begin{aligned}
F_0 &= 0 \\
P_0 &= 1
\end{aligned}
$$

4.8.2 Linearisation of MSZM Model for EKF

As the MSZM model is not linear, UKF or EKF instead of the standard Kalman filter must be used. In order to use EKF,the main model equation 3.2 has to be linearised. The derivation of the formulas is shown below.

Prediction step

$$
\begin{aligned}
F_{t-1}^{-} &= \gamma_F F_{t-2} \\
P_{t-1}^{-} &= \gamma_F P_{t-2} \gamma^t + Q
\end{aligned}
$$

Measurement update step

$$
\begin{aligned}
\alpha_t^{-} &= \overline{\alpha} F_{t-1}^{-} - k_t + b(F_{t-1}^{-})^2 \\
\beta_t^{-} &= \overline{\beta} + F_{t-1}^{-} \\
v_t &= r_t - \left(\alpha_t^{-} + \beta_t^{-} r_{m,t}\right) \\
H &= 2bF_{t-1}^{-} + \overline{\alpha} + (r_{m,t} - r_t) \\
S_t &= HP_{t-1}^{-} H^t + R \\
K &= \frac{P_{t-1}^{-} H^t}{S_t} \\
F_{t-1} &= F_{t-1}^{-} + K v_t \\
P_{t-1} &= P_{t-1}^{-} - KHP_{t-1}^{-}
\end{aligned}
$$

Whereas Mamaysky, Spiegel, and Zhang (2003) use only EKF for their estimation, we use also UKF which theoretically gives better estimations for the state and requires no computation of Jacobian matrices.

4.9 Summary

The Kalman filter is a special case of a Bayesian filter, assuming linear state and observation equations and Gaussian noises. The Extended Kalman Filter (EKF) is a linearisation of the Kalman filter for nonlinear state and/or observation equations which uses a first-order Taylor expansion. Still, it assumes Gaussian noises. In the case of linear equations, the EKF is equivalent to the standard Kalman filter. The Unscented Kalman Filter (UKF) is another approach to nonlinear systems with Gaussian noises. Unlike EKF, it does not linearise the equations and therefore needs no Jacobians, but it approximates the state variable by using an unscented transformation to it. Another category of filters which are not discussed here are Particle filters. Instead of a Gaussian approximation of the a posteriori state, they use Markov-Chain Monte-Carlo simulations and are therefore also used in the case of Non-Gaussian noises. However, the higher flexibility leads to an increase of complexity, also in terms of processing power.

5 Tests and Results

5.1 Data

5.1.1 Dependent Variables

As dependent variable, monthly returns of long/short equity indices from Hedge Fund Research Monthly Indices (HFRI) and CSFB/Tremont are used. Data is taken from January 1994 until June 2005. As internal analysis by RMF shows, Cusum, Mosum and F tests indicate that there is a structural breakpoint in February 2000 (figure 5.1). This breakpoint causes problems for the convergence of the Kalman filter's maximum likelihood estimation, especially since the number of available datapoints is already small. For this reason, the returns from January 2000 to March 2000 are removed from the dataset.

5.1.2 Regressors

For tests against market return, monthly returns on S&P 500 are chosen. The monthly data for the Fama/French factors SMB (Small Minus Big) and HML (High Minus Low) as well as for UMD (Momentum Factor) is taken from 'Data Library' on Kenneth R. Frenchs homepage (French 2005). As risk-free asset 3-month LIBOR in USD is taken.

5.2 Implementation

The algorithms presented in chapter 3 were implemented in Matlab. For maximisation (respectively minimisation) of the likelihood function, the fmincon algorithm was used. The implementation of UKF was adapted from a demonstration-software by Nando de Freitas for the paper by van der Merwe, Doucet, de Freitas, and Wan (2000). De Freitas himself used the code of Julier and van der Merwe.

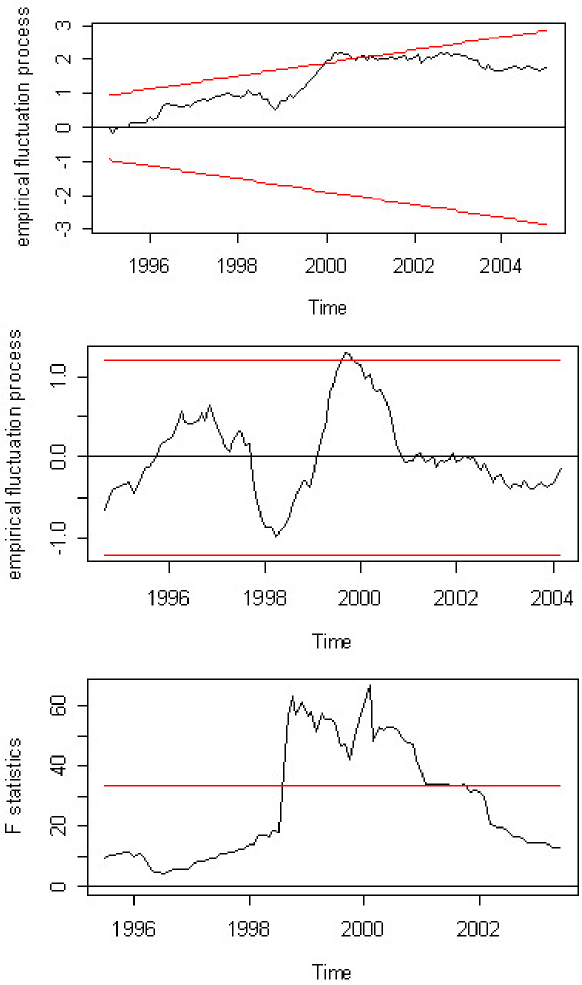

Figure 5.1: Standard Cusum, OLS Cosum and F Test. Source: RMF

5.3 Methodology

5.3.1 Moving Window

The beta forecasts of the different models are compared by using a moving window that contains a fixed number of data points. The parameters are estimated by maximising the likelihood within this window, then the filter is run until the last datapoint of the window. The alpha and beta forecast for the next month is then taken to estimate the performance. The parameter estimation especially for the more complex models requires a certain number of datapoints. For this reason and based on tests the window size was set to 80 datapoints.

5.3.2 Parameter Estimation

As described in section 4.14, maximum likelihood estimation (MLE) is used to estimate the unknown hyperparameters. In Matlab, the function fmincon provides numerical minimisation for nonlinear, unconstrained problems. While the estimation of the simpler models like RWM is straightforward with this function, the estimation of the parameters of MSZM caused many problems. Six parameters must be estimated in this model, and with the available number of data points, there exist several local maxima. In particular, it seems that the likelihood is sensitive with regard to the values of b, $\overline{\alpha}$ and γ. Under certain circumstances MLE of these parameters is also sensitive to the initial value of state and covariance. In order to overcome this problem, several approaches were tried. First, parameter ranges were restricted were it seemed sensible, especially γ was allowed to vary only between 0 and 1. In addition, a grid containing three starting values for each of the three most critical parameters was constructed. This leads to 27 different starting values from which the minimisation is started. Finally, the lowest of the 27 results is taken. Despite these extensions, parameter estimations between two moving windows, which differ in only two data points, sometimes still resulted in significantly different parameter values. Hence, as next remedy, the parameter estimations from the previous moving window estimation were used as starting values of the optimisation for the next window. This helped to remove some of the diverging parameter estimations, but not all. Finally, as Mamaysky et al. report that b seems to be close to zero in most cases, we tried to estimate the parameters by using a grid, but fixing b to zero in a first step. The minimisation of this constrained model converges much better than when b is also estimated. In a second step, the received optimal parameters from the former estimation were then used as starting values for the minimisation of the function with all parameters, but only one set of starting values. This method seems to provide continuous parameter values for adjacent windows in most cases, although the success still depends on the underlying dataset. It should be mentioned that the resulting likelihood is sometimes smaller than the one with non-continuous parameter values, but it is probable that the latter are erroneous and caused by the small number of observations. It seems that this happens when the estimation of the variance of the system equation is very low, resulting in some kind of oscillation of the likelihood function.

5.3.3 Performance Selection Criteria

When it comes to the selection of the best model, it is important to decide whether the evaluation is made in-sample or out-of-sample. While most studies in chapter 2 use R^2, i.e. an in-sample criteria, we use out-of-sample criteria since we are interested in exposure forecasts and a good in-sample performance is no guarantee for a good out-of-sample performance. There is actually no single or best choice for the evaluation of forecasts. Authors of former studies use several criteria and argue that the best model should perform best in most categories. However, the overall conclusion may be biased by the personal preferences of the author, especially when the different criteria provide ambiguous results.

As Yao and Gao (2004) and as Wells (1996), we will use the following criteria to evaluate the performance of the models:

Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) At the end of every moving window, the model estimate of next period's alpha and beta is made. Once next period's index and market return become available, the error between the estimated return using alpha and beta forecasts and the realised return can be calculated.

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T} v_t^2} \tag{5.1}$$

$$MAE = \frac{1}{T}\sum_{t=1}^{T} |v_t| \tag{5.2}$$

where v_t is defined as in 4.8, computed with the respective alpha and beta forecasts.
It should be stressed here that the above criteria use the one step ahead market returns for the computation of the forecasted one step ahead asset return which is then compared to the realised one step ahead asset return. Not using this one step ahead market return would require to forecast also the market return, which is a problem on its own. Since we are not interested in forecasting returns, but alphas and betas, this method should be feasible and provide measurable results.

Akaike Information Criterion (AIC) The Akaike Information Criterion (AIC) is used to compute the reduction of the likelihood considering the number of parameters that are needed to achieve this reduction. Harvey (1989) and Wells (1996) define it in the following way

$$AIC = \sigma^2 \exp \frac{2(k+d)}{T} \tag{5.3}$$

where σ^2 is the estimated variance of the observation equation, k equals the number of hyperparameters, d the state-dimension of the model and T the number of observations. However, the use of estimated σ^2 would be an in sample criterion. As we are interested in the σ^2 of the forecasts, we use the sample variance of the forecast errors within the moving window, which is equal to the squared value of RMSE.

R-square Calculating R^2 from the residuals of the updated state estimations provides an in-sample estimation, which cannot be used as selection criteria for accuracy of forecasts. However it gives an indication of the in-sample fit, so we list it in the tables.

5.4 Simulation

Before running the models with real data, a little simulation can show the performance of some models under artificial conditions. We compare OLS with different window length, RWM and MSZM with UKF. Based on a constantly increasing beta and an alpha kept constant, simulated asset returns are computed with the formula $R_{sim} = \alpha + \beta$.

As figure 5.2 shows, beta OLS approaches the real beta with decreasing window size, while RWM and MSZM are able to predict the real beta almost accurately.

In the case of alpha, the simulation indicates that RWM and MSZM with UKF are able to predict the constant alpha, whereas the alpha forecasts of OLS with different window size show no convergence (figure 5.3). Hence we conclude that both time-variant models are more accurate at least under non-real world conditions.

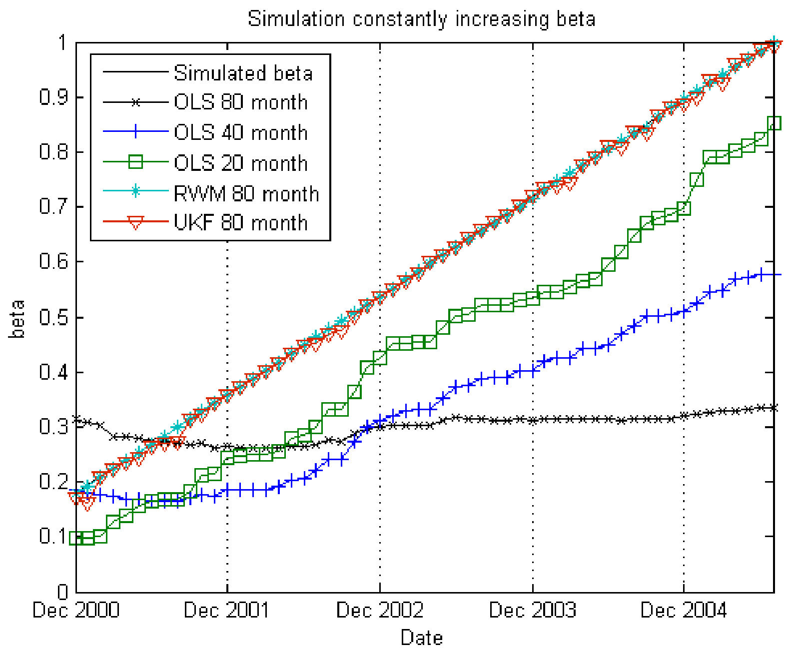

Figure 5.2: Simulation: estimation of continously increasing beta

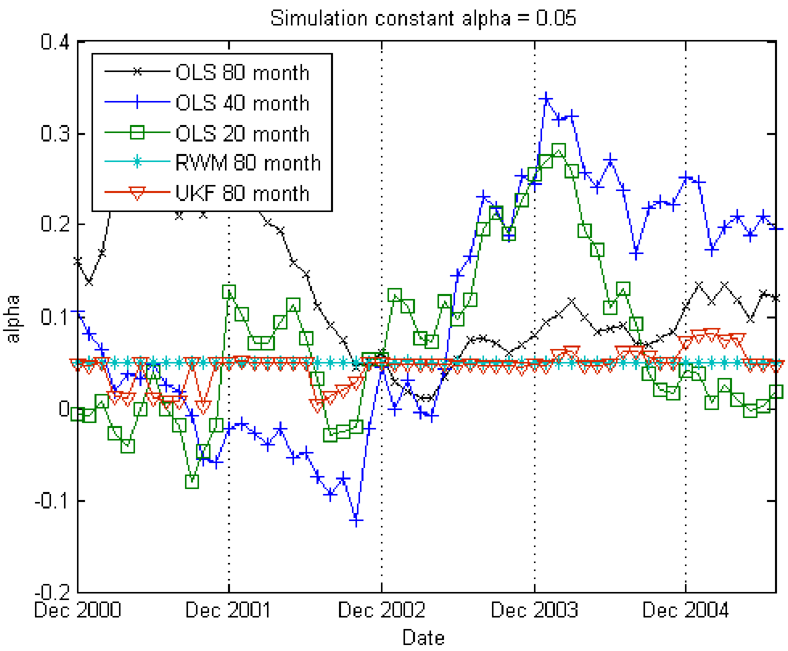

Figure 5.3: Simulation: estimation of constant alpha

5.5 Forecasts with Univariate Regressions

The various models are evaluated with HFRI and CSFB/Tremont long/short equity (and eventually HFRI equity market neutral) indices as dependent variable. As regressors, the three Fama/French factors plus UMD are taken. This corresponds to the four-factor model by Carhart, which extends the Fama-French three-factor model by the momentum factor UMD.

$$r_t - r_{f,t} = \alpha_t + \beta_{r_m,t}(r_{m,t} - r_{f,t}) + \beta_{SMB,t} SMB_t + \beta_{HML,t} HML_t + \beta_{UMD,t} UMD_t + \epsilon_t \quad (5.4)$$

In this section, the models are evaluated in a univariate approach with each of the four regressors independently.

5.5.1 Long/Short Equity against S&P 500

When regressed against S&P 500, the MSZM model performs best for both indices (see tables 5.1 and 5.2). In some categories, the UKF implementation has some advantages, in other it is EKF, however the differences between the two filters are small. In general, the differences between the models (including OLS) are smaller for HFRI than for CSFB/Tremont.

	OLS	**RWM**	**RCM**	**MRM fixed α**	**MRM rw α**	**EKF**	**UKF**
RMSE	1.248	1.321	1.244	1.268	1.279	**1.147**	1.155
MAE	0.953	1.028	0.948	0.973	1.011	0.946	**0.939**
AIC	**1.674**	2.094	1.858	2.075	2.111	1.698	1.720
Av R^2	0.575	0.709	0.659	0.743	0.818	0.890	0.891

Table 5.1: Results HFRI against S&P 500 80 month window

	OLS	**RWM**	**RCM**	**MRM fixed α**	**MRM rw α**	**EKF**	**UKF**
RMSE	1.907	1.492	1.910	1.514	1.534	1.372	**1.319**
MAE	1.433	1.101	1.427	1.120	1.110	1.020	**0.988**
AIC	3.911	2.669	4.377	2.851	3.035	2.428	**2.243**
Av R^2	0.481	0.777	0.722	0.846	0.852	0.891	0.889

Table 5.2: Results CSFB/Tremont against S&P 500 80 month window

A plot for the squared errors with CSFB/Tremont for the different models is shown in the graphs 5.4, 5.5 and 5.6. OLS serves as a benchmark in each case. Forecasting errors in the period from December 2000 to December 2002 seem to be much higher than the ones in the period from January 2003 to June 2005 for all models. Also, RCM seems to follow OLS closely, which is an indication the RCM does not perform well for this dataset.

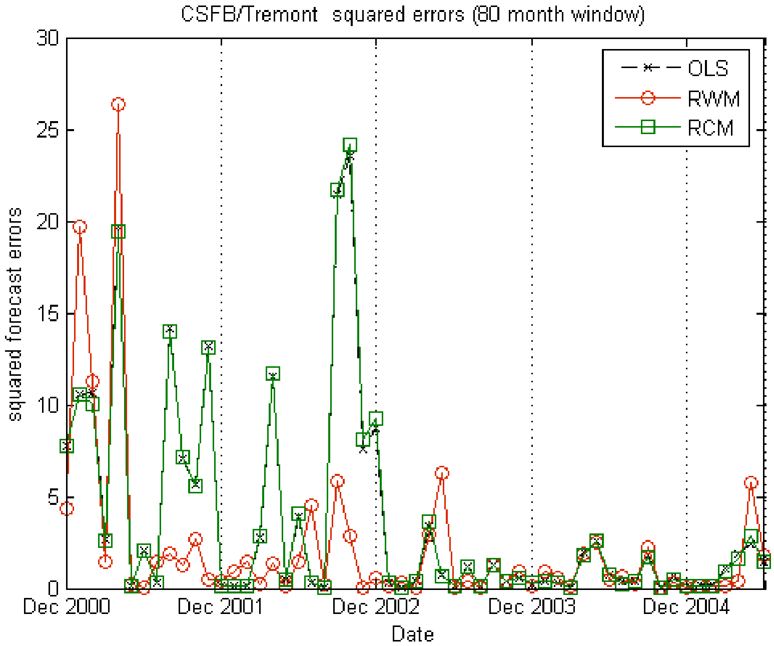

Figure 5.4: CSFB/Tremont squared forecast errors: comparison OLS, RWM, RCM

A comparison of the evolution of betas for OLS, RWM and MSZM with UKF is shown in figure 5.7. As can be seen, beta OLS changes smoothly over time, indicating little change in the exposure to the market. On the other hand, beta RWM follows a u-shaped convex curve which almost reaches zero in December 2002. Intuitively, this seems logical as it indicates that after years of just following the trend of the market with a high exposure, managers reduced their exposure to the market after the stock crash in the beginning of 2000. This is, in a weaker manner, also reflected in the evolution of beta OLS. However, beta RWM indicates that after December 2002, managers started again to increase their exposure to the market. Considering the upward trend of S&P 500 starting beginning of 2003, this seems logical.

For the period from 2002 on, the forecasting errors seem to be small both for OLS as well as for RWM although the betas are quite different, so we cannot automatically deduce that either of the models is more correct. However, we suppose that OLS is reacting slowly to changes in beta and is therefore overestimating (from beginning 2000 until mid 2004) respectively underestimating (from mid 2004 until mid 2005)the real

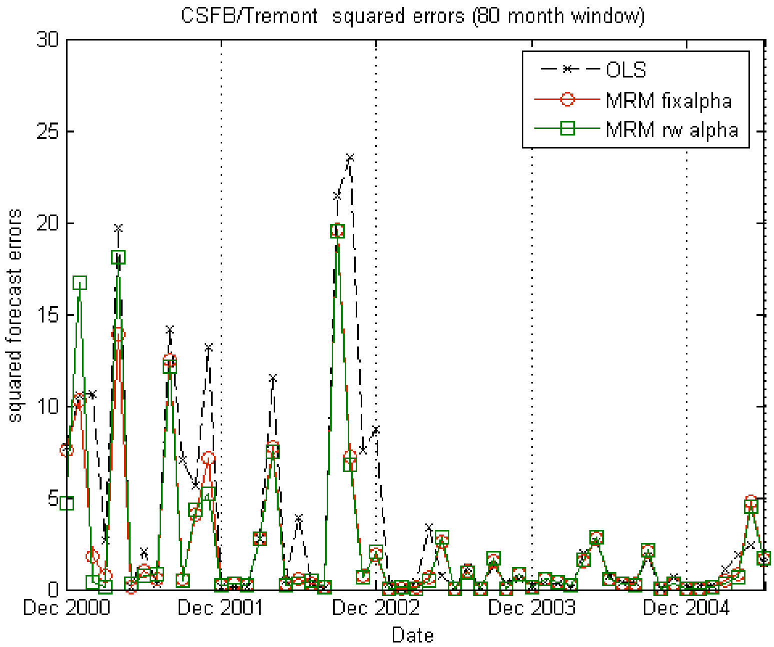

Figure 5.5: CSFB/Tremont squared forecast errors: comparison OLS, MRM with fixed or random walk alpha

value of beta. In order to gain more insight into this topic, we reduce the OLS window size from originally 80 to 40 respectively 20 datapoints, which makes the betas less dependent on past values and thus more dynamic. Figure 5.8 shows that the evolution of beta OLS with smaller window size approaches the evolution of beta RWM. This pattern is very similar to the one in section 5.4 where we show that in a simulation RWM is able to predict the real beta almost accurately, so we conclude that this should also be true here.

The picture is somewhat different for alpha. As figure 5.9 shows, alpha OLS is constantly higher than alpha RWM and most often also higher than alpha UKF. From the previous discussion we conclude that beta RWM provides a better estimation for real beta than beta OLS. Provided that the forecast error of RWM is equal to or smaller than the one of OLS, it follows that also alpha RWM is a better estimator than alpha OLS. This in turn indicates that OLS attributes too much weight to alpha, i.e. overvaluates the stock-selection skills of the manager, whereas RWM takes into account that some of the return comes from change of exposure and that real alpha is therefore smaller.

When the same test as before for beta with a reduction of the OLS window length is performed, alpha OLS shows no convergence to alpha RWM for decreasing window size (figure 5.10). A reason for this could be the increase in forecast errors when the window

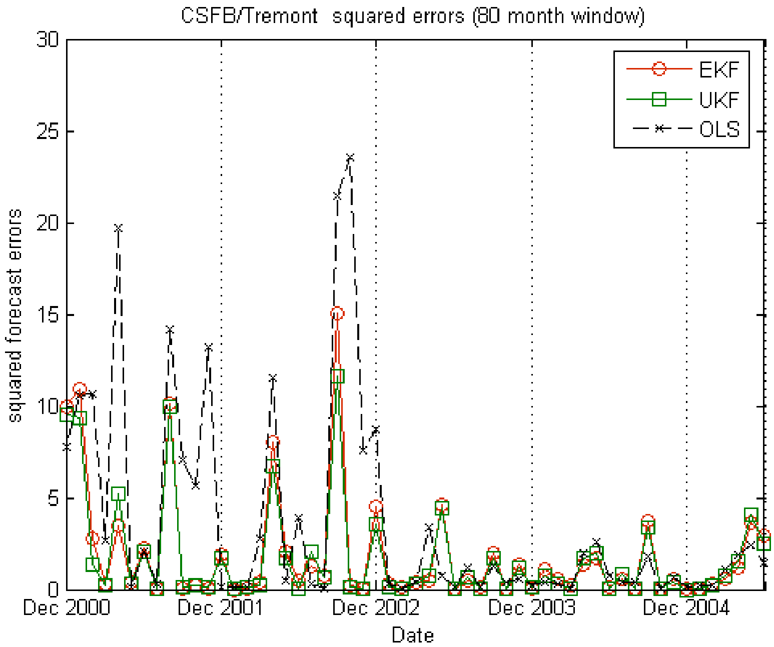

Figure 5.6: CSFB/Tremont squared forecast errors: comparison OLS, EKF, UKF

size of OLS is reduced which leads to a less accurate estimation of alpha (otherwise, we could just reduce OLS window size instead of using Kalman filters in order to get better estimates). Again we remember from section 5.4 that RWM and MSZM with UKF were able to predict the constant simulated alpha more accurate than OLS. Thus we conclude that for CSFB/Tremont, estimates of alpha and beta are better with RWM and MSZM with UKF than with OLS.

Finally, a remark about the parameter estimation of MSZM should be made. From chapter 3.7 we remember that according to Mamaysky, Spiegel, and Zhang (2003) the parameter b measures "the degree to which a fund can go systematically long (short) positive (negative) alpha stocks" and is thus an indication of the manager's timing skills. However, in the above regression we find slightly negative values for MSZM with EKF as well as with UKF. This indicates that for the average fund, there is a negative timing ability, something which was found for mutual funds in several studies. Although the authors do not explicitely comment on it, the average negative return contribution from dynamic betas found in Géhin and Vaissié (2005) leads to the same conclusion. Needless to say that on the individual level, we expect some of the funds to have positive timing ability.

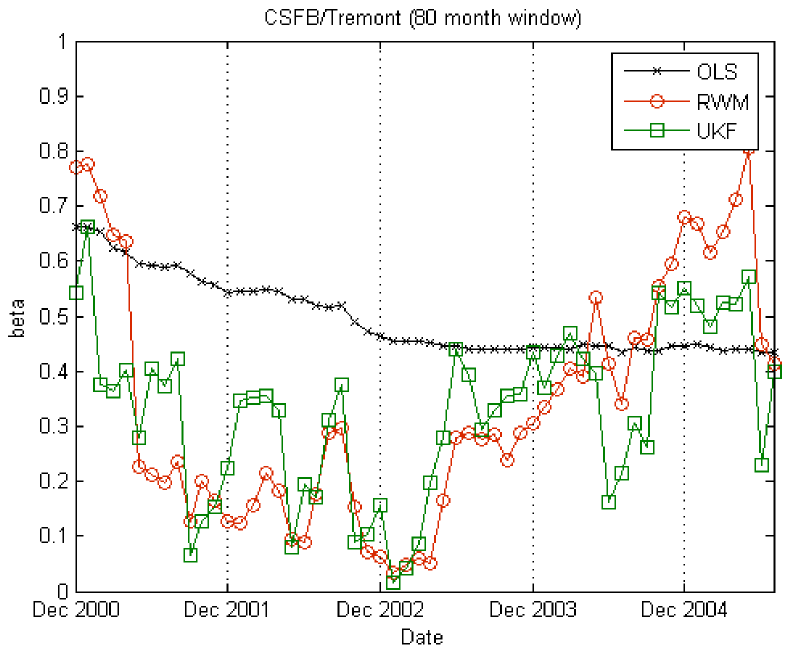

Figure 5.7: CSFB/Tremont: beta OLS, RWM and MSZM/UKF

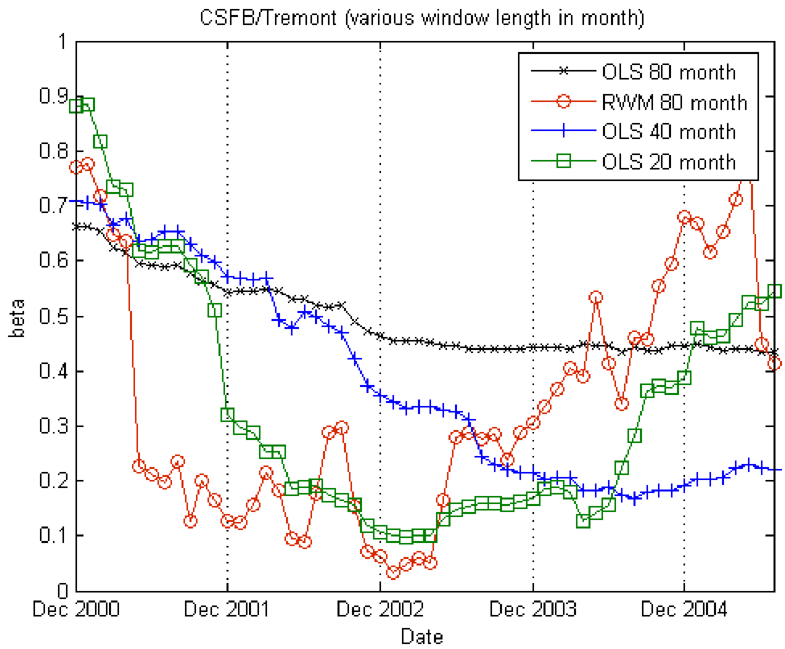

Figure 5.8: CSFB/Tremont: comparison of beta OLS for various window lengths

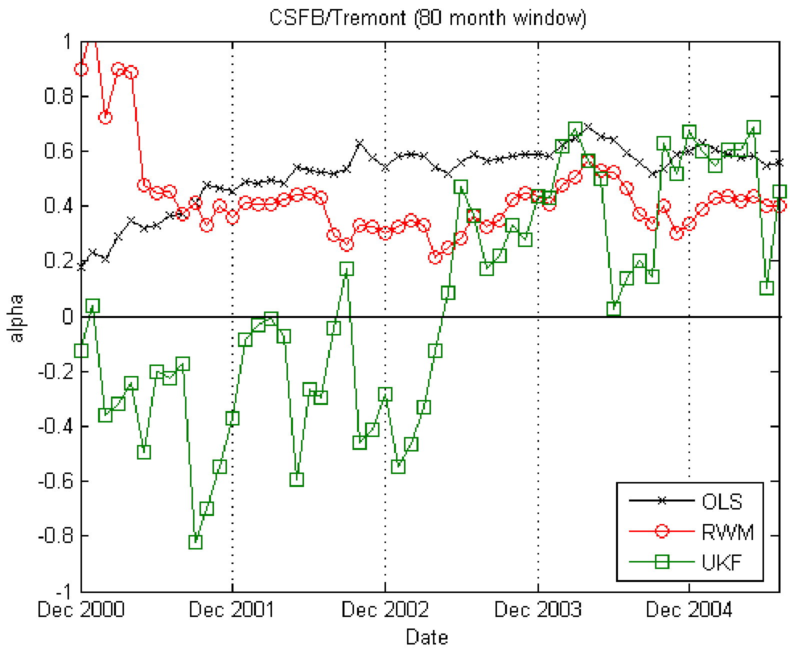

Figure 5.9: CSFB/Tremont: alpha OLS, RWM and MSZM/UKF

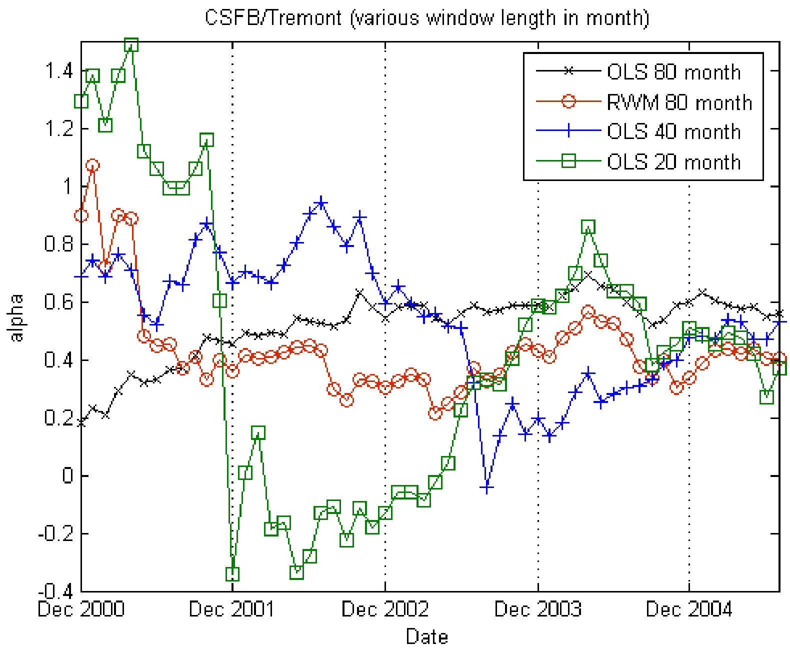

Figure 5.10: CSFB/Tremont: comparison of alpha OLS for various window lengths

5.5.2 Equity Market Neutral against S&P 500

Contrary to long/short equity which is long biased, equity market neutral style should have almost no exposure to S&P 500, which is also found with OLS. It is however interesting to analyse this style with time-variant models in order to find out if they confirm the low exposure or if the OLS result is just an average of positive and negative exposures over time. The results of a regression of HFRI equity market neutral are shown in table 5.3.

	OLS	**RWM**	**RCM**	**MRM fixed α**	**MRM rw α**	**EKF**	**UKF**
RMSE	0.798	**0.782**	0.790	0.791	0.843	0.874	0.847
MAE	0.585	**0.557**	0.579	0.580	0.587	0.627	0.622
AIC	**0.683**	0.729	0.743	0.772	0.909	0.977	0.916
Av R^2	0.018	0.268	0.651	0.655	0.643	0.463	0.488

Table 5.3: Results HFRI equity market neutral against S&P 500 80 month window

We can see that in this case RWM performs only slightly better than OLS, while MSZM performs worse than OLS. Moreover, we find that also for time-varying models, the maximal exposure is around 0.1, which indicates that exposure is indeed close to zero (see figure 5.11).

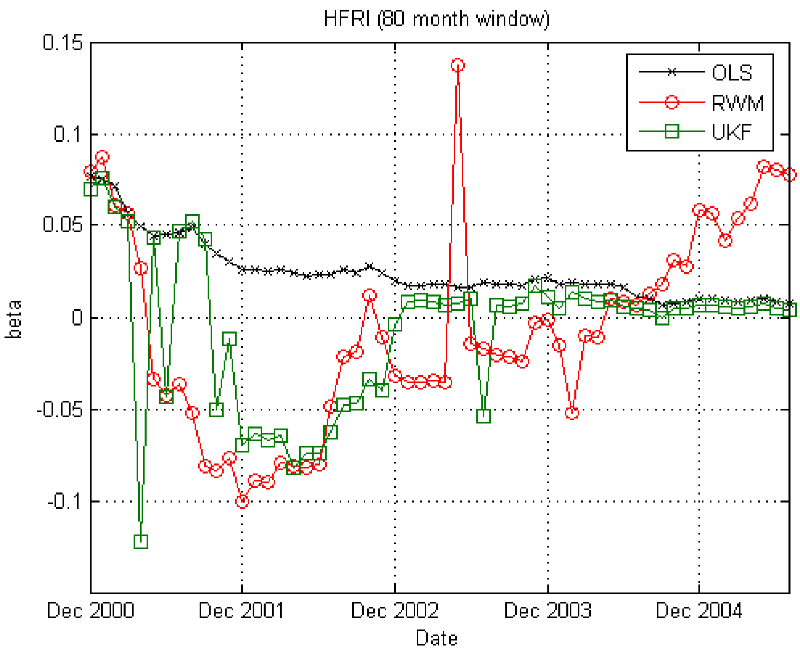

Figure 5.11: HFRI equity market neutral: beta OLS, RWM and MSZM/UKF

This finding may also explain the performance of the various models: While the portfolio weights are kept constant, time-varying betas in the underlying assets may lead to subtle variations in the portfolio exposure. Eventually, managers may also adapt the weights to correct the exposure, but nevertheless the time-variation in the underlyings is the predominant effect. Hence, we would expect that the models that account for this effect like RWM and RCM would perform better, while MSZM should provide no advantage.

5.5.3 Long/Short Equity against SMB

Small minus big (SMB) is the return on a capitalization-based portfolio buying small cap stocks and selling large cap stocks. Hence, portfolios containing small cap stocks will have positive $\beta_{SMB,t}$, while those containing large cap stocks will have negative $\beta_{SMB,t}$. From tables 5.4 and 5.5 we can see that for HFRI, MRM with random walk alpha shows a slight advantage against MSZM with UKF, whereas for CSFB/Tremont, MSZM with UKF performs slightly better.

	OLS	**RWM**	**RCM**	**MRM fixed α**	**MRM rw α**	**EKF**	**UKF**
RMSE	1.906	1.832	1.981	1.994	**1.657**	1.752	1.741
MAE	1.451	1.356	1.497	1.504	**1.289**	1.331	1.336
AIC	3.906	4.024	4.706	4.944	**3.544**	3.961	3.910
Av R^2	0.318	0.515	0.589	0.591	0.750	0.485	0.504

Table 5.4: Results HFRI against SMB 80 month window

	OLS	**RWM**	**RCM**	**MRM fixed α**	**MRM rw α**	**EKF**	**UKF**
RMSE	1.909	2.189	1.940	1.959	1.720	1.708	**1.608**
MAE	1.464	1.470	1.504	1.510	1.339	1.342	**1.289**
AIC	3.919	5.748	4.513	4.774	3.816	3.763	**3.334**
Av R^2	0.226	0.523	0.707	0.707	0.812	0.601	0.783

Table 5.5: Results CSFB/Tremont against SMB 80 month window

The evolution of beta (figures 5.12 and 5.13) shows however a mixed picture. While the evolution of beta with MRM random walk alpha is close to the one of OLS, beta UKF is more erratic and generally below the other ones. However, both curves show that exposure to small cap stocks was reduced after December 2000 and then slightly increased again after December 2002.

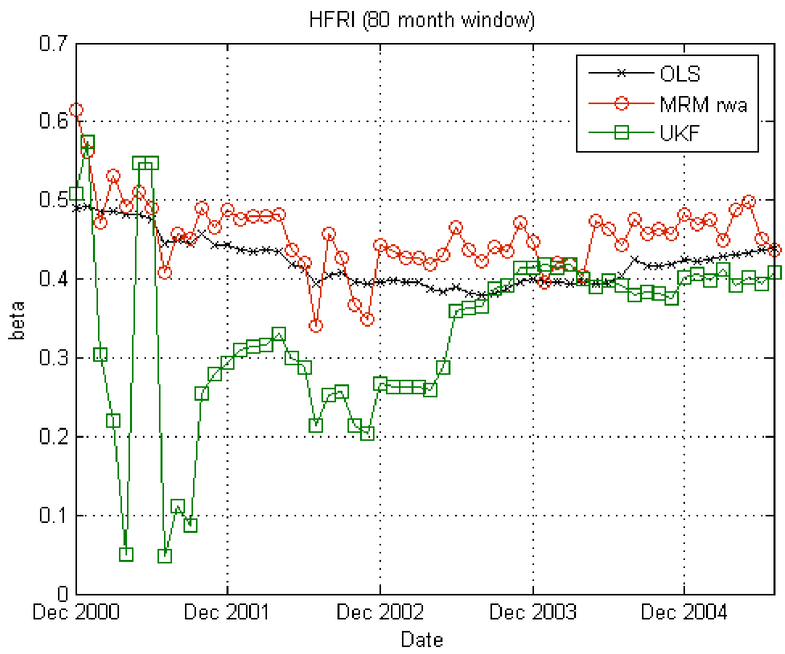

Figure 5.12: HFRI regressed against SMB: beta OLS, MRM with random walk alpha and MSZM/EKF

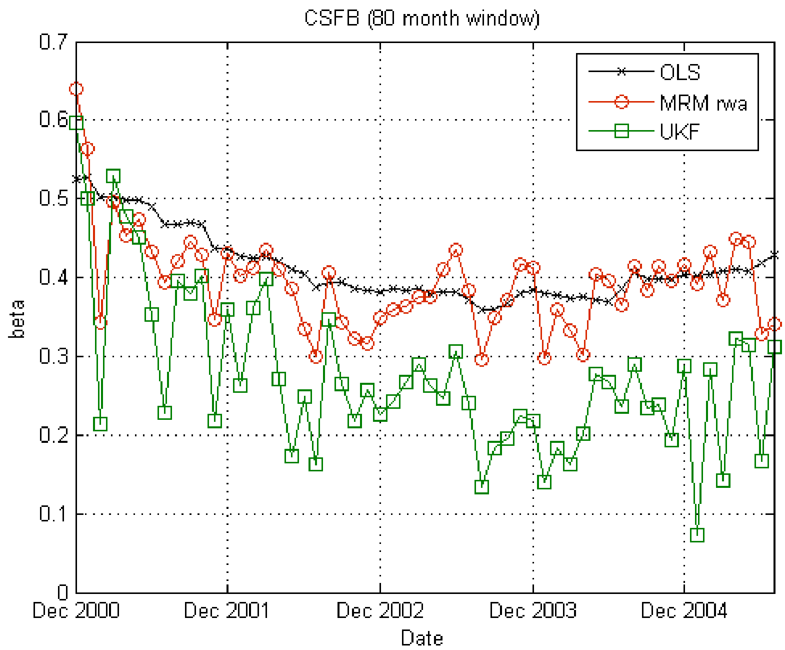

Figure 5.13: CSFB/Tremont regressed against SMB: beta OLS, MRM with random walk alpha and MSZM/UKF

5.5.4 Long/Short Equity against HML

High minus low (HML) characterises the return on a portfolio that buys high book-to-market stocks and sells stocks with a low book-to-market ratio (growth stocks). For this reason, $\beta_{HML,t}$ is high for portfolios of value stocks and negative for portfolio of growth stocks. Tables 5.6 and 5.7 show the results of a regression of HFRI respectively CSFB/Tremont against HML. In this case, RWM performs best in most cases, although the differences for HFRI are again rather small. Figures 5.14 and 5.15 show the comparison of β_{HML} for OLS, RWM and MSZM with EKF/UKF. The exposure measured with OLS is negative and fairly constant, indicating a constant exposure to growth stocks. Yet measured with MSZM/EKF and RWM, the exposure is constantly increasing, starting from a negative value and ending up in the positive area. This indicates that managers changed their exposure away from growth to value stocks.

	OLS	**RWM**	**RCM**	**MRM fixed α**	**MRM rw α**	**EKF**	**UKF**
RMSE	2.179	2.081	2.173	2.161	2.148	**2.067**	2.105
MAE	1.686	**1.638**	1.684	1.685	1.667	1.645	1.647
AIC	**5.107**	5.196	5.663	5.810	5.951	5.509	5.716
Av R^2	0.432	0.578	0.529	0.527	0.577	0.577	0.581

Table 5.6: Results HFRI against HML 80 month window

	OLS	**RWM**	**RCM**	**MRM fixed α**	**MRM rw α**	**EKF**	**UKF**
RMSE	2.340	**1.850**	2.328	2.165	2.083	2.125	2.101
MAE	1.764	**1.375**	1.761	1.545	1.527	1.620	1.580
AIC	5.888	**4.105**	6.501	5.829	5.597	5.824	5.695
Av R^2	0.387	0.702	0.756	0.724	0.736	0.790	0.766

Table 5.7: Results CSFB/Tremont against HML 80 month window

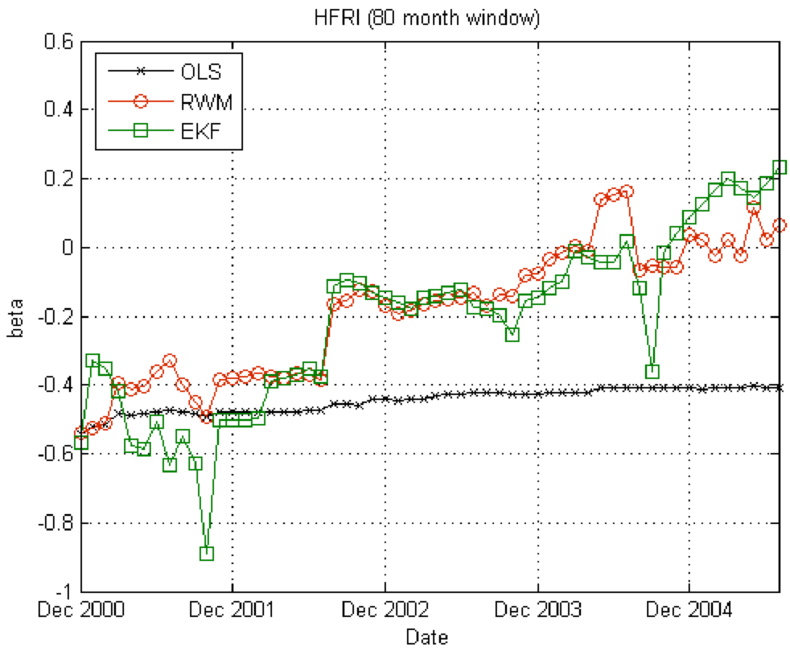

Figure 5.14: HFRI regressed against HML: beta OLS, RWM and MSZM/EKF

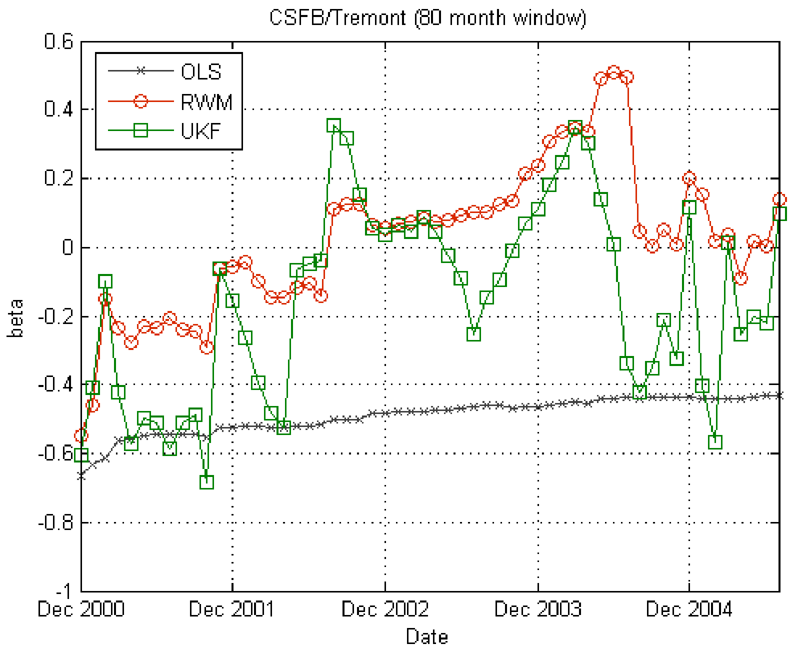

Figure 5.15: CSFB/Tremont regressed against HML: beta OLS, RWM and MSZM/UKF

5.5.5 Long/Short Equity against UMD

UMD is the return on a portfolio that buys stocks which have high past one-year returns and sells stocks with low past one-year returns. Thus, β_{UMD} is positive for a portfolio of stocks with high past one-year returns and negative for a portfolio with low past one-year returns. Tables 5.8 and 5.9 show the values resulting when HFRI respectively CSFB/Tremont is regressed against the Momentum Factor UMD. In this case, the best model on average is MRM with randowm walk alpha for CSFB/Tremont and depends on the selection criteria for HFRI.

	OLS	**RWM**	**RCM**	**MRM fixed α**	**MRM rw α**	**EKF**	**UKF**
RMSE	2.458	2.450	2.467	**2.087**	2.144	2.922	2.202
MAE	1.862	**1.470**	1.899	1.546	1.485	1.751	1.619
AIC	6.498	7.202	7.297	**5.420**	5.927	11.010	6.253
Av R^2	0.012	0.508	0.430	0.502	0.534	0.600	0.496

Table 5.8: Results HFRI agaist UMD 80 month window

	OLS	**RWM**	**RCM**	**MRM fixed α**	**MRM rw α**	**EKF**	**UKF**
RMSE	2.476	2.244	2.502	2.115	**2.037**	2.310	2.253
MAE	1.770	1.427	1.806	1.535	**1.460**	1.605	1.559
AIC	6.590	6.040	7.510	5.565	**5.352**	6.881	6.545
Av R^2	0.078	0.458	0.406	0.466	0.481	0.460	0.463

Table 5.9: Results CSFB/Tremont against UMD 80 month window

Average R^2 for OLS is quite low (0.01 for HFRI and 0.08 for CSFB/Tremont), which causes a low beta for OLS in figures 5.16 and 5.17. The graphs also indicate that although there are differences in values between the different models with Kalman filter, they all show a trend towards lower beta values in the first period and increasing values in the second period which come either close to the OLS value (CSFB/Tremont) or even exceed it (HFRI). For HFRI this indicates that at the beginning of 2001, managers reduced their exposure to a slightly negative value. Only after mid 2003, they started to shift their exposure back towards a positive exposure. For CSFB/Tremont, the same pattern is present, but to a lesser extent.

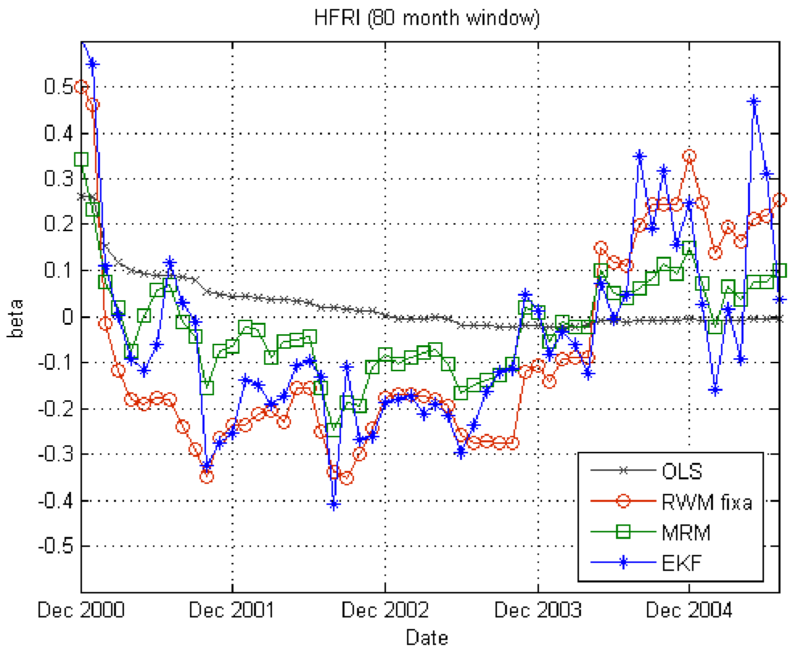

Figure 5.16: HFRI regressed against UMD: beta OLS, RWM, MRM with fixed alpha and MSZM/EKF

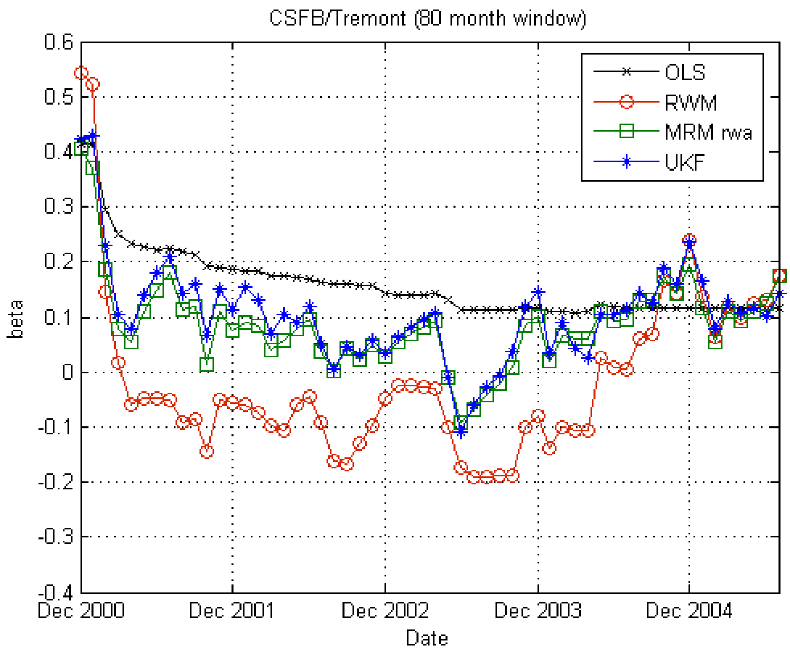

Figure 5.17: CSFB/Tremont regressed against UMD: beta OLS, RWM, MRM with random walk alpha and MSZM/UKF

5.6 Forecasts with Multivariate Regressions

In this section, we run now a multivariate regression of the Carhart model (equation 5.4). As mentioned in 5.3.2, the parameter estimation for the more complex models like MSZM is already difficult in the univariate case. As the multivariate case would substantially increase the numbers of parameters to estimate, we do not perform the regression for MSZM.

	OLS	**RWM**	**RCM**	**MRM fixed** α	**MRM rw** α
RMSE	1.031	1.123	1.060	1.060	**1.012**
MAE	0.816	**0.797**	0.830	0.830	0.824
AIC	**1.276**	1.882	1.869	2.163	2.045
Av R^2	0.832	0.941	0.946	0.950	0.967

Table 5.10: Results multivariate regression HFRI 80 month window

	OLS	**RWM**	**RCM**	**MRM fixed** α	**MRM rw** α
RMSE	1.435	**1.259**	1.429	1.430	1.432
MAE	1.146	**0.977**	1.141	1.149	1.153
AIC	2.381	**2.364**	3.396	3.933	4.095
Av R^2	0.807	**0.943**	0.968	0.974	0.977

Table 5.11: Results multivariate regression CSFB/Tremont 80 month window

Compared to the univariate cases and as expected, R^2 increases for all models, and the difference between OLS and the Kalman filter models is getting smaller. For CSFB/Tremont, RWM still provides better results than OLS while RCM and MRM are converging to OLS and provide almost the same result. The evolution of the betas for RWM is shown in figures 5.18 and 5.19. When compared with the evolution in the univariate case, the general form of the curve for CSFB/Tremont against S&P 500 is similar, although the extremes are less pronounced in the multivariate case. The same is true for HML and SMB.

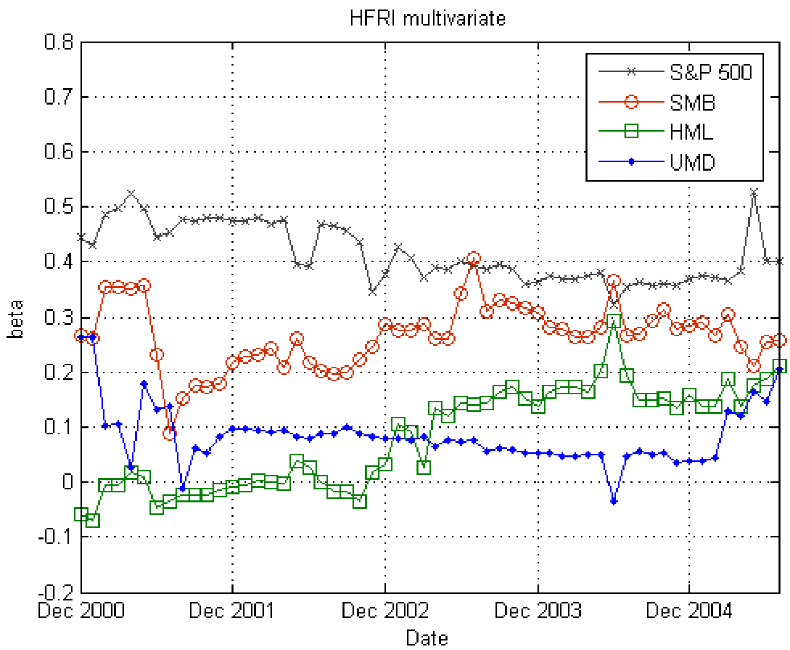

Figure 5.18: HFRI multivariate regression: betas RWM

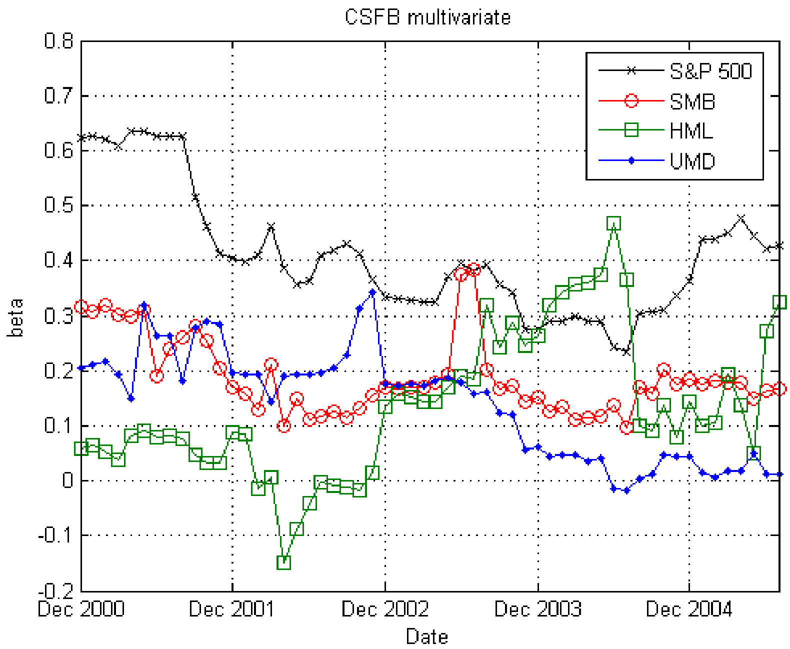

Figure 5.19: CSFB/Tremont multivariate regression: betas RWM

6 Conclusions

At the beginning of this thesis we claimed that dynamic asset allocations of hedge fund managers could not be captured using static regression methods. We argued that as a consequence of this, time-varying exposure of hedge funds as measured by beta would be computed incorrectly with these methods. In the following, we used different state space models which allow for time-varying abnormal returns (alpha) and exposures (beta). We then estimated the models for long/short equity styles by using Kalman filter respectively Extended and Unscented Kalman filter (EKF/UKF). Finally, the estimated models were evaluated in terms of their ability to forecast alpha and beta.

While we found that there is no explicit model that performs best in all cases, results show that with the exception of the Random Coefficient model (RCM), models with time-varying alpha and beta generally perform better than OLS. On average, the most complex model by Mamaysky et al. (MSZM) seems to perform best, especially when UKF is used as filter for estimation. Also models with mean-reverting beta (MRM) perform well in many cases, and even the relatively simple Random Walk model (RWM) is still superior to OLS in most cases. Since MSZM assumes dynamic management of portfolio-weights by the manager, the results hint that hedge fund exposures are actively managed. For the period analysed and the average of funds contained in the indices we obtain however a negative value for the MSZM parameter measuring the timing skills in the regression against S&P 500. This could underline the necessity of a thorough assessment of individual funds. Moreover, the performance of the different models also seems to vary depending on which market factor is used as regressor. This could indicate that managers use different dynamic strategies for different risk factors.

When the models are estimated with the multivariate four factor model originally introduced by Carhart, the advantages of models with time-varying exposure against OLS decline while the forecast error becomes smaller than in the univariate case for both OLS as well as time-variant models.

Concerning the evolution of beta over time, results indicate that time-varying models are better able to estimate the current value of beta than OLS. When we analyse the exposure against the market from 2001 to 2005, time-varying betas indicate that after the market decline of 2000, hedge fund managers reduced their exposure drastically, but started to increase it back after 2002, something that cannot be seen in the evolution of OLS beta.

As a summary, we conclude that models with time-varying alphas and betas provide better results than simple OLS in forecasting alpha and beta of long/short equity hedge fund indices and offer more insight into the dynamic evolution of exposure and hereby the behaviour of hedge fund managers. Moreover, results indicate that time-variation is at least partly caused by an active changement of portfolio-weights by the manager. This confirms that the value generated by hedge fund managers comes not only from their asset selection skills, but also from their dynamic management of the portfolio.

For further analysis, it would be interesting to reproduce above results on a cross section of individual hedge funds and with further hedge fund styles. Also, the application of Particle filters might improve the results as this class of filters is able to cope with non normal errors, although at the cost of increased complexity. Results however suggest that there is potential for more accurate estimation of time-varying exposure in hedge funds when time-varying exposure is considered in the estimation models.

Bibliography

Agarwal, V., and N. Y. Naik (2004): "Risks and Portfolio Decisions involving Hedge Funds," *Review of Financial Studies*, 17.

Arulampalam, M. S., S. Maskell, N. Gordon, and T. Clapp (2002): "A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking," *IEEE Transactions on Signal Processing*, 50(2).

Brealey, R. A., and E. Kaplanis (2001): "Changes in the Factor Exposure of Hedge Funds," *London Business School*.

Capocci, D. (2002): "An Analysis of Hedge Fund Performance 1984-2000," *Working paper, University of Liège*.

Carhart, M. (1997): "On persistence in mutual fund performance," *The Journal of Finance*, 52.

Ferson, W. E., and R. W. Schadt (1996): "Measuring Fund Strategy and Performance in Changing Economic Conditions," *The Journal of Finance*, 51(2).

French, K. R. (2005): "Kenneth R. French - Data Library," http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Fung, W., and D. A. Hsieh (1997): "Empirical Characteristics of Dynamic Trading Strategies: The Case of Hedge Funds," *The Review of Financial Studies*, 10(2).

——— (2003): "Estimating the Risk Premium in Equity Long/Short Hedge Funds," *Working Paper, London Business School and Duke University*.

——— (2004a): "Extracting Portable Alphas from Equity Long/Short Hedge Funds," *Journal of Investment Management*, 2(4).

——— (2004b): "Hedge Fund Benchmarks: A Risk Based Approach," *Financial Analysts Journal*, 60(5).

GÉHIN, W., AND M. VAISSIÉ (2005): "The right place for alternative betas in hedge fund performance: an answer to the capacity effect fantasy," *Edhec Risk and Asset Management Research Center.*

HARVEY, A. C. (1989): *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge University Press.

JAEGER, L., AND C. WAGNER (2005): "Factor Modelling and Benchmarking of Hedge Funds: Are passive investments in hedge funds possible?," *Working paper, Partners Group, Baar-Zug.*

JAVAHERI, A., D. LAUTIER, AND A. GALLI (1980): "Filtering in Finance," *WILMOTT magazine*, 45.

KAT, H. M., AND J. MIFFRE (2003): "Performance Evaluation and Conditioning Information: The Case of Hedge Funds," *Working Paper # 0006, Cass Business School, City University London.*

KRAIL, R. J., C. S. ASNESS, AND J. M. LIEW (2001): "Do Hedge Funds Hedge?," *The Journal of Portfolio Management.*

MAMAYSKY, H., M. SPIEGEL, AND H. ZHANG (2003): "Estimating the Dynamics of Mutual Fund Alphas and Betas," *Yale ICF Working Paper 03-03.*

MCGUIRE, P., E. REMOLONA, AND K. TSATSARONIS (2005): "Time-varying exposures and leverage in hedge funds," *BIS Quarterly Review.*

SHARPE, W. F. (1992): "Asset Allocation: Management Style and Performance Measurement," *Journal of Portfolio Management.*

VAN DER MERWE, R., A. DOUCET, N. DE FREITAS, AND E. WAN (2000): "The Unscented Particle Filter," *Technical Report CUED/F-INFENG/TR380, Cambridge University Engineering Department.*

WELCH, G., AND G. BISHOP (2001): "An Introduction to the Kalman Filter," *ACM Computer Graphics (SIGGRAPH'2001).*

WELLS, C. (1996): *The Kalman Filter in Finance*, vol. 32. Kluwer Academic Publishers.

YAO, J., AND J. GAO (2004): "Computer-Intensive Time-Varying Model Approach to the Systematic Risk of Australian Industrial Stock Returns," *Australian Journal of Management*, 29(1).