PREDICTIVE ABILITY OF AUTOREGRESSIVE CONDITIONAL BETA MODELS FOR REALIZED VOLATILITIES

Javier Moles Palomar

Trabajo de investigación 25/007

Master en Banca y Finanzas Cuantitativas

Tutor: Dr. Carles Bretó Martínez

Universidad Complutense de Madrid

Universidad del País Vasco

Universidad de Valencia

Universidad de Castilla-La Mancha

www.finanzascuantitativas.com

Predictive Ability of Autoregressive Conditional Beta models for Realized Volatilities

Master's Degree in Banking and Quantitative Finance

Author

Javier Molés Palomero

Tutor

Carles Bretó Martínez

June 2025

Abstract

This thesis carries out an extended investigation of the modeling of conditional volatility by score-driven models, with a particular focus on the Autoregressive Conditional Beta framework developed in Blasques et al. (2024). We also consider several other models and fit them to multiple return data using factor regressors. We filter volatilities that we then use to build a set of models that are statistically similar in predictive ability. We find that Autoregressive Conditional Beta models excel at conditional volatility modeling: while other models can be considered equal in predictive ability, Autoregressive Conditional Beta models are systematically ranked first.

Keywords: conditional volatility, factor models, score-driven models, Autoregressive Conditional Beta; predictive ability comparison; Model Confidence Set; Fama-French.

JEL Codes: C01, C14, C22, C58, C61, G17, .

Contents

List of Figures

List of Tables

1 Introduction

Statistical modeling in finance has advanced significantly in the incorporation of variable dynamics, for example, with the use of Score-Driven or Quasi-Score-Driven models. These models establish the notion of not only score-driven time-varying volatilites but as of recently also take into consideration time-dynamic slope coefficients (betas). The use of these betas has been evaluated mostly for the conditional means of data, overlooking the dynamics and effects regarding conditional variances or volatilities. A high-frequency analysis of realized volatility can capture rapid market movements and potentially enhance the precision of models such as the Autoregressive Conditional Beta (ACB) model. However, the ACB model entails higher computational costs than alternative models. In this thesis, we aim to expand on the ACB model by studying its predictive ability for realized volatilities and comparing it to that of other models by use of the Model Confidence Set procedure for predictive ability. We find that the ACB model outperforms alternative models, for standard confidence levels, in terms of predictive ability.

The building of statistical models serves as a way to bring out the intricate inner-workings of the markets in a way that is understandable, quantifiable and useful for certain purposes and decision making. Improvements in these aspects of statistical models coming from fields like econometrics are especially useful in these times of technological growth, computing and artificial intelligence. We have come a long way in the analysis of time series from AR models (Box et al., 2013) to established frameworks like GARCH (Bollerslev, 1986; Engle, 1982; Francq and Zakoian, 2019; Francq and Zakoïan, 2012, 2015), etc. However, all these share a common weakness, that is, the assumption that model parameters themselves remain constant over time. This is why recent advances in financial modeling (Creal et al., 2013) revolve around time-varying parameters (Koopman, 2012; Durbin and Koopman, 2012; Harvey, 2002; Cipra, 2020), as such variation is consistent with what one sees in the real world. One example of such advancements is Blasques' ACB framework.

Blasques' recent work (Blasques et al., 2024) is a great step in the incorporation of variable dynamics. In this work, this is done via the use of a Score-Driven model (Creal et al., 2013), making use of an updating gradient equation to adjust parameters over time. More specifically, to create time-varying slope coefficients, referred to as betas (Sharpe, 1964; Fama and MacBeth, 1973; Fama and French, 1993, 2015) for a regression to reflect the evolving sensibility of an asset's returns to several risk factors. Blasques' betas are part of the conditional mean, the expected value of a regression model given the data and Blasques' work is mostly centered on establishing the ACB model itself and its properties, as well as on the modeling and practical applications of the ACB model, as well as how it performs against other models.

The problem that we want to tackle is that the matter of conditional volatility has not been properly looked into within this framework. Blasques' paper delves into detail in all matters regarding the conditional mean but there are barely any results for the conditional variance, which is just as central to financial analysis and on which we plan on expanding. Our contribution to solving this shortcoming of Blasques' work is to replicate Blasques' analysis to then extend it to predict realized volatility (Andersen and Teräsvirta, 2009; Da and Xiu, 2021; Zhang et al., 2005). We then test the predictive ability (Giacomini and White, 2006; Diebold and Mariano, 2002; Hansen et al., 2011) of the ACB model against alternative models, highlighting the better fit of the ACB model as measured by favorable log-likelihood ratios.

Solving this problem is significant since the ACB model (as well as score-driven models in general) are not as straightforward as existing alternative models. By expanding on the limited existing results regarding ACB models with results on volatility, we hope to yield a deeper understanding of ACB models, with the ultimate goal of identifying shortcomings and weaknesses of ACB models. Another goal is to contribute to the spread of ACB models, considering that the models' benefits have already been hinted at (Blasques et al., 2024). Failing to consider the dynamics in ACB models could result in less effective hedging strategies and risk management.

The value of this thesis is thus the additional investigation of the use of Score-Driven models for predicting 5-min. realized volatility with tests for predictive ability between models. Furthermore, we emphasize that ACB models might lead to significant improvements in log-likelihood when compared to models that keep betas constant.

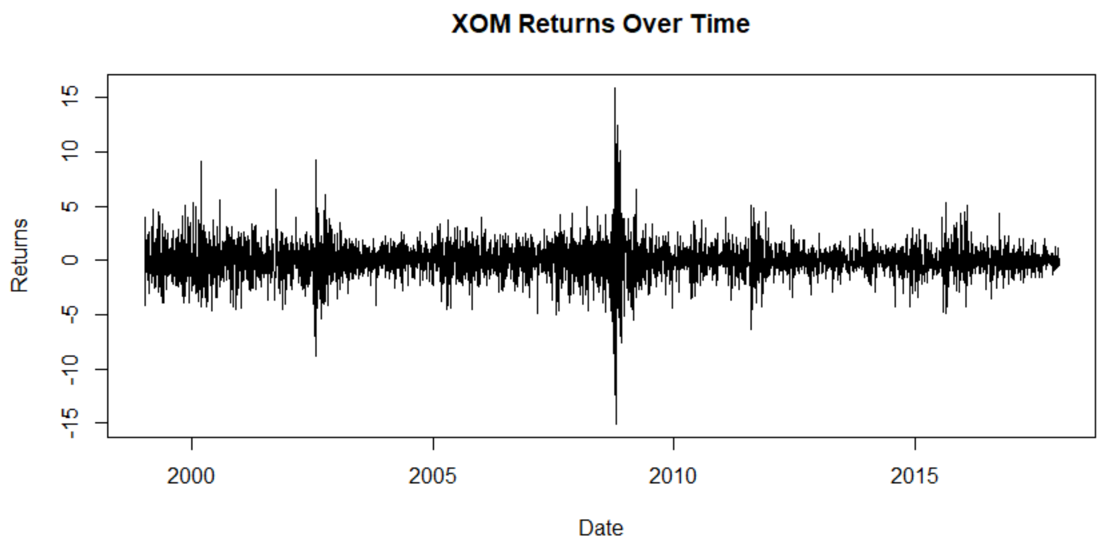
2 Data

2.1 Assets

2.1.1 S&P 500

Blasques in his paper takes 16 US stocks (randomly selected among the most liquid ones) from the S&P 500, we have chosen a selection based on the ones that appear in his paper, though it is not entirely irrelevant since like we will explain later, not all of the stocks in his article were chosen in the same time window. The ones we have selected for our paper include ExxonMobil Corporation (XOM) American Express Company (AXP), The Boeing Company (BA), Microsoft Corporation (MSFT), McDonald's Corporation (MCD), and Pfizer (PFE). Like in standard practice, the actual values we will use in the empirical process will be the logarithmic returns.

Figure 1: XOM logarithmic returns



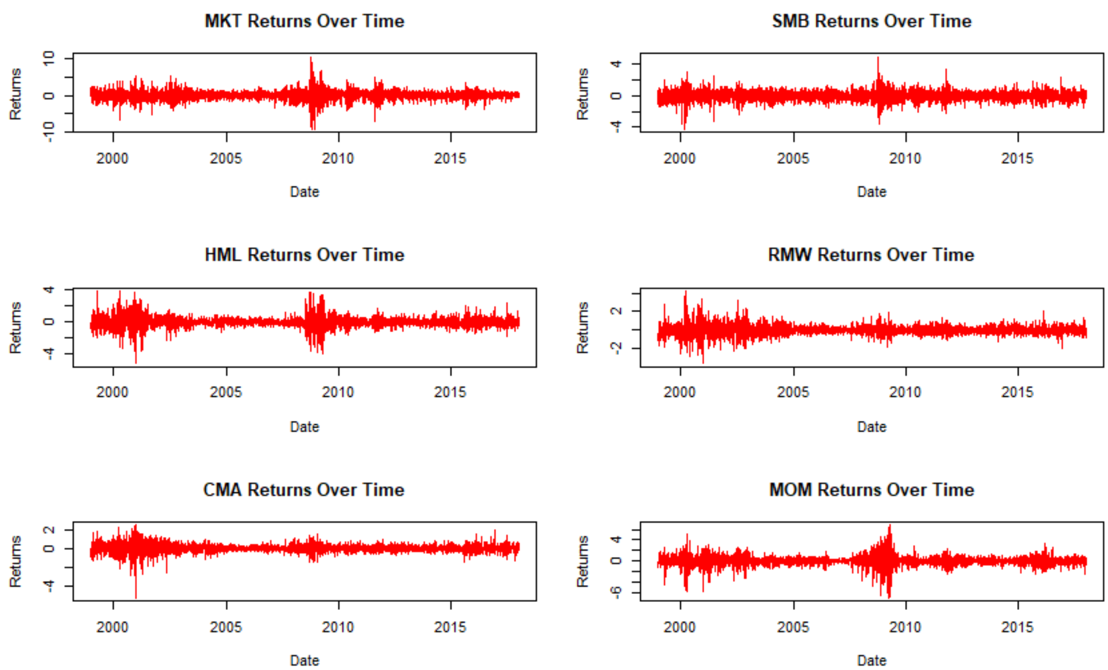
2.1.2 Factors

Factor models are widely used in finance to price an asset's returns. The beginnings of factor models go back to the Capital Asset Pricing Model (CAPM) by Sharpe (1964) who proposed a single factor model, this being the market risk, represented as a single beta that drives the asset's expected return. Over the years, new factors have been studied to improve the single beta model, most famously the theory developed by Fama and French (Fama and MacBeth, 1973; Fama and French, 1993), which led to the creation of multifactor models. In particular, the three-factor model (Fama and French, 1993), which expands on the risk factors that explain asset returns. Over time, more factors such as momentum (Carhart, 1997) are included in these types of models.

We will be taking six factors for the betas, the exact same ones Blasques uses in his paper, the risk factors of the three-factor model of Fama and French (1993), i.e., the market factor represented by the log-returns on the S&P500 index, the standard Fama-French size, the (Small Minus Big) and (High Minus Low) value factors. The other three factors are the (Robust Minus Weak) and (Conservative Minus Aggressive) factors initially proposed by Fama and French (2015) as well as (Momentum) as in Carhart (1997), we obtain the data for these factors, from Dacheng's website[1], these factors come in arithmetic returns formed from portfolios that replicate each of these factors in 5 minute intervals from 9:30 to 16:00 hours, turning these into log returns and adding them by making use of the log properties we are able to obtain the daily log returns for all 6 factors that we will be using as our regressors, we can see this in figure 2.

[1]https://dachxiu.chicagobooth.edu/

Figure 2: Factor returns over time



2.2 Realized Volatility

Realized volatility is a financial metric that measures the price fluctuations of an asset over a specific time interval, since the underlying latent time-continuous variance is unobservable (Andersen and Bollerslev, 1998). Introduced in Andersen and Bollerslev (1998), the most basic measure for realized volatility as a non-parametric measure is the sum of squared returns. This estimator however, is subject to the market's microstructure noise (Andersen and Teräsvirta, 2009; Zhang et al., 2005), thus over time more research has gone into developing noise-robust estimators. Inference through parametric approaches have been used to improve upon the modeling of realized volatility like the use of ARCH, stochastic volatility models, etc.(Andersen and Teräsvirta, 2009). More recently a paper by Da and Xiu (2021) builds a robust likelihood-based approach for realized volatility that serves as the focus for our work.

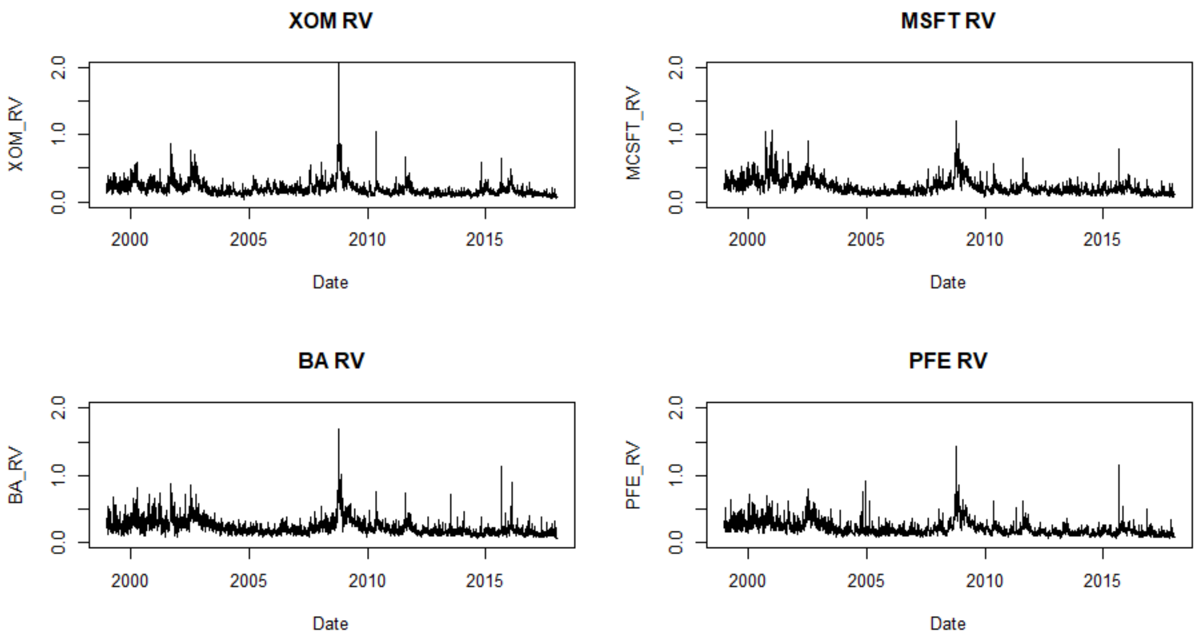
One of the cornerstones of this thesis is the combined use of Blasques' model for time-varying conditional betas, with Dacheng's work on realized volatility (Da and Xiu, 2021). The concept of realized volatility appears when looking into higher frequencies of trading. It is defined as the changes in price of an asset over a particular amount of time and taking it into account can be very important since it measures the intraday volatility of a market day in which even if the price of the stock started and ended on the same price, heavy fluctuations could have happened in-between. This is useful for many reasons, including volatility forecasting and forecast evaluation, which is what we will be doing.

That said, Dacheng's work contributes a way to measure this type of volatility in such a way that helps remove noise from the estimations. They assume that the observed transaction price follows a time-continuous Itô-semimartingale and estimate volatility, defined as the integrated variance of said semimartingale, by maximizing the likelihood of a moving-average model whose order is based on an information criterion (Da and Xiu, 2021).

Unlike the standard non-parametric ways of obtaining realized volatility, which like we said is usually just the addition of squared log returns over the desired period of time, their approach consists on using a quasi-maximum-likelihood estimator by fitting a misspecified parametric Gaussian moving average model MA(q) to the vector of high-frequency returns; pretending that the logarithm of the efficient price is a Brownian motion with constant volatility but no drift (Da and Xiu, 2021), with the noise following said Gaussian MA(q) model of q order to be determined using information criteria.

Their work on realized volatility is extremely helpful since the data is readily available in Dacheng Xiu's website as well, that is where we obtain our time series data for realized volatility that we will use in the forecasting process. We download the realized volatility series for the same selected stocks under the same time window we specify below. We obtained the realized volatilty series for our selected stocks with the added 5-min. realized volatility figures which do not need any extra work done on our part before being used for our purposes.

Figure 3: Realized Volatility time series for XOM, MSFT, BA and PFE



We see that generally XOM seems to have higher realized volatility peaks in the time frame around the 2008 financial crisis, while the other stocks remain generally and steadily below this, with MSFT appearing the least volatile. We must also note that there is small discrepancy in terms of scale between RV and our filtered results.

2.3 Time Windows

We have obtained data in a similar manner to Blasques' paper, since much of our work consists on replicating his work and extending it, so we have data for the period starting in January 1999 to December 2017, in Blasques' article he mentioned a reduced time window for a few of the stocks but since we are not using the complete list of 16 stocks we will keep to just this particular window which would come at 4780 observations, 4779 since we account for losing one data point when building the log returns (like we mentioned previously, we aggregated the 5-minute log-returns on both the stocks and the factors).

3 Methodology

The estimation process for the ACB model is mainly carried out by a multistep QML method with the following steps:

- First, fit and estimate by standard Quasi-Maximum Likelihood (QML) the univariate GARCH (1,1) parameters $(\mu_{0i}, \omega_{0i}, \alpha_{0i}, \beta_{0i})$, for $x_{i,t}$, so that we can obtain the conditional mean $\mu_{i,t}$ and variance $g_{i,t}^2$. Requiring an initial value [$\tilde{g}_{i_1}^2 = \tilde{g}^2 \geq 0$] starting the sum at t = 2.

- The next step is to estimate the remaining parameters of the ACB model, specifically for the betas. Again, by minimizing a function for all its parameters via QML, we use the Nelder-Mead algorithm for the optimization process. For the starting values, we use those that bring the model closer to a constant beta, that is, $\xi_i \approx 0$, $c_i = 1$, and $\varpi_i \approx 0$. For the betas' starting values we use the ones obtained from the GARCH model with factors, specifically the 'mxreg' values for the betas and the 'mu' for the intercept. We say the model is complex because the optimization process is heavy on computer resources, since it requires fitting 32 parameters for each stock, it is very time consuming. This is why we want to compare this model with others in terms of performance to see if we can achieve similar results to avoid the heavy computational costs.

- In terms of code, this implementation in R is run with the packages "rugarch" for the GARCH fitting process, the packages "gas" or "gasmodel" provide the tools and algorithms for working with score-driven models though we found it really hard to use those for the ACB model, since we could not exactly specify the distribution of the mean and variance needed for the ACB model within those packages, so instead we run an optimization function built just for optimizing the ACB parameters via the QML objective.

3.1 Models

3.1.1 Generalized Autoregressive Score

Our work in this paper and all results stem from an interest in the Score-Driven models, a framework initially developed and published by (Creal et al., 2013) who wanted to extend the research into the modelling of multivariate time-varying volatilities and parameters, from their study we obtain the most basic form of the score-driven, time-varying updating equation for the factor from which all other papers, including ACB, base their work on:

$$f_{t+1} = \omega + \sum_{i=1}^{p} A_i s_{t-i+1} + \sum_{j=1}^{q} B_j f_{t-j+1},$$

Here ω is a vector of constants, A_i and B_j are coefficient matrices with appropriate dimensions i=1,...p and j=1,...q respectively. s_t is the scale function of current and past data, which is formed as a derivative of the density function at time t with respect to the parameter vector f_t. We can see below that the scaling function of the Generalized Autoregressive Score (GAS) model is formed by two components, a scaling matrix S_t and a ∇_t derivative of the log-conditional density with respect to the parameter. Given the data up to time t, f_{t+1} is known since it is a deterministic function (Creal et al., 2011).

$$s_t = S_t \cdot \nabla_t, \quad \nabla_t = \frac{\partial \ln p(\nu_t | f_t, \mathcal{F}_t; \theta)}{\partial f_t}, \quad S_t = S(t, f_t, \mathcal{F}_t; \theta)$$

The choice of S_t scaling matrix can come from several options and will all lead to different GAS models, for example the most usual option is to set S_t as the inverse of the Fisher information matrix which leads the GAS model to a closer, more familiar, GARCH specification; different specifications of this function lead to different models like for example using the square Inverse information of Fisher (Creal et al., 2011). It is relevant to mention how Creal's GAS model framework is a more general expression of other models, for example if p=q=1 and the p density is Normal, with mean 0 and the f time-varying factor is the variance, then it causes the GAS recursion to become the standard GARCH(1,1) model.

$$f_{t+1} = \omega + A_1(y_t^2 - f_t) + B_1 f_t,$$

The methodology for estimating the GAS model is relatively simple, since it is an observation driven model a convenient property of them is the straightforward estimation of the parameters via maximum likelihood, by putting the log-likelihood contribution in a maximizing algorithm and adding the sum. The GAS specification of the parameters adapts naturally to different configurations of the density and parameters. If instead of making a model on dynamic volatility we want to use one to measure betas of a regression, the GARCH dynamics will automatically adapt (Creal et al., 2011).

All this being said, the GAS model provides us with a unified likelihood-based framework for designing many different time series models. By using the score of the density it leads to efficient and robusts time-varying dynamic updates, and it is able to handle varied recursions like heavy tails, asymmetry, skewness, etc. And since the likelihood is closed form the estimation via quasi-maximum likelihood is straightforward. However the shortcomings of this framework comes when working in high-dimensional multivariate settings, since this implies having to estimate an exponentially growing number of parameters which takes a toll on computational resources, a limitation shared by similar observation-driven approaches.

3.1.1.1 Autoregressive Conditional Betas

The model that this paper is mostly centered around is Blasques' particular case for GAS, based on a structure to form time-varying betas for a regression of six factors. The updating equations are the key aspect of the GAS for Autoregressive Conditional Betas model and are closely related to the scaling function for each of these betas which are composed of an intercept, a score-driven function formed by scale and score, the derivative of the log-likelihood contribution which in this case it is obtained from a Gaussian distribution, multiplied by a parameter of magnitude, and finally an autoregressive component. All of these parameters are estimated at the same time by a QML estimator, in our work.

The model is a regression model with score-driven time-varying betas. The model accounts for heteroskedastic errors and stochastic regressors which display GARCH dynamics, these are assumed to be formed by a conditional mean and a conditional variance affected by an i.i.d. random innovation with mean 0 and variance of 1, we can apply this structure to filter said innovations out after estimating the mean and variance of each of the regressors.

If we look at the model's structure we see several components, the first is the equation for the stock return y_t, this being the sum of betas times their factors plus the error term, secondly there is the structure of the equation for the beta itself for a given stock i and time t: the ϖ_i intercept followed by the score function $\frac{\nu_t(\varphi)x_{i,t}}{\mu_i^2+g_{i,t}^2(\varphi)}$ this is the key element that updates the dynamic beta over time, this comes from a derivative of the log-likelihood contribution given by $l_t = -\frac{1}{2}\{ \frac{\nu_t^2}{g_t^2} + \log\big(g_t^2\big)\}$ such that the score function obtained is $S\big(\beta_{i,t}\big) \frac{\partial l_t}{\partial \beta_{i,t}} = \frac{\nu_{i,t}\, x_{i,t}}{\mu_i^2 + \sigma_{i,t}^2}$ the final component is the autoregressive element of the equation multiplied by the c_i parameter for the magnitude of effect of the previous beta.

$$y_t = \beta_{1,t}\, x_{1,t} + \cdots + \beta_{p,t}\, x_{p,t} + \nu_t, \quad \nu_t = g_t\, n_t, \tag{1}$$

$$\beta_{i,t+1} = \varpi_i + \xi_i \frac{\nu_t\, x_{i,t}}{\mu_i^2 + g_{i,t}^2} + c_i\, \beta_{i,t} + \gamma_{1,i}\, z_{1,t} + \cdots + \gamma_{Q,i}\, z_{Q,t}, \tag{2}$$

$$g_{t+1}^2 = \omega + \alpha\, \nu_t^2 + \beta\, g_{i,t}^2, \tag{3}$$

$$x_{i,t} = \mu_i + g_{i,t}\, \varepsilon_{i,t}. \tag{4}$$

$$g_{i,t+1}^2 = \omega_i + \alpha_i (x_{i,t} - \mu_i)^2 + \beta_i g_{i,t}^2. \tag{5}$$

The final equations (3) and (5) represent GARCH models, however these are not the same and represent different volatilities , the one in equation (5) represents individual GARCH(1,1) models for each of the i factors, this difference is important since this one uses a mean value in the process while the model in equation (3) does not, it is this mean value that appears in the score function of the beta as well as the equation for modeling the factor, which can serve to filter the i.i.d. innovations, so it is not redundant and that mean value will vary for each of the factors. The GARCH model in equation (3) is more standard and only changes between stocks, only lacking a mean value as opposed to the previous model, this GARCH serves us as a way to have an equation to filter the volatilities obtained by the beta process.

In Blasques et al. (2024) the practical applications of this model in hedging and risk management are also explored, wanting to find a strategy to track US stocks using the six-factor model. In the context asset pricing, expected returns on any asset are linear in the betas and only depend upon the risk premiums in the factors (Blasques et al., 2024). Blasques performs a tracking exercise consisting of taking a position at time t in each of the 6 considered factors coming from a step-ahead forecast of the corresponding conditional beta, using each model of each stock to build a different hedging portfolio.

Like we can see and as we mentioned before here, Blasques only explores the implications of this model for conditional means, since the tracking value of each stock at time t is formed only by the aggregate values of the betas times their regressor at time t which, like betas in a regression, are basically the mean of the effect of said factor on the price of the stock. We can see this in the equation below, what Blasques finds in his hedging exercise is that his ACB model is the best model solely based on daily data, much better in function of the tracking errors.

$$
\begin{aligned}
Z_{t+1|t} = {} & \beta_{MKT,t+1|t} MKT_{t+1} + \beta_{SMB,t+1|t} SMB_{t+1} + \beta_{HML,t+1|t} HML_{t+1} \\
& + \beta_{RMW,t+1|t} RMW_{t+1} + \beta_{CMA,t+1|t} CMA_{t+1} + \beta_{MOM,t+1|t} MOM_{t+1}
\end{aligned}
$$

The updating equation is derived from the conditional density of the assumed distribution that the model follows, these equations are made up of a scaling function $S(ft)$ and an $st(ft)$ partial derivative of the log-likelihood of the conditional density for the score-driven model. For the ACB model, this can be seen as a quasi-score-driven model where the update comes from a loss function of a Gaussian log-likelihood, this is regardless of the distribution for the regression error/innovation, which is i.i.d. N(0,1). The basic functionality of this update comes from checking if the beta is either under or overvalued, if the first case occurs then the update will tend to make $\beta_{i,t+1}$ increase.

The scaling function represents the score of the predictive likelihood with respect to the parameter, in this case, the beta. In the literature it is noted that this scaling function should be greater than 0 for the optimal results to hold up, however this is not the case for the current multivariate configuration (Blasques et al., 2024) . Aside from the one mentioned already, several different scaling functions can be used for the same objective, these being identity scaling, inverse information scaling, square root inverse information, etc.

In his paper, Blasques establishes the stationarity of the data the ACB model generates as well as the invertibility of the filters for both the time-varying conditional betas and the conditional volatilities. They note that the conditional volatilities of the regressors do not depend on other filtered parameters, the conditonal betas depend on the filtered conditional volatilities of the regressor only, and the conditional volatility of the observation equation's error depends on all the other filters (Blasques et al., 2024). So they showcase that there exists a solution to the time series such that the error term of the regressors and their univariate volatilities are stationary and ergodic.

Their proof of the model's invertibility is important since it ensures that the model is uniformly invertible, meaning that the filter of the time-varying parameters converge almost surely and exponentially fast to a unique limit solution regardless of the initialization of the filter, and that this result holds uniformly over the parameter space (Blasques et al., 2024), a property that is essential to find a consistent estimator for the parameters.

3.1.1.2 Beta t E-GARCH

We have also decided to include the Beta-Skew-t-EGARCH model, proposed by (Harvey and Chakravarty, 2008), like its name indicates, it is a way to take into account volatility clustering, since it is typical for episodes of high volatility to be followed by equally high volatilities, and the opposite for the episodes of low volatilities; thus this model, which also contains a score-driven updating component is a good way to incorporate this skewness. The model presents several helpful characteristics, like robustness to jumps as well as outliers, in addition, it also accounts for heavy tails and skewness in the conditional return and for a leverage and a time-varying long-term component in the volatility specification (Sucarrat, 2013). In our case we will be using a regular Beta t E-GARCH estimated using only the stock data, but also one that is estimated on the residuals of a linear regression using all the six factors, that way we incorporate all the data to observe whether it is significant when predicting realized volatility.

Like with the ACB model, which we base much of our work on, the Beta-t-EGARCH model can be viewed as an unrestricted version of the Generalised Autoregressive Score model of Creal, which means there is a score-driven updating component in the model's framework, u_t represents said conditional score, the derivative of the log-likelihood of y_t with respect to λ_t.

$$
\begin{aligned}
&y_t = \exp(\lambda_t)\varepsilon_t = \sigma_\varepsilon \varepsilon_t, \quad \varepsilon_t \sim \text{st}(0, \sigma_\varepsilon^2, \nu, \gamma), \quad \nu > 2, \quad \gamma \in (0, \infty), \\
&\lambda_t = \omega + \lambda_t^\dagger, \\
&\lambda_t^\dagger = \phi_1 \lambda_{t-1}^\dagger + \kappa_1 u_{t-1} + \kappa^* \text{sgn}(-y_{t-1})(u_{t-1} + 1), \quad |\phi_1| < 1.
\end{aligned}
$$

This equation forms the structure for the single component Beta t E-GARCH, the y_t is formed by an exponential of the lambda component times a ε_t i.i.d. innovation that follows a t-student distribution with ν degrees of freedom, a mean μ component may be included but in much of the literature plus our own work it will appear as 0 (Harvey and Sucarrat, 2014). The λ_t parameter inside the exponential is formed by a ω constant plus a $\lambda^\dagger$ parameter which contains an autoregressive component, the u_t conditional score component; leverage is introduced into the equation with the multiplied κ^* times sgn function which turns into +1 or -1 according to de sign of the resulting value inside the function or

0 if it results in 0. All model parameters are estimated by maximizing the log-likelihood of the observed data under the Student-t density (Harvey and Sucarrat, 2014). In our work we use the R package 'betategarch' for the model estimation, both regular and with the factor regressors(Sucarrat, 2013).

The conditional score function of the Beta t E-GARCH model takes the following form, and it comes from the conditional derivative of the log-likelihood under the t-student's distribution. u_t is bounded in between [-1, ν] and has mean 0, with the ν number of degrees of freedom being positive (Harvey and Sucarrat, 2014). Another aspect to note is the fact that adding the leverage term or not does not change the score function, since it is only one extra piece in the dynamic equation and has no effect on the likelihood from which the score is derived from, the leverage term only affects how λ_t is updated.

$$u_t = \frac{(\nu+1)(y_t-\mu)^2}{\nu\exp(2\lambda_{t,t-1})+(y_t-\mu)^2} - 1$$

3.1.2 GARCH

3.1.2.1 Standard

A standard GARCH (1,1) model estimated via the 'rugarch' package, though simple, it will serve as a baseline for estimating volatility compared to other models. One thing to keep in mind is that in the process of estimating the ACB model we also make use of two GARCH(1,1) models, these however are not correlated to just using a standard GARCH model for filtering and predicting the realized volatility, which is what we do here, in this case we simply run the standard structure of the generalized autoregressive heteroskedasticity. This model also lends itself to be estimated by Quasi-Maximum Likelihood in a Gaussian structure.

$$\sigma_t^2 = \omega + \alpha_1(\varepsilon_{t-1}-\mu)^2 + \beta_1\sigma_{t-1}^2$$

3.1.2.2 Factors

A GARCH(1,1) model estimated with the six factors taken into account by including them as external regressors in the algorithm of the model. This way we are taking into account the five Fama-French factors as well as the momentum for a new model, which since it includes the same amount of data as the ACB model we have been working on, it should reflect dynamics that represent the realized volatility more accurately when we run it through the contrasts for equal predictive ability. This would provide similar results as estimating a linear regression with all six factors and then using the residuals to estimate the GARCH (1,1) model to then filter the volatilities from it, as we will see next the process for OLS will be slightly different.

$$\sigma_t^2 = \omega + \alpha_1(\varepsilon_{t-1}-\mu)^2 + \beta_1\sigma_{t-1}^2$$

where

$$\mu = \beta_1 x_{1,t} + \cdots + \beta_6 x_{6,t}$$

3.1.3 OLS

A standard linear regression via ordinary least squares, the standard error of this linear regression is estimated after this, like we mentioned previously filtering volatilities from a GARCH(1,1) trained on the residuals of this regressions would yield similar results to the GARCH(1,1) estimated on the six factors so what we are doing here differently is obtain the singular standard error value after estimating the model and then assign said error as a constant sigma derived from this process. This result seems unrealistic and contrary to what we want to study, which is the dynamics of volatility over time, but it still works as the most basic form of estimation for volatility, and as we will see in the results, is not too out of the question when comparing the predictive ability.

$$y_t = \beta_1 x_{1,t} + \cdots + \beta_6 x_{6,t}$$

$$\sigma_t = \sqrt{\frac{1}{n-2} \cdot \frac{\sum (y_i - \hat{y}_i)^2}{\sum (x_i - \bar{x})^2}}$$

4 Results

Here we will begin displaying the results obtained from our endeavors, the main results come in the form of the filtered volatilities from the diverse range of models we have considered, in the case of the ACB model, this also includes the results of the filtered time-varying betas for several stocks in a similar way to how Blasques presented them in his paper. The other main kind of results are those derived from the Model Confidence Set for predictive ability contrast, whose details will be extended further below.

4.1 Filtering

4.1.1 Betas

The first objective of our work was to successfully replicate the betas that appear in Blasques' ACB paper, in his paper, he presents the results for all 6 factor betas plus an intercept component, since Blasques in his paper provided results for the parameters of 1 stock, it was straightforward to build a function that filtered the betas and volatility according to the structure of the equation which we see below, following an observation-based structure. Here are the betas for EXXON obtained by Blasques followed by our replication:

$$\hat{\beta}_{i,t+1} = \hat{\varpi}_i + \hat{\xi}_i \frac{\hat{\nu}_t x_{i,t}}{\hat{\mu}_i^2 + \hat{g}_{i,t}^2} + \hat{c}_i \hat{\beta}_{i,t}$$

The challenge was obtaining the same results for other stocks since his paper does not provide results for all of them, we need not only betas that behave similarly but also a comparable log-likelihood value, like we mentioned previously, no R package provided us with a clear way to estimate Blasques' model so we created a minimizing function by hand that would optimize all 32 parameters of a model for 1 stock, which would also compile the log-likelihood contributions as a way to estimate how close we were to the actual results.

Figure 4: Original betas obtained for Exxon by Blasques

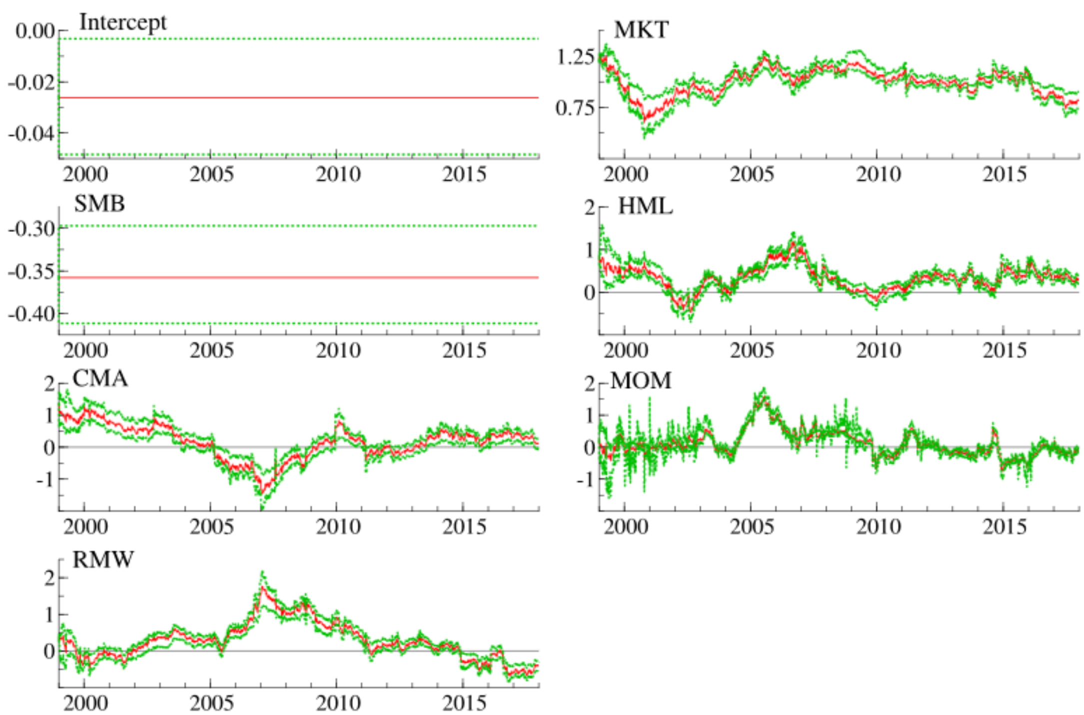
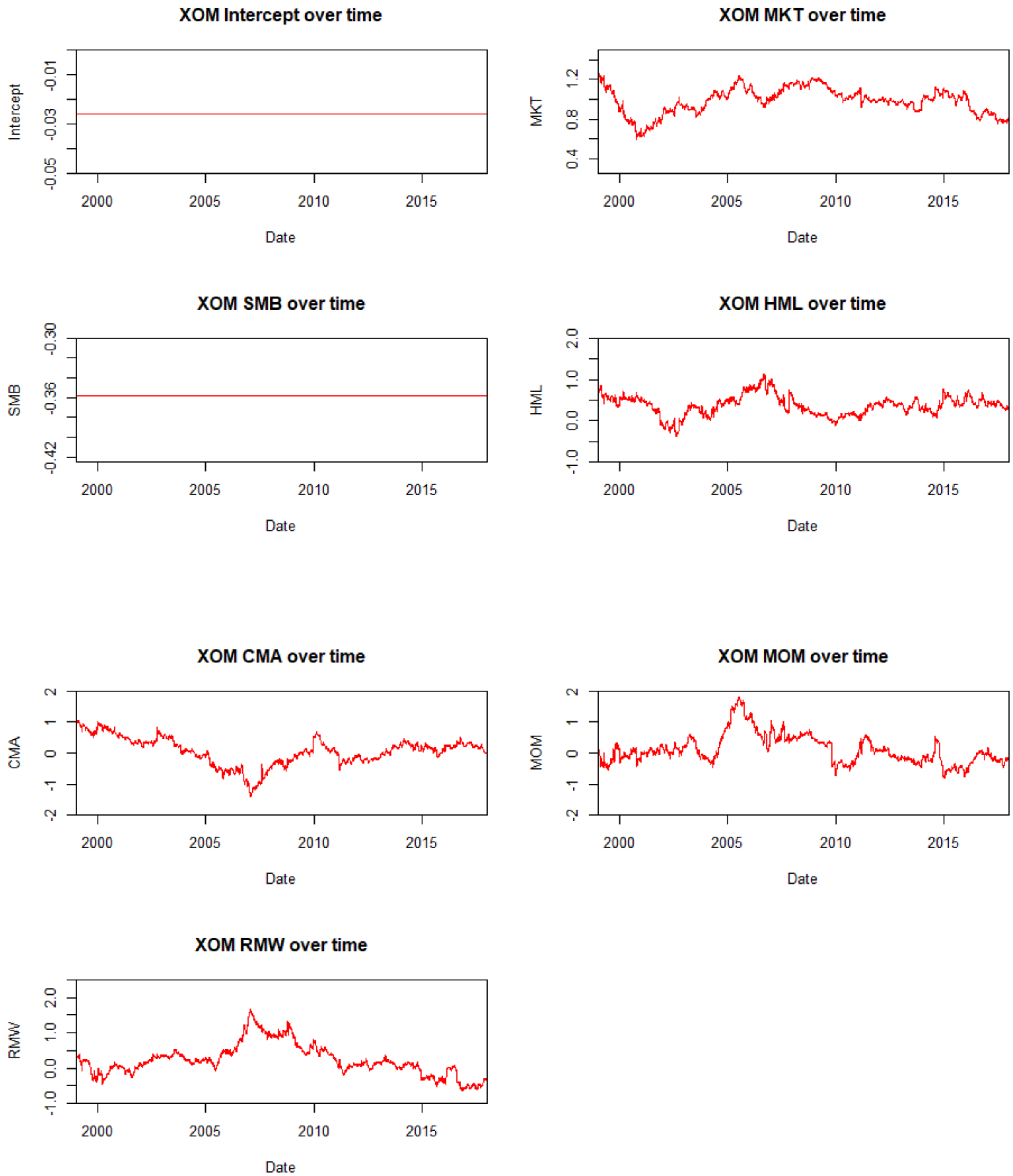
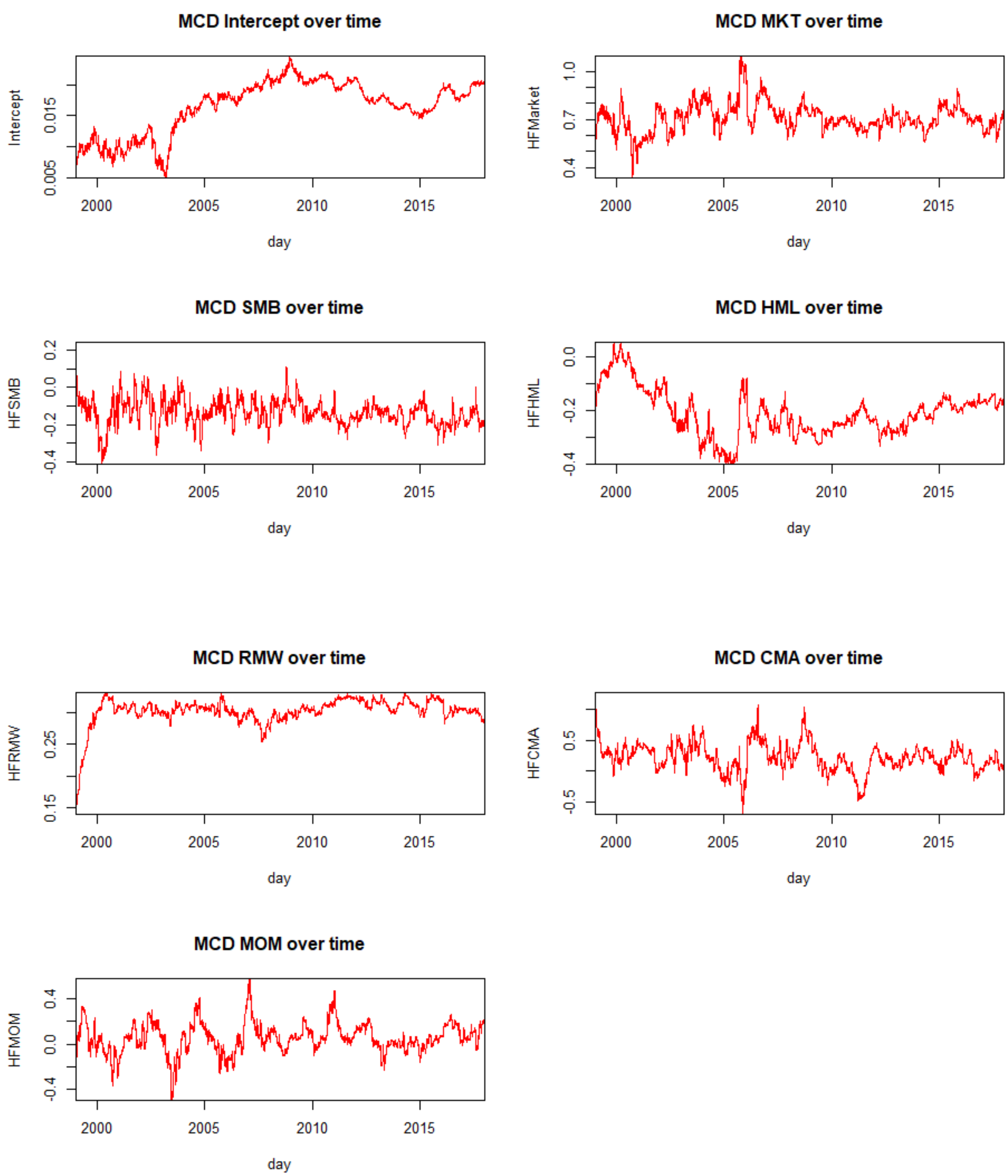


Figure 5: Betas for XOM replicated by our optimizing function



The intercept and SMB betas remain constant just as per Blasques' results, the intercept ϖ_i is used as the constant instead of letting it vary in time.

Figure 6: Betas for MCD replicated by our optimizing function



In this other example for MCD we can see that we have left the intercept and SMB vary in time since we did not have the reference results from Blasques. The rest of the time-varying betas we estimated are present in the appendix section.

As we can see if we look at the graphs, our function succesfully replicates the time-dynamic betas very closely to Blasques, we coroborate this by looking at the log-likelihood values, Blasques obtains a log-likelihood for his estimated ACB on XOM of -6107.4 while we obtain a value of -6140.19 using his provided parameters, the the difference may seem significant but we were limited by the amount of starting values we could try in our process so, with more computing resources we are confident this value would be much closer to that of his paper.

However, when we used our own optimizing function (where we did not let any beta be constant and allowed everything to vary over time unlike in Blasques' article), we find a log-likelihood of -6117.095. The purpose of this is to see that we can successfully replicate Blasques' work so that we can use the same procedure with different stocks for which we do not have previous results, knowing that the methodology is correct for more assets.

However, in addition to this, these log-likelihoods serve as a way to start looking at the performance of models, we can compare the log-likelihood from the ACB model with the ones obtained from the GARCH factor model, which keeps betas constant. An approach that Blasques also briefly mentions in his article.

Table 1: Likelihood Comparison: ACB vs GARCH_factors Models

Stock	**ACB Model**	**GARCH_factors Model**
XOM	**-6117.095**	-6626.135
MSFT	**-7749.92**	-7873.703
BA	**-8214.728**	-8261.78
AXP	**-7588.296**	-7703.226
MCD	**-7290.248**	-7324.708
PFE	**-7219.209**	-7387.017

We see that for all of the stocks, the ACB log-likelihood is larger than the one from the GARCH factors model, according to the Xi^2 distribution, for the difference in-between the two models to be significant at a 95% level of confidence the difference should be equal or higher than 8.67, which all stocks achieve, this is strong evidence in favor of the ACB model when compared to a model that keeps betas constant.

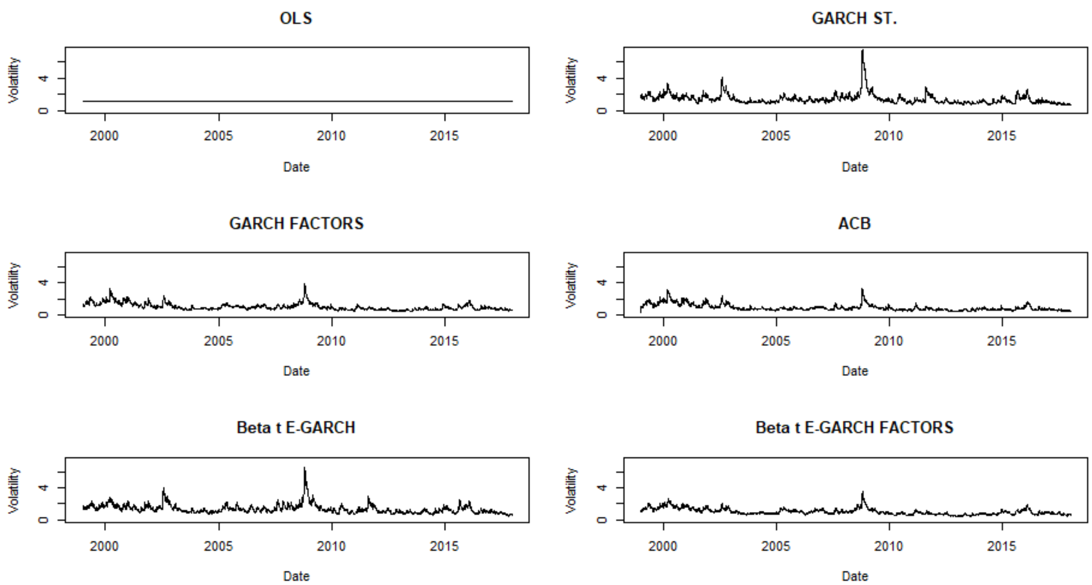
We can take a look at all of our parameters results in table 11 and 12 in the appendix, where we see the results obtained when we let everything vary in time.

4.1.2 Volatility

Figure 7 contains the filtered volatilities for EXXON in all of the six considered models expressed in percentages, we can use this graph to analyze how each model behaves during the same time windows, OLS uses just a constant volatility so there are no episodes of volatility clustering or any variation over time. The ACB model uses the following equation to filter out volatilities.

$$\hat{g}^2_{t+1} = \hat{\omega} \; + \; \hat{\alpha}\, \hat{\nu}^2_t \; + \; \hat{\beta}\, \hat{g}^2_{i,t}$$

Figure 7: Filtered volatilities for XOM



For the non-constant models there seems to be a shared episode of high volatility at the start of the time series around the year 2000 as well as a high peak before the year 2010 which would coincide with the start of the 2008 financial crisis. The GARCH(1,1) standard model seems to reach the highest peak in this latter episode compared to the other ones, whereas the regular Beta t E-GARCH seems to be the highest in the beginning episode, these both have in common the fact that they just use the standard log returns time series data in its estimation, for the other models the values during these sections appear to be more balanced, it is ACB and both the GARCH and Beta t E-GARCH using all factors that show the most steady volatility values and appear to be quite similar overall.

We can verify the points mentioned in the caption further with a statistical summary for the filtered volatilities, in table 2, indeed, the ACB, GARCH factor, and Beta t E-GARCH factor models have lower mean and median volatilities than the rest, with the GARCH(1,1) and Beta t E-GARCH standard models reaching the highest values like we see in the plots.

Table 2: Summary Statistics for Different Model Specifications

Model	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
OLS	1.108	1.108	1.108	1.108	1.108	1.108
GARCH_st	0.612	1.013	1.259	1.395	1.592	7.458
GARCH_factors	0.467	0.745	0.913	1.037	1.222	3.927
ACB	0.4536	0.6946	0.8215	0.9488	1.056	3.2645
Betategarch	0.600	1.126	1.429	1.673	2.016	5.610
Betategarch_factors	0.420	0.761	0.932	1.040	1.237	3.568

To finalize, we have also grouped all filtered volatilities for both XOM individually and all other stocks together, giving a better view of the filtered results. We can see these in figures 8 and 9 in the appendix.

4.2 Model Confidence Set

Evaluating a model's performance is something tricky and many authors have developed their own methodology. Diebold and Mariano (2002) develop a broadly applicable test for the null hypothesis of no difference in the accuracy of two competing forecasts even with forecast errors that follow different distributions. Loss functions can take on many forms, including squared error, absolute error, etc. Later, Giacomini and White (2006) innovated by implementing conditional evaluation of forecasts, meaning which forecast will be more accurate in the next period given all previous information (instead of which model forecasts better on average like the previous literature). Finally, Hansen et al. (2011) establish the framework for the Model Confidence Set (MCS), which allows multiple forecasts to be compared at the same time, thus creating a set of models that have statistically the same predictive ability.

For our thesis, we will use the Model Confidence Set criteria established by (Hansen et al., 2011). They introduce this process as a way to compare candidate models of forecasting without assuming none of them are the "true" best one, since choosing a particular model is an onerous task. So, the key factor is to create a set of models that contains the statistically best ones under a specified confidence level meaning that at that level it can be said that their predictions are all equally good, similar to a confidence interval used for measuring the "true" value of an estimated parameter.

The way the MCS procedure evaluates this predicting ability and this statistic is with the Loss function, this is the loss of accuracy of the model when forecasting. It is obtained by comparing the forecasted value of the model at time t and comparing it to its time-adjacent value on the data we wish to predict, in this case it would compare a volatility value predicted by the ACB model with the realized volatility value at time t. The loss function is what determines the criteria for obtaining this accuracy loss.

There are several ways to obtain said loss, first one and the most common method is the mean squared error, we will be using this and comparing it with the absolute error in our analysis to study whether the changing of the loss function has significant changes in the construction of the superior set of models with equal predictive ability and the ranking system for the models within this SSM.

4.2.1 MCS methodology

The MCS procedure generates a contrast for the equal predictive ability of the superior set of models via the R package 'MCS', the main objective of this contrast is to build a superior set of models (SSM), this is achieved via a statistic of our choice, depending on the statistic we choose the procedure will eliminate the models that fall outside of the bounds marked by the significance degree and then restimate again with the remaining models, this will keep on going until a SSM is created where the null hypothesis is that all of the remaining models have the same predictive ability and the alternative is this not being true. After the algorithm starts to work we obtain a series of columns as a result, these are:

- Rank_M, represents the rank of the model, a rank closer to 1 means the model has a lower statistic value, thus being better than the rest.
- v_M, represents the value of the statistic from T_{max} related to the Rank, it grows lower with the rank, thus 1 is the lowest value with it often being negative and higher ranks, meaning "worse" models, make the value of the statistic grow.
- MCS_M, is a measure of probability related to a model being in the confidence set under a certain percentage of confidence, lower values will mean the model does not belong in the set.
- _R, these variations come from the TR statistic, both Rank_R, v_R, and MCS_R mean the same but applied to the other statistic, depending on the one we select the code will check either this or the _M columns for the elimination process.
- Tmax/TR, relates to the statistic used to estimate the Equal Predictive Ability of the models, and correspond with $T_{max,M} = \max_{i \in M} t_i$ and $T_{R,M} = \max_{i,j \in M} |t_{ij}|$ respectively, where those t values correspond with these expressions from (Bernardi and Catania, 2018):

$$t_{ij} = \frac{\bar{d}_{ij}}{\sqrt{v\hat{a}r\left(\bar{d}_{ij}\right)}} \quad \text{and} \quad t_i = \frac{\bar{d}_{i,\cdot}}{\sqrt{v\hat{a}r\left(\bar{d}_{i,\cdot}\right)}} \quad \text{for} \quad i, j \in \mathbb{M}$$

where

$$d_{ij,t} = \ell_{i,t} - \ell_{j,t}, \quad i, j = 1, \ldots, m, \quad t = 1, \ldots, n,$$

$$d_{i,t} = (m-1)^{-1} \sum_{j \in \mathbb{M}} d_{ij,t}, \quad i = 1, \ldots, m,$$

- For TR, the null hypothesis is according to (Bernardi and Catania, 2018):

$$\begin{aligned} \mathrm{H}_{0,\mathrm{M}} &: c_{ij} = 0, \quad \text{for all } i, j = 1, 2, \ldots, m \\ \mathrm{H}_{\mathrm{A},\mathrm{M}} &: c_{ij} \neq 0, \quad \text{for some } i, j = 1, \ldots, m, \end{aligned}$$

  Where c_{ij} = $\mathbb{E}[d_{ij}]$ we can see that it looks at all possible pairs of models for the contrast. Meanwhile T_{max}

$$\begin{aligned} \mathrm{H}_{0,\mathrm{m}} &: c_i = 0, \quad \text{for all } i = 1, 2, \ldots, m \\ \mathrm{H}_{\mathrm{A},\mathrm{m}} &: c_i \neq 0, \quad \text{for some } i = 1, \ldots, m, \end{aligned}$$

  Where c_i = $\mathbb{E}[d_i]$ so it checks for the mean of different models.

If we consider first a situation where the data contains little information in such a way that the test lacks power and the elimination rule may cause a superior model to be chosen before the elimination of all inferior models, this lack of power causes the procedure to finish too early on average and thus the MCS will contain a large number of model including several inferior ones.

We view this as a strength of the MCS procedure, since lack of power is tied to a lack of information in the data, the Model Confidence Set should be large when there is insufficient information to distinguish good and bad models. In the opposite situation, where data is informative, the test is powerful and will reject all false hypotheses and on top of that the elimination rule will not stop until all inferior models have been eliminated.

The T_{max} and TR statistics offer different ways to build the superior set of models, like we can see in the equations above, both statistics have a different structure, the TR is formed by a maximum of the absolute value of t_{ij} with it being a standard shape we see often in other frameworks like (Diebold and Mariano, 2002) of a mean divided by the square root of the variance, with d_{ij} being the difference between the loss of two values of data, estimated and the true value we want to predict at time t.

The TR statistic performs a two-tail comparison on all pairs of models possible, which is why it uses the suffixes i and j, so in order to reject the hypothesis we need the differences to be large, regardless of which one we are in favor of, to lean towards the alternative that there is indeed a difference.

T_{max} on the other hand changes this structure for a similar looking standardization form, but without the absolute value, here instead of just taking the value of d as just a difference of two values, it also calculates the mean of these differences. So when we are contrasting that the mean is 0 we may find two situations, either all differences are 0, or one model is clearly better than others but also worse than others, in such a way that by offsetting each other the mean becomes 0.

T_{max} thus keeps more models because it does not eliminate models that are clearly below the best one but that are not as clearly below the rest of the models which are also below the best one. This is consistent with our results underneath, which shows that T_{max} keeps more models in the Superior Set than TR, and that TR eliminates more models much quicker, since it simply compares each model with the best one instead of taking into account how many models remain close to the best.

4.2.2 MCS results

We are going to start by displaying the MCS results in the following tables. The information displayed contains the MCS results for six stocks each containing the models that remain in the superior set of models along with the ones that were eliminated, the statistic values for each model in each of the stocks and finally the p-value that the superior set of models itself obtained:

Table 3: Model Confidence Set (MCS) Results using Tmax Statistic and Squared Error

Stock	Model	Rank_M	v_M	MCS_M	Loss	Status
XOM	ACB	1	-6.4107	1.0000	0.6799	Superior Set
	OLS	2	0.8990	0.5396	0.8415	Superior Set
	BETATEGARCH-Factors	3	1.5259	0.2088	0.8205	Superior Set
	Factors	4	1.7888	0.1312	0.8442	Superior Set
	STD	-	-	-	-	*Eliminated*
	BETATEGARCH	-	-	-	-	*Eliminated*
MSFT	ACB	1	-2.5310	1.0000	1.2985	Superior Set
	BETATEGARCH-Factors	2	-0.0883	1.0000	1.3636	Superior Set
	Factors	3	0.2119	0.9686	1.3748	Superior Set
	OLS	4	0.8311	0.5844	1.4352	Superior Set
AXP	BETATEGARCH-Factors	1	-4.1149	1.0000	1.2699	Superior Set
	ACB	2	-0.6387	1.0000	1.4071	Superior Set
	OLS	3	0.6907	0.6928	1.5061	Superior Set
	Factors	4	1.6789	0.1652	1.5346	Superior Set
BA	ACB	1	-4.3030	1.0000	1.5307	Superior Set
	Factors	2	-1.0841	1.0000	1.5816	Superior Set
	OLS	3	0.4735	0.8208	1.6473	Superior Set
	BETATEGARCH-Factors	4	1.5130	0.2038	1.6809	Superior Set
MCD	ACB	1	-3.3123	1.0000	1.1571	Superior Set
	Factors	2	-1.9240	1.0000	1.1811	Superior Set
	OLS	3	0.2291	0.9516	1.2431	Superior Set
	BETATEGARCH-Factors	4	1.9143	0.0928	1.3193	Superior Set
PFE	BETATEGARCH-Factors	1	-0.7171	1.0000	1.1639	Superior Set
	OLS	2	0.0406	0.9976	1.1925	Superior Set
	ACB	3	0.5621	0.7328	1.2133	Superior Set

All the models will appear in the XOM row, but we will remove the eliminated ones in the following rows.

What we find is that all the superior sets of models we have built contain the same models for all stocks, these being OLS, GARCH(1,1) with Factors, ACB and Beta t E-GARCH using factors except PFE which also eliminates the GARCH(1,1) using the factors; with the standard GARCH and Beta t E-GARCH being eliminated in all instances. Some aspects we need to keep in mind are, in this table we use the T_{max} statistic, which like we mentioned tends to favor more models even if some are inferior, this means that for this case it seems the defining factor when modeling realized volatility is the use of the factor regressors in the estimation, both of the models that get eliminated in all cases are the ones that were built with only the stock data, which highlights the importance of said factors when building models.

The ACB model is ranked highly in each superior set, always being ranked first save for AXP and PFE where Beta t E-GARCH with factors is favored.

It is specially interesting that this importance of the factors applies even to the OLS model that only contains a singular value constant over time, but since it does not get eliminated from the set it is not redundant to take it into account.

Table 4: MCS Process Summary Statistics Squared Error - Tmax

Stock	Models Eliminated	Statistic	P-value
XOM	2	T_max	0.1312
MSFT	2	T_max	0.5844
AXP	2	T_max	0.0676
BA	2	T_max	0.2038
MCD	2	T_max	0.0928
PFE	3	T_max	0.7328

By looking at the p-values for each of the superior set of models we observe that, indeed, we cannot reject the Null hypothesis that all of the models in the set have statistically the same predictive ability. This means, in relation to predicting realized volatility, on average it is statistically the same to predict it using an OLS constant model or a more sophisticated ACB or Beta t E-GARCH model as long as all the factor regressors are being used in the estimation.

Table 5: Model Confidence Set (MCS) Results using TR Statistic and Squared Error

Stock	Model	Rank_R	v_R	MCS_R	Loss	Status
XOM	ACB	1	-2.6319	1.0000	0.6799	Superior Set
	OLS	-	-	-	-	*Eliminated*
	Factors	-	-	-	-	*Eliminated*
	BETATEGARCH-Factors	-	-	-	-	*Eliminated*
	STD	-	-	-	-	*Eliminated*
	BETATEGARCH	-	-	-	-	*Eliminated*
MSFT	OLS	3	1.2968	0.3726	1.4352	Superior Set
	ACB	1	-1.1287	1.0000	1.2985	Superior Set
	BETATEGARCH-Factors	2	1.1287	0.4662	1.3636	Superior Set
AXP	OLS	2	1.6672	0.0950	1.5061	Superior Set
	BETATEGARCH-Factors	1	-1.6672	1.0000	1.2699	Superior Set
BA	OLS	2	1.2405	0.2144	1.6473	Superior Set
	ACB	1	-1.2405	1.0000	1.5307	Superior Set
MCD	OLS	2	0.9034	0.3598	1.2431	Superior Set
	ACB	1	-0.9034	1.0000	1.1571	Superior Set
PFE	OLS	2	0.2966	0.9518	1.1925	Superior Set
	ACB	3	1.0979	0.4984	1.2133	Superior Set
	BETATEGARCH-Factors	1	-0.2966	1.0000	1.1639	Superior Set

All the models will appear in the XOM row, but we will remove the eliminated ones in the following rows.

The next set of tables, 5 and 6 looks at the same loss stats obtained with squared error but with the TR statistic instead, like we described earlier, this method is calculated differently so we expect to see widely different results in the formation of the superior set of models. The difference in eliminated models is much more varied in this case, instead of the regular 2 eliminations we were seeing earlier, for XOM all models besides the ACB are eliminated, for the rest of the stocks the number of removed models differs heavily.

We see here the ACB model being eliminated for the first time for AXP, for the other stocks, ACB keeps a good ranking position in general. The presence of the OLS constant model also stands out, and in several cases it is estimated to be just as good as the ACB model.

Table 6: MCS Summary Statistics Squared Error - TR

Stock	Models Eliminated	Statistic	P-value
XOM	5	TR	0.0076
MSFT	3	TR	0.3726
AXP	4	TR	0.095
BA	4	TR	0.2144
MCD	4	TR	0.3598
PFE	3	TR	0.4984

The p-values remain consistent with our previous results for the sets of models that contain more than one, we do not reject the null hypothesis so these are statistically equal in predictive ability, however since we now have a stock where only 1 model remains this p-value is considerably lower, which is to be expected since only having one model makes the set pretty much redundant.

Table 7: Model Confidence Set (MCS) Results using Tmax Statistic and Absolute Error

Stock	Model	Rank_M	v_M	MCS_M	Loss	Status
XOM	ACB	1	-12.1714	1.0000	0.7524	Superior Set
	OLS	-	-	-	-	*Eliminated*
	Factors	-	-	-	-	*Eliminated*
	BETATEGARCH-Factors	-	-	-	-	*Eliminated*
	STD	-	-	-	-	*Eliminated*
	BETATEGARCH	-	-	-	-	*Eliminated*
MSFT	ACB	1	-0.1097	1.0000	1.0710	Superior Set
	BETATEGARCH-Factors	2	0.1097	0.9072	1.0721	Superior Set
AXP	BETATEGARCH-Factors	1	-8.9878	1.0000	1.0032	Superior Set
BA	ACB	1	-8.4480	1.0000	1.1779	Superior Set
MCD	ACB	1	-5.3445	1.0000	0.9953	Superior Set
PFE	OLS	4	2.1814	0.0508	1.0871	Superior Set
	Factors	3	0.1057	0.9954	1.0233	Superior Set
	ACB	1	-3.4260	1.0000	0.9856	Superior Set
	BETATEGARCH-Factors	2	-2.5901	1.0000	0.9917	Superior Set

All the models will appear in the XOM row, but we will remove the eliminated ones in the following rows.

For tables 7 and 8 we now use the absolute error to calculate the loss, we are back to using the T_{max} statistic, the results this time are outstandingly different, more models seem to get eliminated now, except for PFE; the absolute error as the loss function seems to leave ACB as the sole best predictor for realized volatility except for AXP which from what we have seen until now seems to favor the Beta t E-GARCH model with factors, this is consistent with our previous results.

Table 8: MCS Summary Statistics Absolute Error - Tmax

Stock	Models Eliminated	Statistic	P-value
XOM	5	T_max	0.00
MSFT	4	T_max	0.9072
AXP	5	T_max	0.00
BA	5	T_max	0.00
MCD	5	T_max	0.00
PFE	2	T_max	0.0508

Seeing this, the p-values obtained are mostly redundant since most sets only have 1 model, for the ones that do have more than one we still do not reject the null hypothesis of equal predictive ability. Using the absolute error instead of the squared error seems to make the loss difference between ACB and the other models pretty significant, where in most sets this is the sole remaining model.

Table 9: Model Confidence Set (MCS) Results using TR Statistic and Absolute Error

Stock	Model	Rank_R	v_R	MCS_R	Loss	Status
XOM	ACB	1	-7.6576	1.0000	0.7524	Superior Set
	OLS	-	-	-	-	*Eliminated*
	Factors	-	-	-	-	*Eliminated*
	BETATEGARCH-Factors	-	-	-	-	*Eliminated*
	STD	-	-	-	-	*Eliminated*
	BETATEGARCH	-	-	-	-	*Eliminated*
MSFT	ACB	1	-0.1066	1.0000	1.0710	Superior Set
	BETATEGARCH-Factors	2	0.1066	0.9150	1.0721	Superior Set
AXP	BETATEGARCH-Factors	1	-9.0534	1.0000	1.0032	Superior Set
BA	ACB	1	-8.7389	1.0000	1.1779	Superior Set
MCD	ACB	1	-8.6835	1.0000	0.9953	Superior Set
PFE	ACB	1	-1.2430	1.0000	0.9856	Superior Set
	BETATEGARCH-Factors	2	1.2430	0.2120	0.9917	Superior Set

All the models will appear in the XOM row, but we will remove the eliminated ones in the following rows.

Finally, tables 9 and 10 have the same loss function with the absolute error but use the TR statistic again, the trend is similar to what we saw in the previous tables, ACB or Beta t E-GARCH with factors being the most favored models with ACB often being ranked first, the superior sets are formed by the same models as table 7 except for the stock PFE which loses two models in this procedure.

Table 10: MCS Summary Statistics Absolute Error - TR

Stock	Models Eliminated	Statistic	P-value
XOM	5	TR	0.00
MSFT	4	TR	0.915
AXP	5	TR	0.00
BA	5	TR	0.00
MCD	5	TR	0.00
PFE	4	TR	0.212

These results remain consistent, for the 2 sets with more than 1 model, ACB and Beta t E-GARCH with factors have, statistically, the same predictive ability.

What we gather in general after these results is that ACB and Beta t E-GARCH with factors are favored heavily in the MCS process, for basically all stocks ACB is always present in the superior set of models and ranked first, this highlights the importance of score-driven models when predicting parameters and makes ACB potentially the best model for predicting realized volatility. It is also worth mentioning the fact that OLS, even as just a constant over time, also appears to form part of the superior set in some ocations, meaning a simpler model is sometimes also a correct choice.

4.3 Future lines of work

We were limited by time constraints and computing resources, so this work has some limitations, and in the future we would like to study, aside from more models estimated for more stocks, risk management and hedging performance based on the conditional volatility of the models, VaR backtesting, portfolio building, etc. We are satisfied with our work and results for the time being and we prove that the ACB model is a very complete framework whose use we hope to spread with these results, as the model's practical benefits were also explored by Blasques himself (Blasques et al., 2024).

5 Conclusion

In this thesis, we talk about score-driven models, and a recent one like Blasques' ACB model, along with others like Beta t E-GARCH. We describe the realized volatility structure and explore how these models perform in terms of conditional volatility (something that was missing in the original literature) by forecasting realized volatility values; we then compare these using the model confidence set framework.

Score-driven models are a big step in modeling parameters that are not constant over time, since what we see in the real world is that parameters are constantly changing. What we found in our results is that letting parameters vary over time combined with the use of good-quality relevant data, such as the Fama-French factors, is very useful in making a good model.

We also find empirical evidence of the ACB framework being a very strong model based on the log-likelihood compared to that of a model with constant betas, proving that time-varying parameters offer better modeling potential. Furthermore, we find strong evidence in favor of score-driven models, with the model confidence set procedure often having ACB and Beta t E-GARCH with factors as equal in terms of predictive ability. As a final addition, seeing the OLS constant in some superior sets gives reason to using simpler models in certain situations if the circumstances demand for it.

Future research could expand these results with more models across more stocks, while studying more practical applications based on conditional volatility for risk management, hedging, etc.

6 Bibliography

References

Andersen, T. G. and Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International economic review*, pages 885–905.

Andersen, T. G. and Teräsvirta, T. (2009). *Realized Volatility*, pages 555–575. Springer Berlin Heidelberg.

Bernardi, M. and Catania, L. (2018). The model confidence set package for r. *International Journal of Computational Economics and Econometrics*, 8.

Blasques, F., Francq, C., and Laurent, S. (2024). Autoregressive conditional betas. *Journal of Econometrics*, 238.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327.

Box, G. E., Jenkins, G. M., and Reinsel, G. C. (2013). *Time series analysis: Forecasting and control: Fourth edition*.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of finance*, 52(1):57–82.

Cipra, T. (2020). *State Space Models of Time Series*, pages 373–390. Springer International Publishing.

Creal, D., Koopman, S. J., and Lucas, A. (2011). A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. *Journal of Business & Economic Statistics*, 29:552–563.

Creal, D., Koopman, S. J., and Lucas, A. (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28.

Da, R. and Xiu, D. (2021). When moving-average models meet high-frequency data: Uniform inference on volatility. *Econometrica*, 89(6):2787–2825.

Diebold, F. X. and Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 20:134–144.

Durbin, J. and Koopman, S. J. (2012). *Time Series Analysis by State Space Methods*. Oxford University Press.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50:987.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33:3–56.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of financial economics*, 116(1):1–22.

Fama, E. F. and MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81:607–636.

Francq, C. and Zakoian, J.-M. (2019). *GARCH models: structure, statistical inference and financial applications*. John Wiley & Sons.

Francq, C. and Zakoïan, J.-M. (2012). Qml estimation of a class of multivariate asymmetric garch models. *Econometric Theory*, 28:179–206.

Francq, C. and Zakoïan, J.-M. (2015). Risk-parameter estimation in volatility models. *Journal of Econometrics*, 184:158–173.

Giacomini, R. and White, H. (2006). Tests of conditional predictive ability. *Econometrica*, 74.

Hansen, P. R., Lunde, A., and Nason, J. M. (2011). The model confidence set. *Econometrica*, 79:453–497.

Harvey, A. and Chakravarty, T. (2008). Beta-t-(e)garch. *Cambridge Working Papers in Economics*, CWPE 0840.

Harvey, A. and Sucarrat, G. (2014). Egarch models with fat tails, skewness and leverage. *Computational Statistics and Data Analysis*, 76.

Harvey, A. C. (2002). Forecasting, structural time series models and the kalman filter. *Wilmott*, 2002:14–15.

Koopman, S. (2012). *State-Space Methods*. Wiley.

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

Sucarrat, G. (2013). Betategarch: Simulation, estimation and forecasting of beta-skew-t-egarch models. *R Journal*, 5.

Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, 100(472):1394–1411.

7 Appendix

Table 11: Parameter Estimates for Six Stocks

Parameter	XOM	MSFT	AXP	BA	MCD	PFE
ϖ_1	-0.000107	0.000032	0.000140	0.000038	0.000041	-0.000517
ϖ_2	0.011376	0.017049	0.011912	0.019267	0.009718	0.005084
ϖ_3	-0.014418	-0.007159	-0.010549	-0.002981	-0.003485	-0.002348
ϖ_4	0.003878	-0.001377	0.004142	0.000071	0.000168	0.000214
ϖ_5	0.000033	0.000262	0.000605	0.001406	0.002973	0.008886
ϖ_6	0.001203	-0.002761	0.003252	0.000747	0.001472	0.001787
ϖ_7	-0.000024	-0.000561	-0.000558	-0.000035	0.000771	0.000474
ξ_1	-0.006963	0.001232	-0.007924	0.001568	0.000156	-0.015120
ξ_2	0.022620	0.016760	0.013041	0.011043	0.011242	0.001757
ξ_3	0.002969	0.007745	-0.011008	0.005888	0.007486	0.000774
ξ_4	0.033501	0.003639	0.003282	0.005716	-0.002496	0.005356
ξ_5	0.006768	0.015440	-0.008857	-0.002721	0.000602	0.026354
ξ_6	0.017521	0.012840	-0.008288	-0.004907	0.008844	0.014863
ξ_7	0.016495	0.009239	0.021017	0.010842	0.010407	0.004421
c_1	0.996012	1.001895	0.985138	0.999974	0.997724	0.968779
c_2	0.988366	0.983906	0.990308	0.982008	0.986392	0.994450
c_3	0.957145	0.974450	0.926177	0.976263	0.975938	0.992525
c_4	0.984773	0.997679	0.989739	0.999867	1.000075	0.999795
c_5	0.999231	0.996925	0.996467	0.994952	0.990241	0.949662
c_6	0.992522	0.997569	0.968941	0.999586	0.990905	0.995489
c_7	0.997005	0.996360	0.994954	0.988569	0.991573	0.996424
$\beta_{i1,1}$	0.006202	-0.070565	0.319697	-0.077377	0.006473	-0.265634
$\beta_{i1,2}$	0.978278	-0.169524	2.963808	0.389640	0.434627	0.487712
$\beta_{i1,3}$	-0.370277	-2.294882	2.618651	0.246316	0.244469	-1.507927
$\beta_{i1,4}$	0.456871	-0.117999	0.151144	-0.050980	-0.195328	-0.799677
$\beta_{i1,5}$	-0.133053	-0.110459	0.046418	-0.418503	0.139686	1.734629
$\beta_{i1,6}$	0.785243	-2.811756	2.712229	0.501350	1.275382	-0.091866
$\beta_{i1,7}$	-0.251906	0.223239	0.003453	-0.058431	-0.115093	0.455388
g_1	1.312000	1.312000	1.312000	1.312000	1.312000	1.312000
ω	0.008584	0.009380	0.033440	0.020861	0.003130	0.030968
α	0.063796	0.019415	0.110218	0.036017	0.017592	0.124314
β	0.930123	0.975302	0.880945	0.953911	0.979964	0.868604

Table 12: Shared μ Parameters from Factor GARCH(1,1) models (Same For All Stocks)

Parameter	Value
μ_{INT}	1.000000
μ_{MKT}	0.059815
μ_{SMB}	0.017875
μ_{HML}	0.002508
μ_{RMW}	0.001169
μ_{CMA}	-0.000461
μ_{MOM}	0.017010

Figure 8: Filtered volatilities for XOM combined

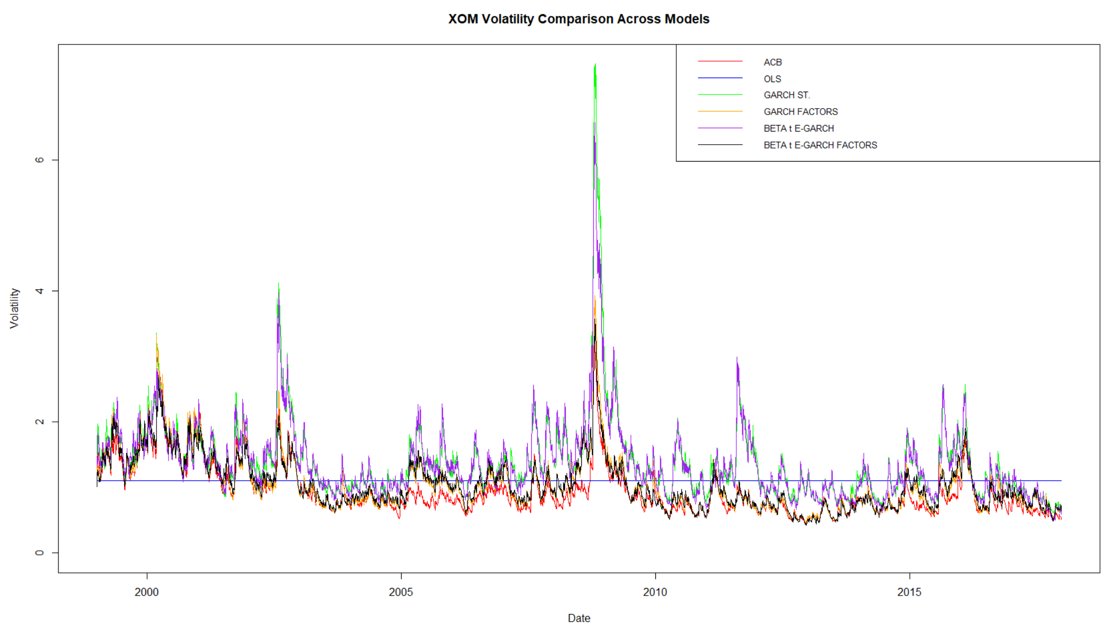
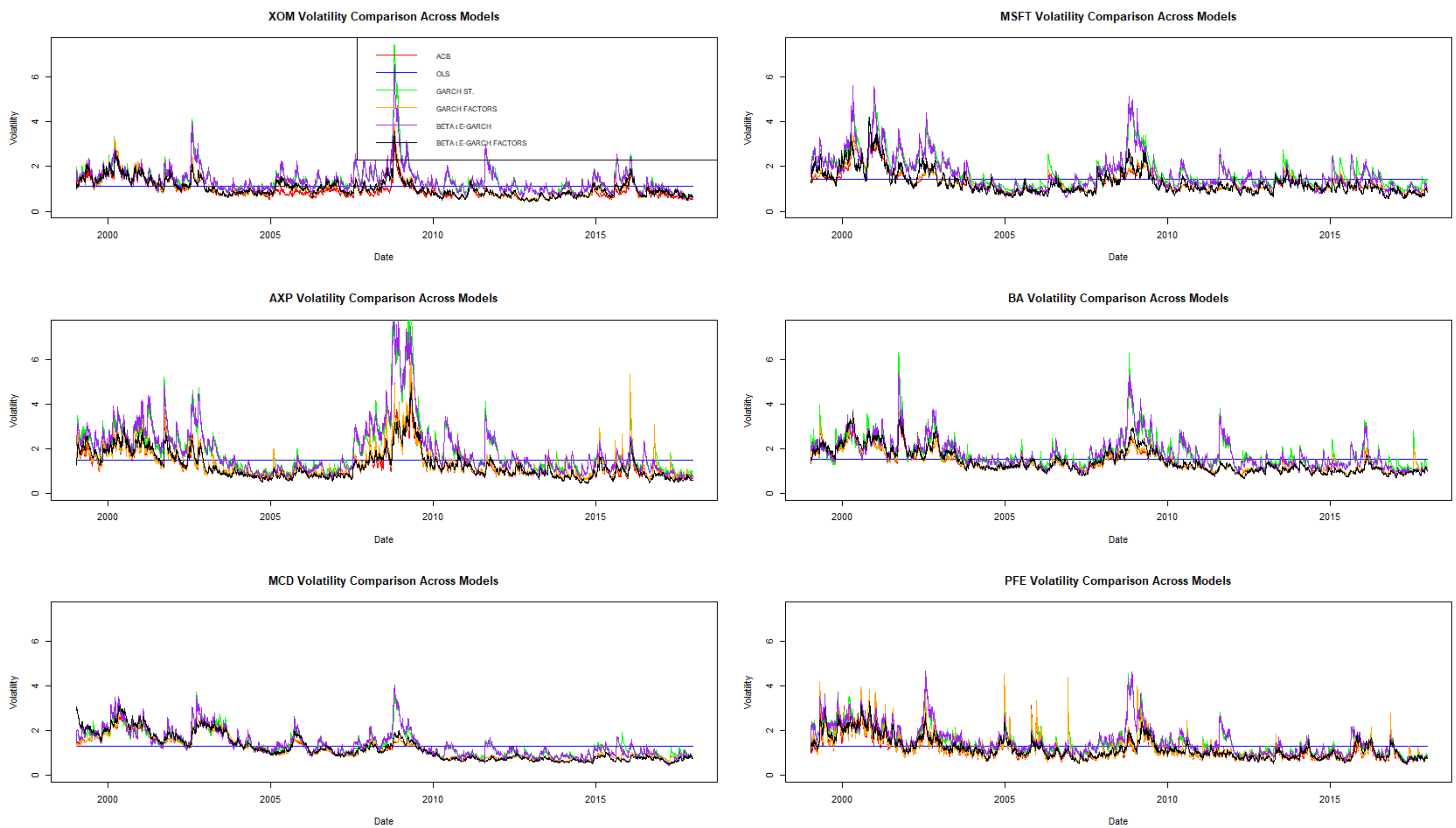


Figure 9: All filtered volatilities



AXP seems to be the most volatile stock according to our filtered data all in the same scale, while MCD appears to be the least. In general the trends among models appear to be consistent in all stocks, with slight variations like for example the GARCH(1,1) with factors being more noticeable for PFE, etc.

Figure 10: Betas for MSFT replicated by our optimizing function

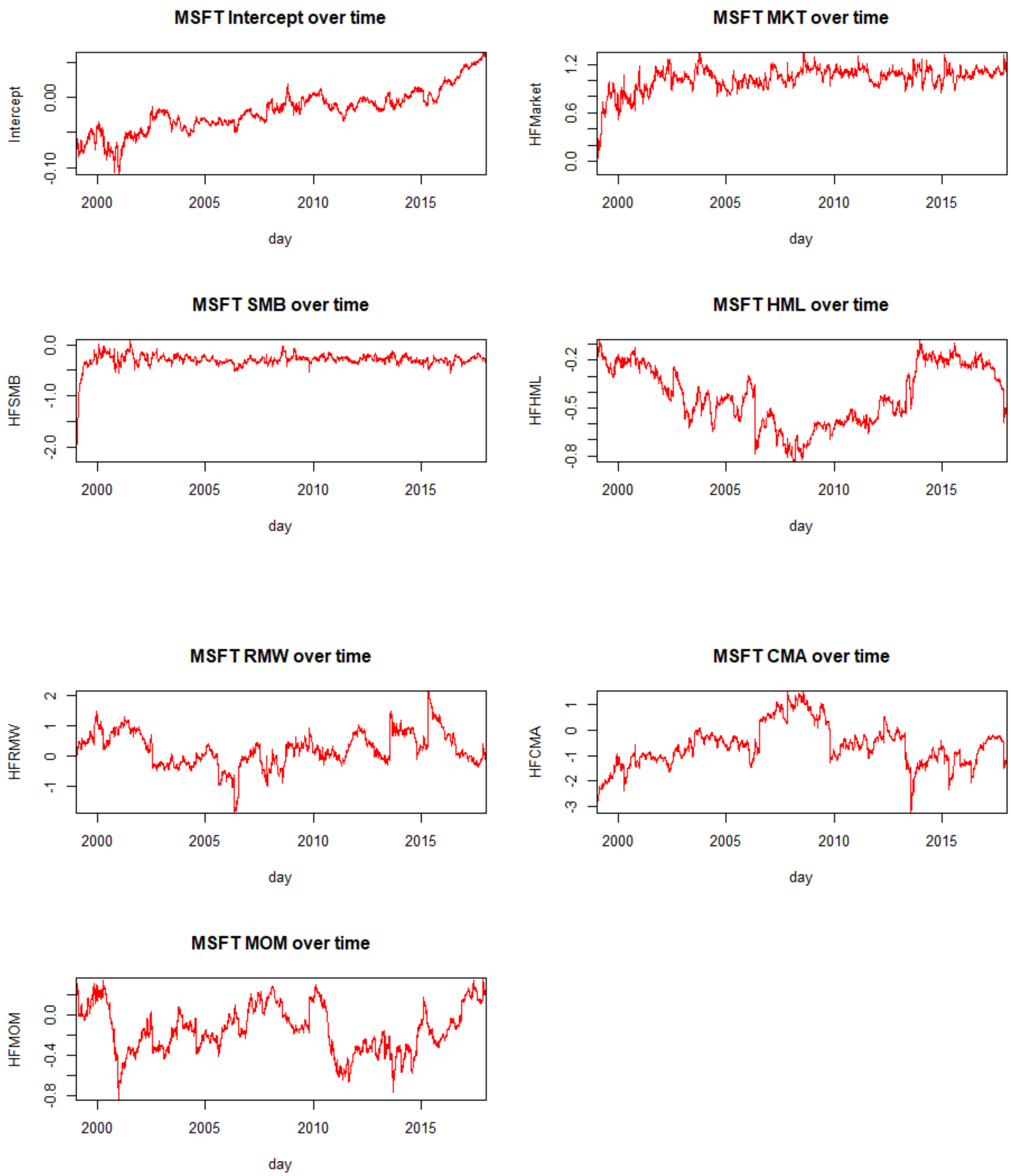


Figure 11: Betas for BA replicated by our optimizing function

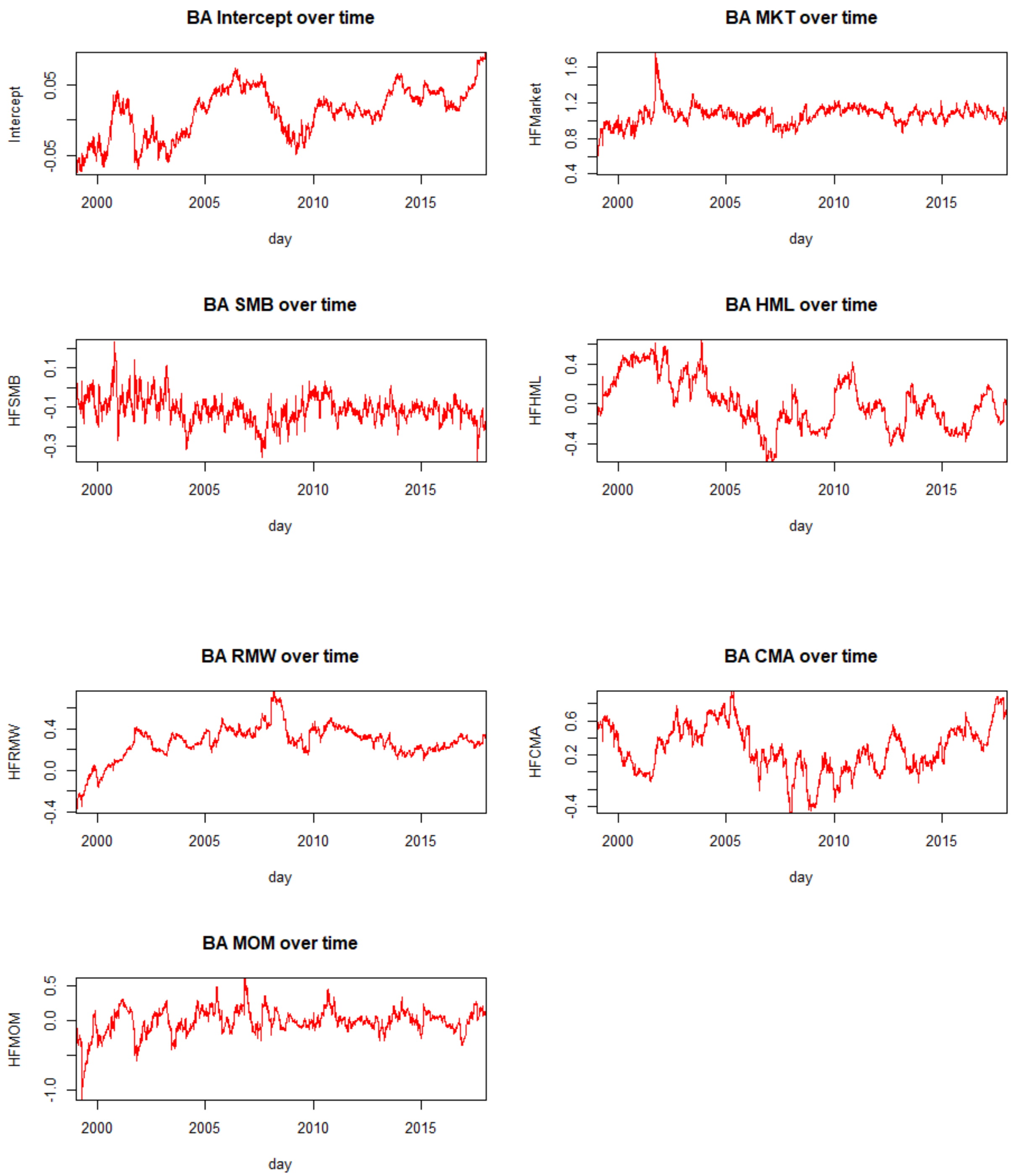


Figure 12: Betas for AXP replicated by our optimizing function

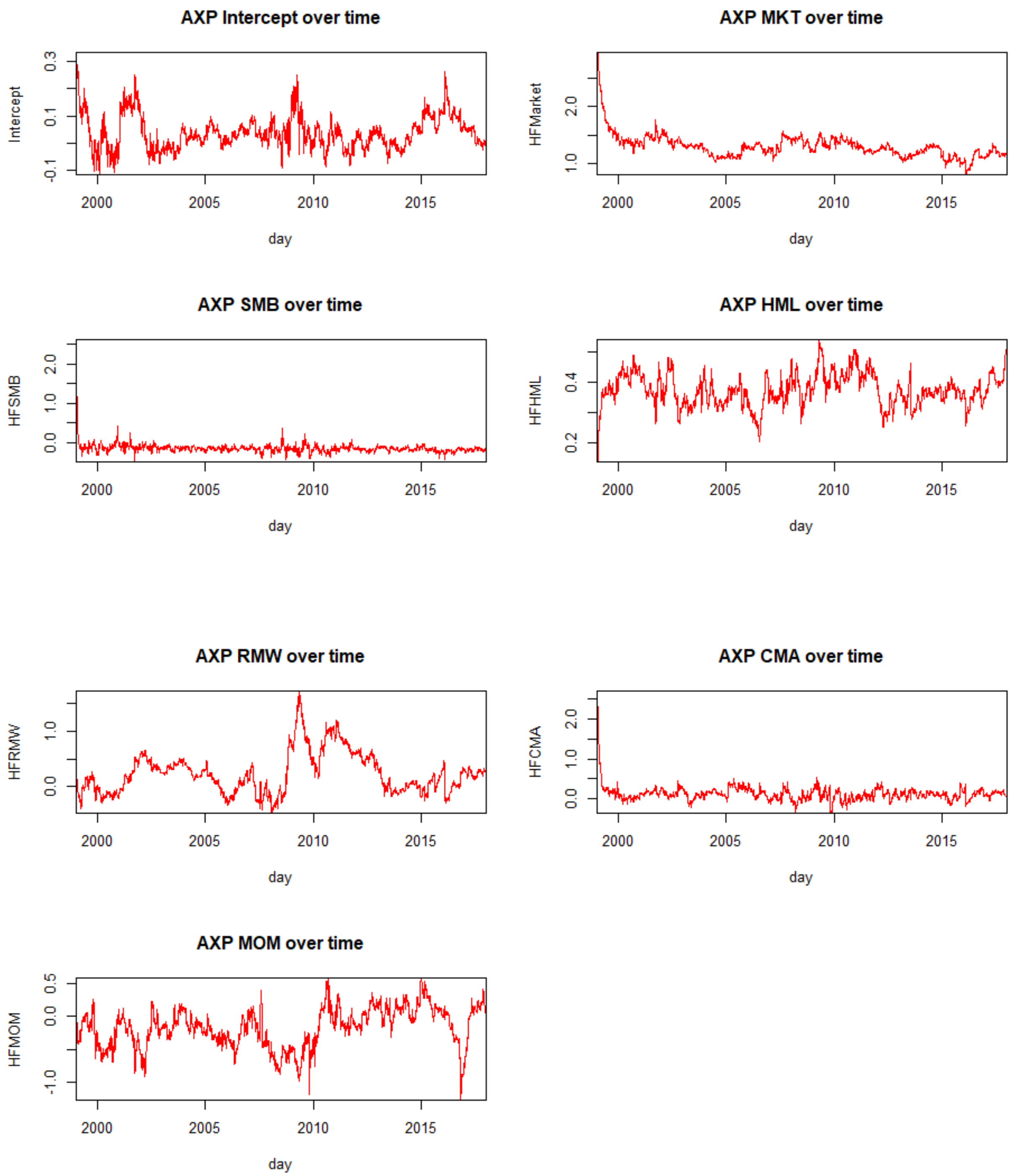


Figure 13: Betas for PFE replicated by our optimizing function

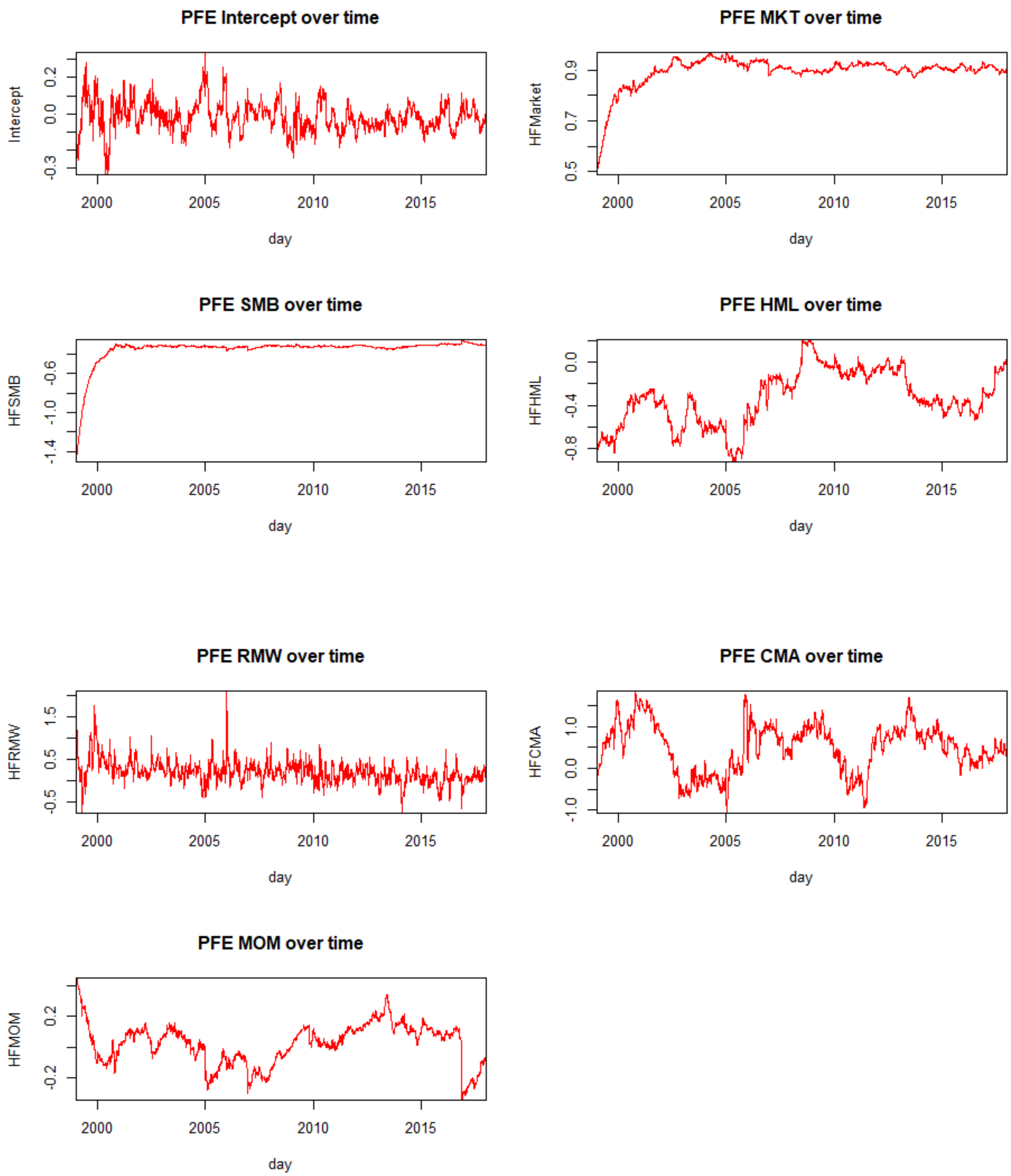


Figure 14: Betas for XOM replicated by our optimizing function