The Next Generation of Multi-Factor Models: Time Heterogeneity

Shiyu Du

Xinhui Gu

Qingyao (Patrick) Sun

David Thai

A thesis submitted in partial fulfillment for the

degree of Master in Financial Engineering

in the

Haas School of Business

University of California Berkeley

Supervisor: Prof. Martin Lettau

March 2018

Contents

Abstract

Although factor models are currently well-understood and widely used, the majority of the implementations are in the domain where coefficients (exposure to factors) remain constant over time, which is a strong assumption when using data that span long time horizons. In this paper, we explore the relatively less well-traversed area of time-varying coefficients. We start by developing the traditional Fama-French factor models, with simple rolling regressions to allow for time-dependency of the coefficients. Then we move onto models where coefficients vary with time, using one separate state equations for each factor in the Kalman Filter. To our knowledge we are the first to compare the effectiveness of the Kalman Filter technique on 5 factor models with other time-varying coefficient models. For this comparison, we have designed a series of tests. First we check the consistency across models, to analyze whether our models' outputs are drastically different. Then we move onto asset pricing aspect where we analyze the R-squared and mean squared error over time, and use cross sectional data to estimate the implied factor risk premium, keeping in mind that correct coefficient estimates should lead to positive risk premium on average. Finally we will test the models in an asset management perspective, where we construct portfolios based on minimum variance and mean-variance optimization and compare the outcome of these portfolios over time.

Introduction and Data

1 Introduction

1.1 Background

Factor models have been at the forefront of financial empirical research for the past 50 years. Beginning with their initial form as the Capital Asset Pricing Model(CAPM) in the 1960s, modern factor models have hundreds of factors and incorporate a vast variety of approaches, such as using multiple stage regressions to filter out one set of factors at a time. These models have gained such popularity because they are easy to compute and often have clear interpretations that align with our intuition. Many use factor models not only for return prediction, but also as a way to quantitatively analyze different types of risks that a particular portfolio is exposed to, and to gain insights to its risk return characteristics. Nowadays most of the research in the vicinity of factor models focus on the discovery of innovative alpha signals after adjusting for well known factors, and on the customized constructions of factors themselves instead of the loadings on the factors. The alpha signals have significant impact because they are directly related to profitability, after neutralizing exposures to other factors. Hence active managers are constantly seeking new signals in the market that are not yet well exploited, which can then serve as investment opportunities. At the same time there also exist many ways of constructing the same factor. These methods capture subtlety different risks that are specific to each investor, hence constructing tailored factors is also of interest to professionals in the industry. However less is done in terms of time-varying coefficients/exposures to these factors and the predictive power of the factor model measured as prediction accuracy for example. We believe the time-varying exposures are exactly what makes factor models interesting. This is because most active managers have some kind of re-balancing strategy which can be rule based for a quantitative fund or judgment based, but in essence this is because they believe their portfolio's exposures to the risk factors are changing over time. Hence in order to consistently hedge away the risks that are not part of their active bet, they need to periodically re-balance their portfolio. The issue is that without a model for the constantly changing risk exposure, the re-balancing is only an approximation to what they are really looking to achieve. Our paper aims to improve the traditional factor models by recognizing the fact that not only the factors, but also the exposures to them vary over time.

1.2 Literature Review

The capital asset pricing model (CAPM) originally developed by Willam Sharpe (1964) and John Lintner (1965) laid the foundation of modern asset pricing theory. 30 years later (1993), Eugene F. Fama and Kenneth R. French came up with two new factors, namely size and value, in addition to the market factor. In the subsequent years (2014), Fama and French improved their model by introducing two additional factors, forming the well-known Five-Factor Asset Pricing Model. This model is able to capture the exposure to value, size, profitability, and investment patterns in any particular stock or portfolio return series. These models are still widely used by practitioners which often assume constant factor loadings. However, the validity of this assumption is becoming more and more questionable. Ferson and Harvey[1] (1999) has shown that the time variation in the loadings of factors has significant explanatory power over portfolio returns. Fama and French[11] (1997) has also suggested that the large standard errors of parameter estimates are due to imprecise estimations of the factor loadings. This lead us to believe that factor loading in fact tends to vary over time significantly.

There has already been some work on the topic of time-varying factor models in the form of regime-shifting models, as proposed by Diebold and Rudebusch[3] (1994), where they used two regimes modeled as a Markov process with time-varying transition probabilities. Tong and Lim[4] (1980) also developed regime shifting methods in the form of threshold models, where they had different models when certain parameters passes certain thresholds and hence the name. However, Lettau and Van Nieuwerburgh[5] (2006) has shown that regime shifting models in general have poor out-of-sample performance because first the shifts between regimes are difficult to detect, and the mean values in different regimes are estimated with large uncertainty. Therefore, in this paper we take a different view on the problem: instead of having a discrete number of regimes, we allow the factor loadings to vary in real time, which essentially corresponds to a continuum (infinite number) of states instead of discrete jumps between states.

Existing literature in this area includes Andrian and Franzoni's work "Learning about Beta: Time-Varying Factor Loadings, Expected Returns and the Conditional CAPM" [6](2004,2005) where they applied the Kalman Filter procedure to estimate the CAPM's beta across time. Two years later, inspired by their work, Peng Huang was the first to implement this learning type model in a multi-risk scenario (4-factor model) in "Applications of Time-Varying-Parameter Models to Economics and Finance"[7], 2006. In their paper Huang modeled the 4 factors (bond term faction and Fama French 3 factors) each as an auto-regressive process with no correlations between them, however in our model we will also explore the possibility of factor loadings varying around a fixed constant. Nieto, Orbe and Zarraga [9](2014) also compared rolling regression and the Kalman Filter approach to capture time varying coefficients, however their work used the CAPM which has only one factor on the Mexican stock market. We will extend this comparison to the 3 and 5 factor models in the U.S. stock market in our paper.

1.3 Problem Definition

The goal of this paper is to compare a variety of time-varying coefficient factor models to traditional non-time-varying models. More specifically, we will build the time-homogeneous CAPM, Fama-French 3 and 5 factor models along with their time-varying counterparts, and compare their performance in a set of metrics including R-squared, mean squared error and more importantly cross sectional estimation of factor premium, and portfolio construction based on the output of the models. These measures will be discussed in detail in the final section of Execution Plan.

1.4 Execution Plan

Benchmark

As our first step, we build the plain vanilla multi-factor model using OLS regression. We will choose a period at the beginning of our data as the training period and test the coefficients out of sample, assuming they are constant outside the training period. The length of the training period (60 months) is chosen so that it is in-line with other models that we build, so no model have an competitive advantage over others at the beginning. Then we will construct the most basic type of time-varying model using a rolling regression method. Our simple portfolio construction strategy's back-testing performance along with standard factor model statistics will serve as our benchmark. Here we will also analyze evolution of our factor exposure over time, which hopefully will provide us with directions in terms of the state equation that is appropriate for different factor exposures in the time-varying coefficient models.

Time-varying Coefficients

For the testing results to be comparable, we will implement the time-varying coefficient factor models with the same set of factors and the first results from the rolling regression will serve as initial conditions for our Kalman Filter model. For the state equations, we will compare two methods: random coefficient and random walk. Random coefficient is when we model $\beta_t = \beta_i + \epsilon_t$ where random walk implies $\beta_t = \beta_{t-1} + \epsilon_t$. The difference being that random coefficients assumes a constant offset in the state equation while random walk uses the previous value as the offset. The choice of the state equations will be based on our economic intuition and also the insights that we gain from the plain vanilla model.

Setting the Evaluation Metrics of Factor Model

We separate the analysis into three parts:

- Consistency Check of the Coefficients:
  Here we will look at the correlation of the coefficients over time from different models, and whether the pattern over time coincides with our basic intuition. For example during large market draw downs, the portfolios' exposures to market factor should increase.

- Asset Pricing Aspect:
  In this section we look at how well does the model 'explain' the portfolio returns. More specifically, we will test the models in R-squared, mean squared errors. Also once we obtain our coefficients, we perform cross sectional regression of portfolio returns on the estimated coefficients, to calculated the implied factor premium. If the model is accurate, we should expect that the implied factor returns are close to actual factor returns and on average non negative.

- Portfolio Management Aspect:
  Equipped with the factor exposures across time, we can calculate the covariance and expected return of the assets in the next period. This allows us to build portfolios using mean variance analysis. If our models are accurate and consistent, we should expect good performance when we construct our portfolios. For example if we construct a minimum variance portfolio, a good factor model should produce a portfolio with low variance.

One problem worth pointing out with the R-squared is that because we are building time-varying coefficient models, predictions each period are results of different regressions, hence the conventional R-squared definition will not be valid because it is not necessarily bounded by 0 and 1. Instead, we use the unconditional variance method as described by Harvey, Solnik, Solnik and Zhou [8](2002), where the proportion or explained variance is defined by:

$$VR1 = \frac{Var(\mathbb{E}_M(R_i))}{Var(\mathbb{E}_Z(R_i))} \tag{1.1}$$

where $\mathbb{E}_M(R_i)$ is the model prediction of the return and $\mathbb{E}_Z(R_i)$ is the prediction obtained from the regression:

$$\mathrm{R}_{i,t} = \alpha_i + \beta_{i1}\mathrm{f}_{1,t} + \beta_{i2}\mathrm{f}_{2,t} + \beta_{i3}\mathrm{f}_{3,t} + \beta_{i4}\mathrm{f}_{4,t} + \beta_{i5}\mathrm{f}_{5,t} + \epsilon_{i,t} \tag{1.2}$$

Note here $R_{i,t}$ and $\mathrm{f}_{k,t}$ are both values at time t. This ex-post regression can be seen to capture the true realized exposures. Therefore $VR1$ is a way of estimating the proportion of time variation of the predictions that is explained by our model, the closer to 1 the better. Similarly, we can define:

$$VR2 = \frac{Var(\mathbb{E}_M(\epsilon_i))}{Var(\mathbb{E}_Z(R_i))} \tag{1.3}$$

as the proportion of unexplained variance in the model and we would like this to be as close to 0 as possible.

Desired Results

We believe that with increasing complexity, the performance of the model improves, and hence the Kalman Filter should out-perform the simple rolling OLS model in most metrics if not all.

2 Data Description

2.1 Data Sources

The data that we use can be split into two types. The first type is the Fama-French factor returns time series from their website and their risk-free rate time series, this is the basic constituent of our factor model. The 5 factor returns we use are constructed by Eugene F. Fama and Kenneth R. French in their working paper 'A Five-Factor Asset Pricing Model', which are:

- Market Excess Return:

  Constructed as the value-weighted returns over the risk-free rate of all stocks listed on NYSE, AMEX, or NASDAQ.

- SMB:

  Small minus big, defined as the equal weighted average return of 9 small stock portfolios minus 9 big stock portfolios.

- HML:

  High minus low, defined as the equal weighted average return of 2 value stock portfolios minus 2 growth stock portfolios

- RMW:

  Robust Minus Weak, defined as the equal weighted average return of 2 robust operating profitability portfolios minus 2 weak operating profitability portfolios.

- CMA:

  Conservative Minus Aggressive, defined as the equal weighted average return of 2 conservative investment portfolios minus 2 aggressive investment portfolios.

The second type spans the period between 1963 and 2017 and takes the form of S&P 500 stock daily total returns, calculated from daily prices at market close from WRDS. We use these returns in two ways: first to construct target portfolios that act as the dependent variable in our factor model and second, to perform portfolio optimization to test our models.

2.2 Choice of Universe

The reason for using portfolios to test our factor models is because single stock returns are extremely noisy, any prediction power our model have can be easily swamped by the idiosyncratic noise of the stocks, hence it is standard practice in the industry to test factor model on portfolios.

To form the target portfolios, we sorted the S&P500 stocks according to traded volume from 1964 until 2017. The reason behind choosing traded volume as the differentiator is that this is a characteristics that is not included in the factors we use, and hence should produce interesting coefficient patterns over time. If we were to use portfolios sorted based on one of the factors in the model, it would be very well captured and the coefficients would have little variation over time because majority of the exposure is to that one factor. To create the portfolios, each month we take the 500 stocks and sort them by traded volume into 6 portfolios, and within each portfolio the return is capitalization weighted. At the end of each month we sort again and re-balance the portfolios.

For our final tests under the portfolio management aspect, we would like to restrict ourselves to the most liquid instruments in the equity market, hence the S&P500 is a natural choice. However stocks tends to move in and out of the index over time, and for our portfolio construction we would like to keep the asset universe constant. This is because during reconstitution of the index, there are behavioral phenomena that can occur due to demand or supply shocks, which often lead to temporary large mis-pricings. We will not explore these behavioral anomalies here, therefore instead we simply chose 200 stocks that have continuous data from 1978 until present.

2.3 Portfolio Analysis

Before we dive into the model, we ran some initial analysis on the volume sorted portfolios to gain more insights into the data. First of all, for each portfolio we plotted the portfolio cumulative returns, the average traded volume and average market capitalization over time in Figures 1, 2, 3 and 4. For the cumulative returns, we split the time period into 1964 to 2002 and 2002 to 2018. This is because we can see that before 2002, the higher trade volume, the higher cumulative returns in general with the exception of portfolio 6 which has the highest traded volume. However this pattern is reversed after 2002. This indicates that the liquidity premium did not exist pre-2002. The introduction of this new risk premium into the market occurred contemporaneously with the burst of the internet bubble, and hence people became aware that providing liquidity should in general earn a premium for bearing the risk of not being able to complete desired trades when majority of the market is trading in the same direction.

Figure 1: Cumulative Portfolio Returns from 1964 to 2002

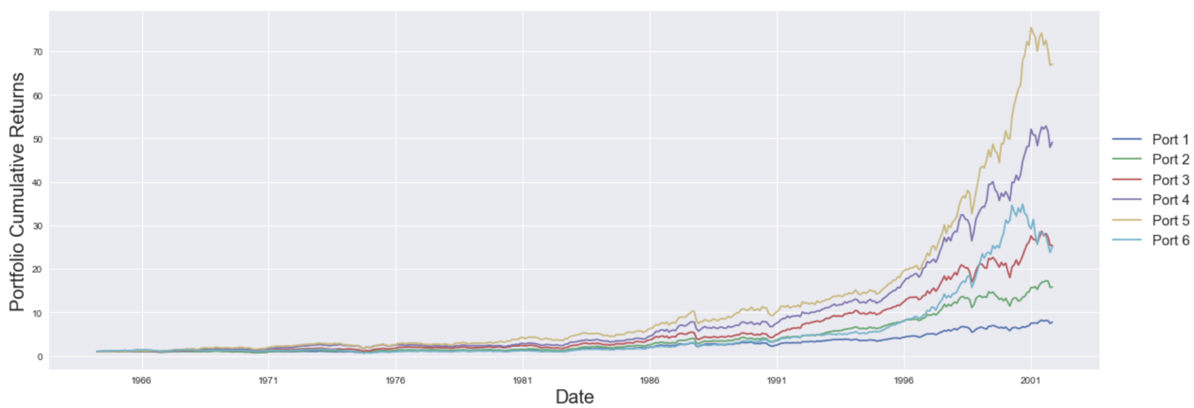


Figure 2: Cumulative Portfolio Returns from 2002 to 2018

Figure 3

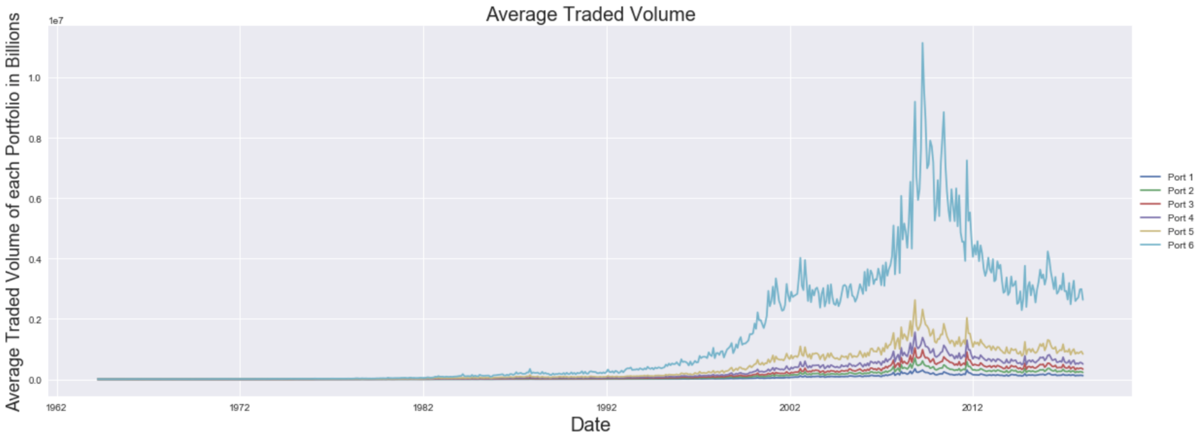
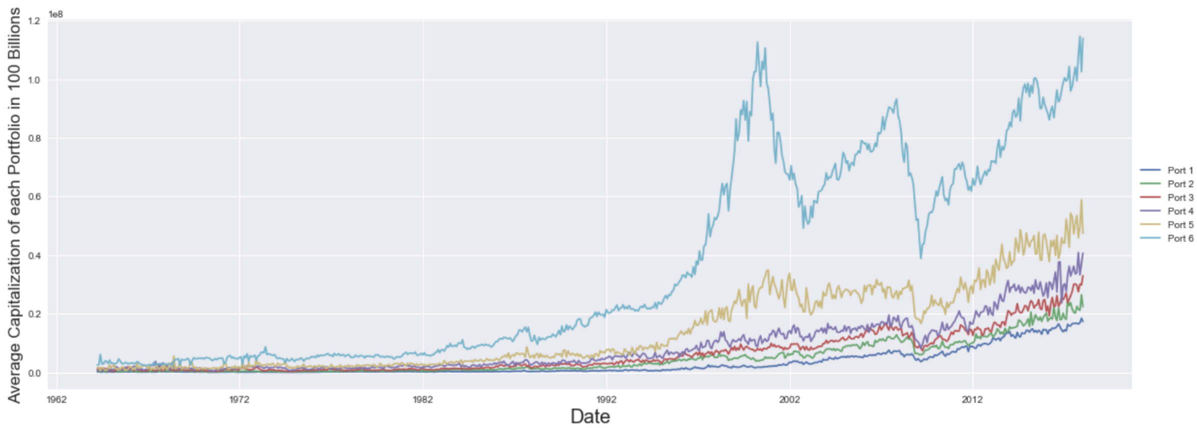


Figure 4



Next we look at the plot of average traded volume and the market capitalization in Figures 3 and 4. Keeping in mind that portfolio 6 has the largest traded volume, we see that its volume is significantly larger than the other five starting from 1990s. Before the 90s due to limitations in technology, the traded volumes for all portfolios are a lot smaller. And unsurprisingly, the market capitalization is also larger for portfolio 6 than all the other 5 combined, since the largest stocks in the market tends to be the most frequently traded. A more interesting question is the composition of these 6 portfolios, because this will help explain their factor exposures over time later on. To do this, we used the Standard Industrial Classification Code (SICC). At the end of each month, we look at the stocks in the portfolio and for each SICC, we count the number of stocks that is in the particular industry, and the industry with the maximum number of stocks is the dominant industry for that portfolio for that month. Then we aggregate across 10 year periods and plot a stacked bar plot where each bar indicates for each portfolio, the number of months that a particular industry is dominant for that 10 year period. The plot is shown in figures 5 to 10:

Figure 5: Portfolio 1 Industry Decomposition

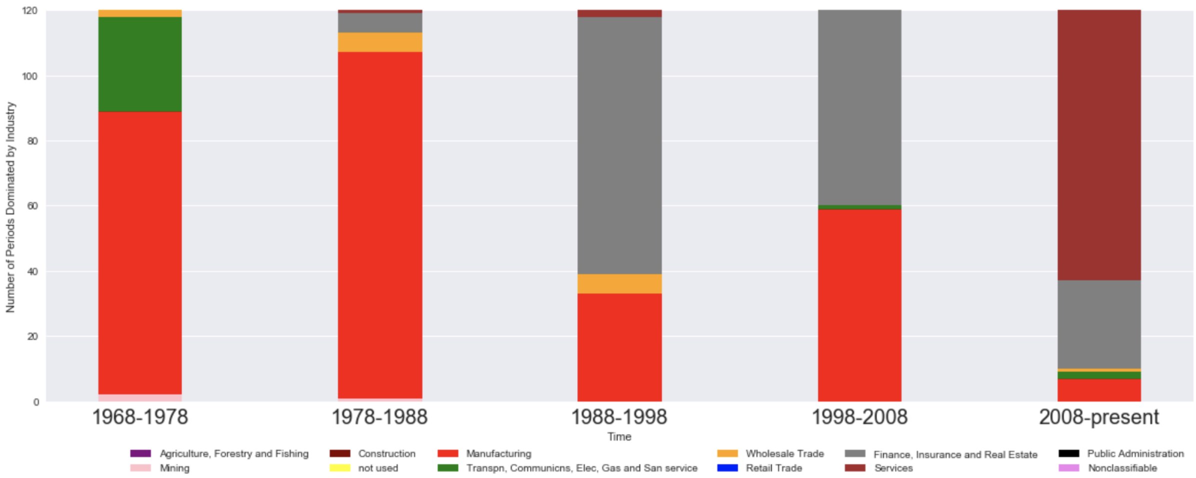


Figure 6: Portfolio 2 Industry Decomposition

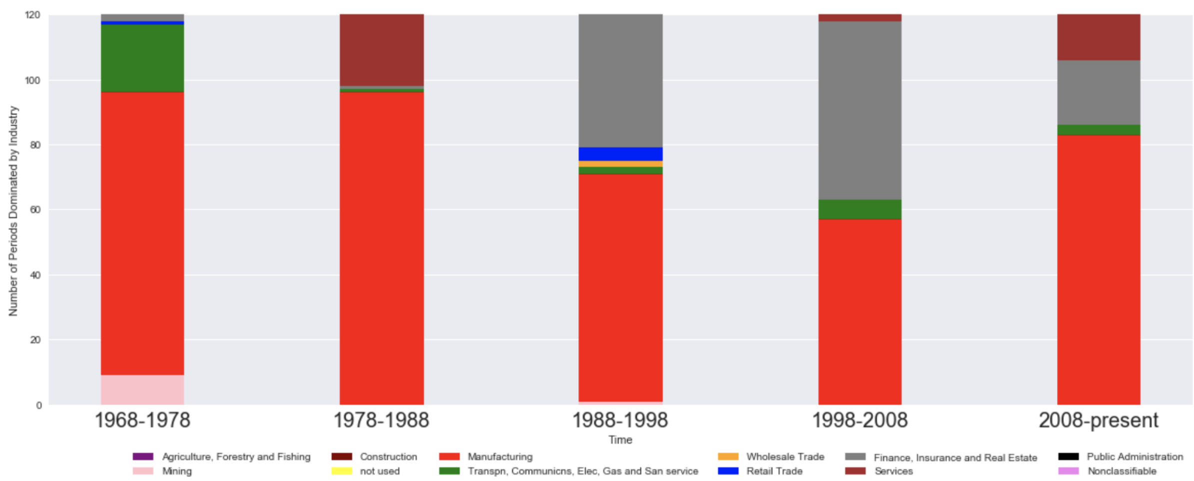


Figure 7: Portfolio 3 Industry Decomposition

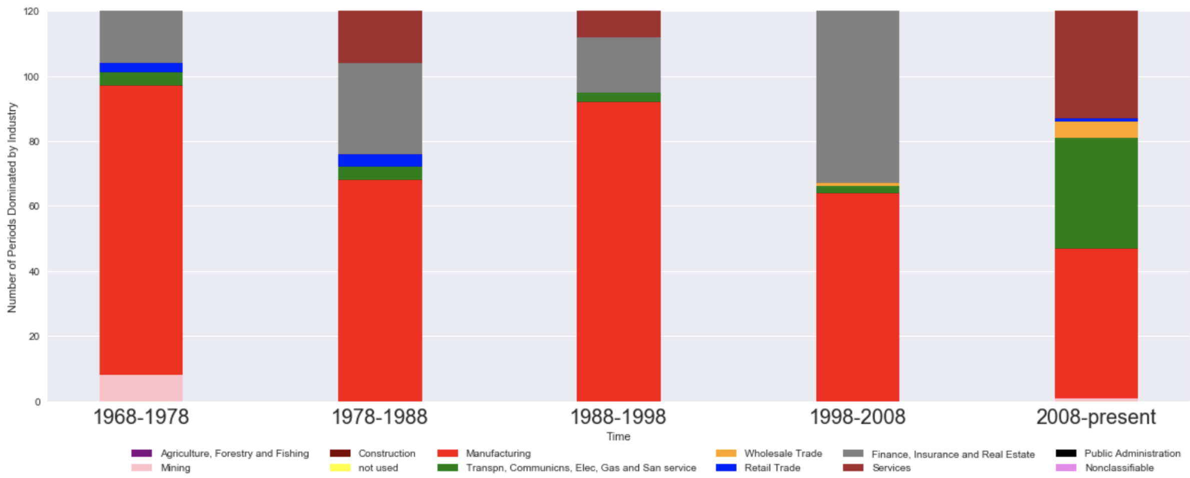


Figure 8: Portfolio 4 Industry Decomposition

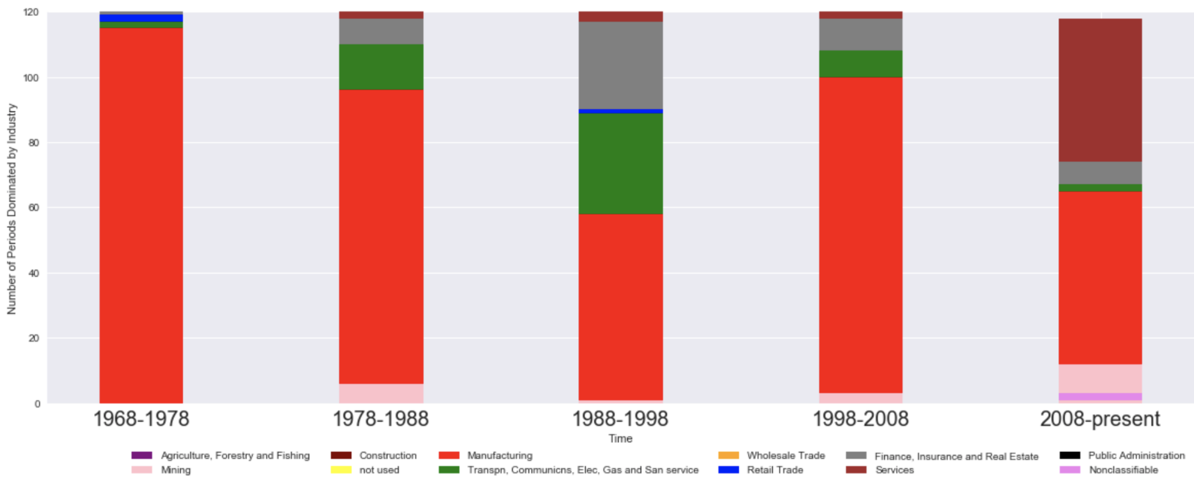


Figure 9: Portfolio 5 Industry Decomposition

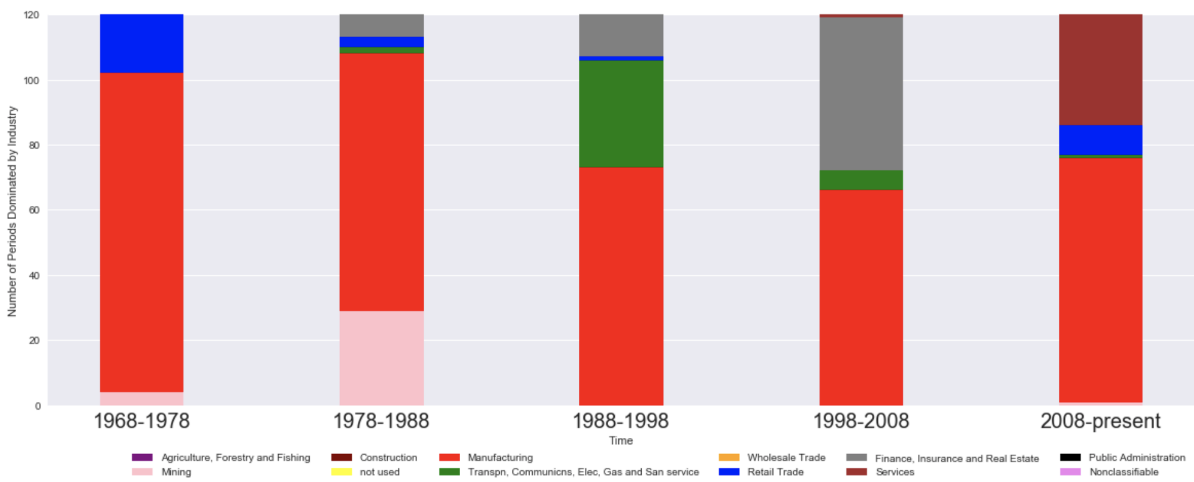
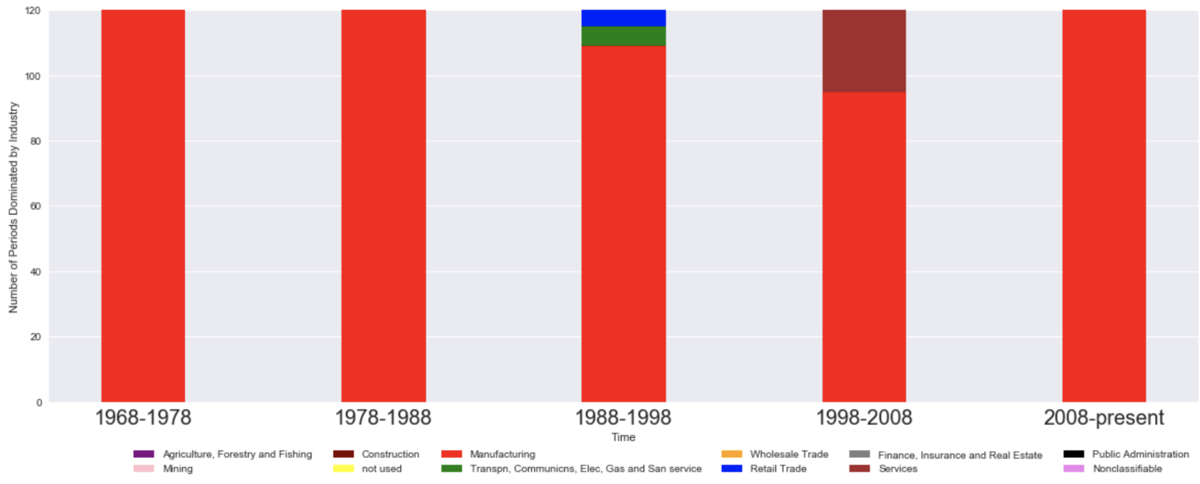


Figure 10: Portfolio 6 Industry Decomposition



Portfolio 6 being the most traded, is dominated by manufacturing (red). Note manufacturing is also the major component for portfolios of all trading frequencies over almost all the five 10 year periods. This is logical because it is the most prevalent industry and there are companies of all sizes within this industry which lead to its appearance in all trading frequencies. We can also observe the increase across time in the amount of months that is dominated by finance (gray). Construction is another significant sector however in recent years it exists in the less traded portfolios.

Models

3 Models

3.1 Basic Model: None Time-Varying Factor Models

Our basic model involves a simple approach that is common practice when building factor models. We split the sample period into training set and testing set, where we use the first 60 months as the training period. From this we obtain our set of coefficients, and assume that they are constant from then on to make predictions of the returns of each stock out of sample. The regression takes following form for a 5-factor model and similarly for 3-factor models and the CAPM:

$$\begin{bmatrix} R_{i,t} \\ R_{i,t-1} \\ \vdots \\ R_{i,t-k} \end{bmatrix} = \alpha\underline{\mathbb{1}} + \begin{bmatrix} f_{1,t} & f_{2,t} & \cdots & f_{5,t} \\ f_{1,t-1} & f_{2,t-1} & \cdots & f_{5,t-1} \\ \vdots & \vdots & \ddots & \vdots \\ f_{1,t-k} & f_{2,t-k} & \cdots & f_{5,t-k} \end{bmatrix} \begin{bmatrix} \beta_{i1} \\ \beta_{i2} \\ \vdots \\ \beta_{i5} \end{bmatrix} + \begin{bmatrix} \epsilon_{i,t} \\ \epsilon_{i,t-1} \\ \vdots \\ \epsilon_{i,t-59} \end{bmatrix} \tag{3.1}$$

(3.2)

where $\underline{\mathbb{1}}$ is a vector of 1s, $f_{i,t-s}, i = 1, ..., 5$ are the 5 factors returns described above respectively at time $t - s$, $\epsilon_{i,t-s}$ is the residual of the regression and R_{it} is the in sample return of stock i at time t. Notice here $\beta_{ik}, k = 1, ..., 5$ does not depend on t. Our prediction for next period is:

$$\overline{\mathrm{R}}_{i,s+1} = \alpha_i + \beta_{i1}\mathrm{f}_{1,s} + \beta_{i2}\mathrm{f}_{2,s} + \beta_{i3}\mathrm{f}_{3,s} + \beta_{i4}\mathrm{f}_{4,s} + \beta_{i5}\mathrm{f}_{5,s} \tag{3.3}$$

where s indicates out of sample period and $\overline{R}$ is the prediction.

3.2 Benchmark Model: Time-Varying with Rolling Regression

In order to stay consistent with the basic model, we begin the rolling regression by using the same 60-month period, and every month we move forward the rolling window, to predict next period's returns.

So, at each time s, for each stock i, we have 5 time series of coefficients:

$$\begin{bmatrix} \beta_{i,1,t} & \beta_{i,2,t} & \cdots & \beta_{i,5,t} \\ \beta_{i,1,t-1} & \beta_{i,2,t-1} & \cdots & \beta_{i,5,t-1} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{i,1,t-k} & \beta_{i,2,t-k} & \cdots & \beta_{i,5,t-k} \end{bmatrix} \tag{3.4}$$

3.3 Advance Models: Time-Varying with Kalman Filter

The above methods assume no models for the coefficients, instead they use only the data and try to choose parameters that fit to the data with a regression based approach. However linear regression is not always the best choice for several well-known reasons. First of all, they tend to be extremely sensitive to outliers in the sample. This problem occurs almost by definition, since linear regressions are done by minimizing the sum of squared error, and outliers can dominate the sum easily. Another problem is that using the rolling window, we are discarding data that is far enough in the past. This can be a desirable property if one believes that data far in the past is irrelevant, but whether that is the case or how far back do we have to go for this assumption to hold is a difficult question to answer and one has to be very careful with data mining when choosing such parameters. These are just a few among many issues that regression based approaches can potentially have.

The Kalman Filter on the other hand takes a Bayesian view of the world, by assuming a prior distribution of the parameters, and given a new observation, updates to a posterior through maximizing the likelihood and makes a prediction for the next period. This method circumvents the problems of regression methods stated above however the disadvantage being that we need to feed in the initial conditions and make distribution assumptions. This issue can be easily mitigated because given that we have long enough data in the past, the accuracy of the initial conditions will not have a big impact because the model learns new information each period and historical information matter less as time moves forward. There are two equations in the model, the state equation, which governs the evolution of the parameters, and the measurement equation, which states the relationship between the observed variables. For us the measurement equation will be the factor model equation and state equation is the coefficient evolution:

$$\boldsymbol{y_t} = \boldsymbol{\alpha_t} + \boldsymbol{F_t}\boldsymbol{\beta_t} + \boldsymbol{\epsilon_t} \qquad \text{(Measurement Equation)} \tag{3.5}$$

$$\boldsymbol{\beta_{t+1}} = \boldsymbol{d_t} + \boldsymbol{T_t}\boldsymbol{\beta_t} + \boldsymbol{u_t} \qquad \text{(Measurement Equation)} \tag{3.6}$$

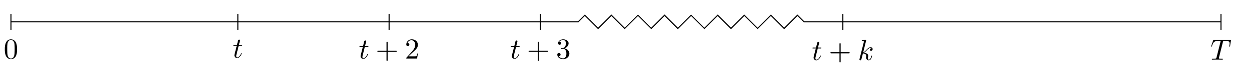
where :

- $\boldsymbol{y_t}$ is the vector of returns on the target portfolio that we are trying to analyze across time;
- $\boldsymbol{\alpha_t}$ is the vector of returns not explained by the factors;
- $\boldsymbol{F_t}$ is the matrix of factor returns;
- $\boldsymbol{\beta_t}$ is the vector of coefficients of the factors at time t;
- $\boldsymbol{\epsilon_t}$ is the vector of measurement noise, which are assumed to have joint normal distribution with mean 0 and covariance matrix $\boldsymbol{R_t}$;
- $\boldsymbol{d_t}$ is the vector of offsets of the coefficients;
- $\boldsymbol{T_t}$ is the matrix that describe the relation between the coefficients;
- $\boldsymbol{u_t}$ is the vector of noise in each state variable and is assumed to be jointly normal with mean 0 and covariance matrix $\boldsymbol{Q_t}$.

This is the most general form of linear state space models. In theory $\boldsymbol{T_t}$ can vary with time, however for now we assume it is constant. In addition, Nieto, Orbe and Zarraga [9](2014) used two sub-categories of this as their time-varying coefficient model with one factor, which are special cases of the general form in the following sense:

- Random Walk Models (RW):
  This is when the state equation is modeled as a random walk, hence $\boldsymbol{d_t} = \boldsymbol{0}$ and $\boldsymbol{T_t} = \boldsymbol{I}$;
- Random Coefficient Models (RC):
  This is when the state equation is modeled as a constant plus white noise, hence $\boldsymbol{T_t} = \boldsymbol{0}$.

In our paper, we will test both RW and RC. The input to the model is the initial estimate of the coefficients and the covariance of the coefficients. In order to be consistent with rolling regression models, we need to make sure the input is the same as the output of the rolling regression at a certain period to make sure neither model has an advantage at the beginning compared to the other. More specifically, as shown by the illustration below:



If at time t we have our first rolling regression result, we continue the rolling model for k periods afterwards until time $t+k$. Therefore, at $t+k$ we have our current coefficient estimate from the rolling regression and the covariance matrix of the $\boldsymbol{\beta}$s from their historical values. At this point we start our Kalman Filter model with these inputs as the initial condition. The estimation process is as follows:

1. Given initial conditions of $\boldsymbol{\beta(0)}, \boldsymbol{\alpha(0)}$, calculate the prediction $\widehat{\boldsymbol{y}}\boldsymbol{(1)}$ as:

   $\widehat{\boldsymbol{y}}\boldsymbol{(1)} = \boldsymbol{\alpha(0)} + \boldsymbol{F(1)}\boldsymbol{\beta(0)}$

2. Observe the true value $\boldsymbol{y(1)}$ and calculate the error as:

   $\boldsymbol{e(0)} = \boldsymbol{y(1)} - \widehat{\boldsymbol{y}}\boldsymbol{(1)}$

3. Calculate the updated coefficient as:

   $\boldsymbol{\beta_{new}(0)} = \boldsymbol{\beta(0)} + K\boldsymbol{e(0)}$

   where K is the Kalman Filter Gain estimated using maximum likelihood method. Note here we have absorbed the $\boldsymbol{d_t(0)}$ in the state equation (3.6) into $\boldsymbol{\beta(0)}$ by adding a column of 1 s to $\boldsymbol{T_t}$.

4. Finally predict the new state estimate as:

   $\boldsymbol{\beta(1)} = \boldsymbol{T}\boldsymbol{\beta(0)}$

This completes one step of filtering and updating, we repeat this process each period to obtain new estimates for the next period which in turn act as our prior for the period after that. We implemented the Kalman Filter using two methods described as above, and they will be labeled RC (random coefficient) and RW (random walk).

Consistency Analysis

4 Consistency Analysis

4.1 Coefficient Statistics

As a preliminary test of the models, we calculated the mean and standard deviations of the coefficients across time for our 6 volume sorted portfolios with the smallest volume as portfolio 1. To estimate the first coefficient value in the stationary and rolling regression, we used data from 1964-03-31 to 1969-02-28. Then, we moved forward the window in a monthly manner from 1969-02-28 to 1974-01-31 to calculate a time series evolution of the coefficients. On 1974-07-31, our real comparison of the models begins, because we take the mean and covariance matrix of the coefficients estimated from rolling regression on that day as the initial condition for the 2 Kalman Filter models. We then continue the estimation of all models until end of November 2017. The mean values and standard deviations of the coefficients over this period is summarized in Tables 1 to 4:

Table 1: Constant Model Coefficient Summary

		Portfolio 1		Portfolio 2		Portfolio 3		Portfolio 4		Portfolio 5		Portfolio 6	
		mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
CAPM	alpha	0.001	0.0	0.005	0.0	0.007	0.0	0.010	0.0	0.009	0.0	0.002	0.0
	mkt	0.828	0.0	0.993	0.0	0.955	0.0	0.818	0.0	0.899	0.0	0.822	0.0
3 Factor	alpha	0.002	0.0	0.006	0.0	0.007	0.0	0.010	0.0	0.009	0.0	0.002	0.0
	mkt	0.742	0.0	0.963	0.0	0.928	0.0	0.841	0.0	0.946	0.0	0.851	0.0
	SMB	0.282	0.0	0.064	0.0	0.078	0.0	-0.110	0.0	-0.202	0.0	-0.073	0.0
	HML	0.242	0.0	-0.007	0.0	0.045	0.0	-0.157	0.0	-0.266	0.0	-0.020	0.0
5 Factor	alpha	0.001	0.0	0.005	0.0	0.008	0.0	0.011	0.0	0.009	0.0	0.000	0.0
	mkt	0.747	0.0	0.969	0.0	0.929	0.0	0.848	0.0	0.931	0.0	0.861	0.0
	SMB	0.346	0.0	0.088	0.0	0.072	0.0	-0.144	0.0	-0.200	0.0	0.081	0.0
	HML	0.478	0.0	0.036	0.0	0.006	0.0	-0.386	0.0	-0.110	0.0	0.578	0.0
	RMW	0.328	0.0	0.126	0.0	-0.028	0.0	-0.159	0.0	-0.010	0.0	0.784	0.0
	CMA	-0.099	0.0	0.032	0.0	0.036	0.0	0.215	0.0	-0.237	0.0	-0.286	0.0

Table 2: Rolling Regression Model Coefficient Summary

		Portfolio 1		Portfolio 2		Portfolio 3		Portfolio 4		Portfolio 5		Portfolio 6	
		mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
CAPM	alpha	0.004	0.004	0.005	0.003	0.005	0.004	0.006	0.004	0.006	0.004	0.004	0.004
	mkt	0.862	0.139	0.904	0.164	0.911	0.164	0.903	0.136	0.892	0.104	1.004	0.073
3 Factor	alpha	0.002	0.004	0.003	0.003	0.004	0.003	0.006	0.003	0.006	0.003	0.005	0.004
	mkt	0.894	0.080	0.972	0.107	0.979	0.094	0.954	0.078	0.939	0.073	1.000	0.062
	SMB	0.120	0.119	0.005	0.088	-0.050	0.088	-0.136	0.165	-0.214	0.071	-0.273	0.084
	HML	0.135	0.276	0.174	0.272	0.148	0.287	0.050	0.249	-0.025	0.187	-0.082	0.176
5 Factor	alpha	0.002	0.004	0.003	0.003	0.004	0.003	0.005	0.003	0.006	0.003	0.005	0.004
	mkt	0.924	0.079	0.999	0.093	0.997	0.083	0.973	0.071	0.964	0.072	1.005	0.062
	SMB	0.143	0.113	0.013	0.083	-0.033	0.098	-0.131	0.178	-0.213	0.083	-0.273	0.098
	HML	0.048	0.238	0.107	0.230	0.116	0.226	0.011	0.216	-0.076	0.116	-0.067	0.268
	RMW	0.088	0.183	0.080	0.186	0.053	0.165	0.063	0.125	0.064	0.147	0.180	0.246
	CMA	0.123	0.189	0.099	0.282	-0.011	0.262	0.064	0.187	0.088	0.194	0.094	0.247

Table 3: KF with RC Model Coefficient Summary

		Portfolio 1		Portfolio 2		Portfolio 3		Portfolio 4		Portfolio 5		Portfolio 6	
		mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
CAPM	alpha	0.002	0.010	0.003	0.010	0.004	0.010	0.006	0.010	0.007	0.009	0.002	0.008
	mkt	0.822	0.001	0.959	0.001	0.933	0.001	0.871	0.001	1.020	0.001	0.966	0.000
3 Factor	alpha	0.001	0.010	0.003	0.011	0.005	0.011	0.007	0.010	0.008	0.010	0.003	0.007
	mkt	0.772	0.001	0.953	0.001	0.996	0.001	0.907	0.001	1.040	0.001	1.036	0.000
	SMB	0.142	0.001	-0.009	0.001	-0.163	0.001	-0.136	0.000	-0.128	0.000	-0.176	0.000
	HML	0.079	0.001	-0.068	0.001	-0.109	0.001	-0.187	0.001	-0.310	0.001	-0.062	0.000
5 Factor	alpha	0.001	0.010	0.002	0.010	0.004	0.010	0.007	0.011	0.007	0.008	0.000	0.013
	mkt	0.796	0.001	0.976	0.001	1.009	0.001	0.935	0.001	1.020	0.001	1.032	0.001
	SMB	0.184	0.001	0.031	0.001	-0.136	0.001	-0.101	0.001	-0.104	0.000	-0.097	0.001
	HML	-0.022	0.001	-0.126	0.001	-0.123	0.001	-0.387	0.001	0.100	0.000	0.477	0.001
	RMW	0.079	0.001	0.149	0.001	0.114	0.001	0.028	0.001	0.369	0.000	0.703	0.001
	CMA	0.184	0.000	0.179	0.000	0.105	0.000	0.273	0.000	-0.265	0.000	-0.189	0.000

Table 4: KF with RW Model Coefficient Summary

		Portfolio 1		Portfolio 2		Portfolio 3		Portfolio 4		Portfolio 5		Portfolio 6	
		mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
CAPM	alpha	0.004	0.014	0.004	0.013	0.005	0.014	0.005	0.013	0.005	0.012	0.004	0.011
	mkt	0.864	0.118	0.920	0.153	0.904	0.148	0.908	0.122	0.906	0.104	1.021	0.074
3 Factor	alpha	0.003	0.012	0.004	0.011	0.004	0.011	0.005	0.011	0.006	0.010	0.005	0.010
	mkt	0.886	0.065	0.967	0.097	0.956	0.074	0.933	0.064	0.925	0.076	1.015	0.054
	SMB	0.092	0.098	-0.010	0.072	-0.089	0.056	-0.146	0.132	-0.206	0.047	-0.293	0.057
	HML	0.127	0.212	0.146	0.220	0.126	0.244	0.040	0.229	-0.052	0.181	-0.111	0.136
5 Factor	alpha	0.002	0.011	0.002	0.010	0.003	0.010	0.005	0.011	0.005	0.010	0.003	0.011
	mkt	0.929	0.071	1.008	0.084	0.985	0.050	0.956	0.060	0.952	0.054	1.032	0.041
	SMB	0.132	0.084	0.026	0.063	-0.056	0.075	-0.126	0.147	-0.189	0.049	-0.249	0.051
	HML	0.009	0.158	0.043	0.165	0.076	0.190	-0.075	0.235	0.035	0.053	0.077	0.223
	RMW	0.126	0.139	0.132	0.132	0.095	0.117	0.053	0.087	0.159	0.112	0.302	0.289
	CMA	0.192	0.063	0.187	0.081	0.073	0.073	0.171	0.067	-0.025	0.173	-0.081	0.088

Note the standard deviation for the constant coefficient model is 0 in Table 1 because after the training period we assume the coefficients do not change. We can see for all models, from portfolio 1 to 6 the average market exposure is increasing, and average exposure to SMB is decreasing, which means more volume implies higher market exposure and larger size. This confirms our intuition that the market is dominated by large capitalization stocks and high volume stocks (high liquidity and hence large capitalization). Alpha in general decrease with the addition of factors, so the factors do help decompose the returns on the portfolios.

Figure 11: Time Series of Coefficients for the 5 Factors Model, 1974 - 2018

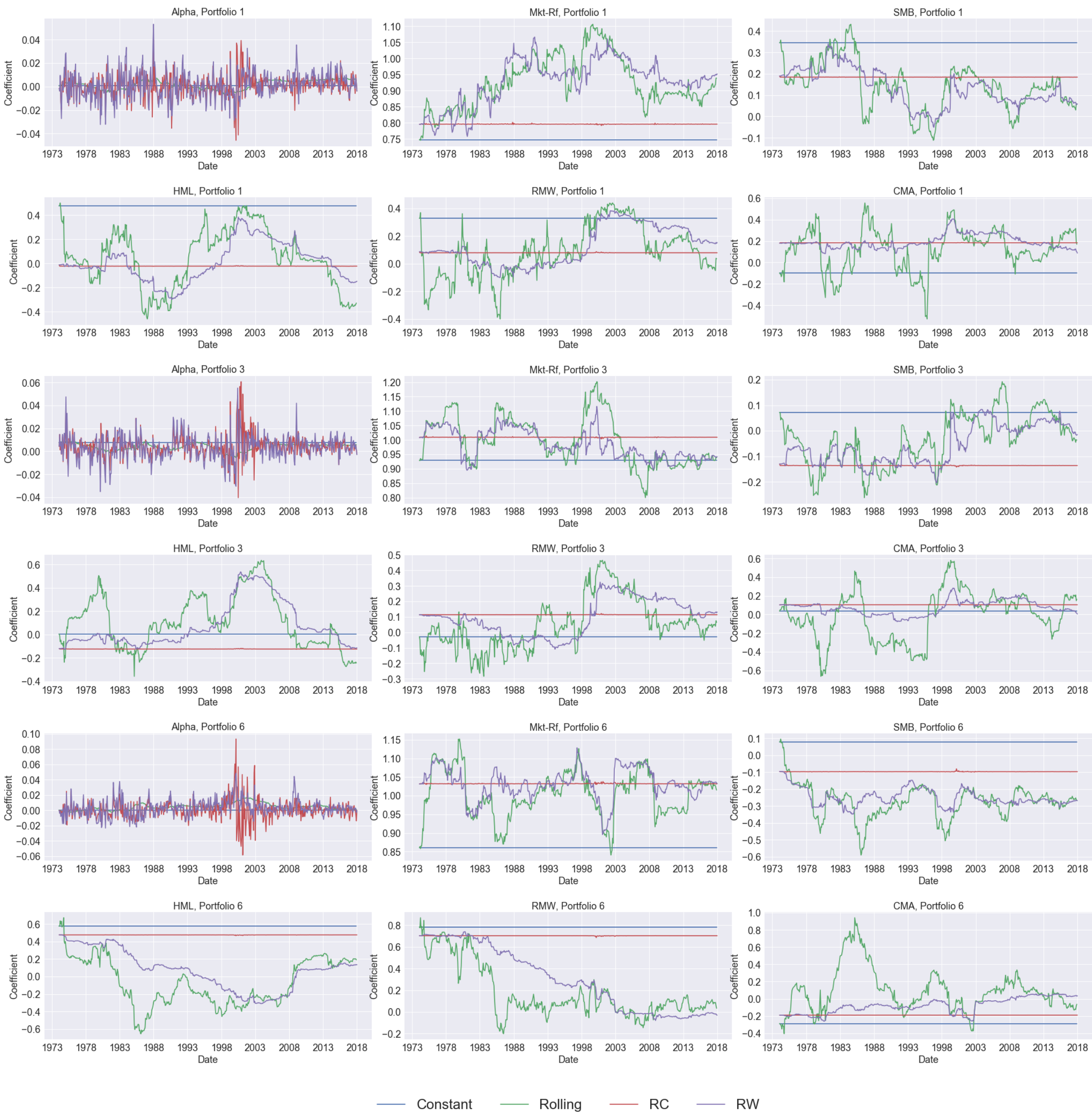


Figure 12: Time Series of Coefficients for the 3 Factors Model, 1974 - 2018

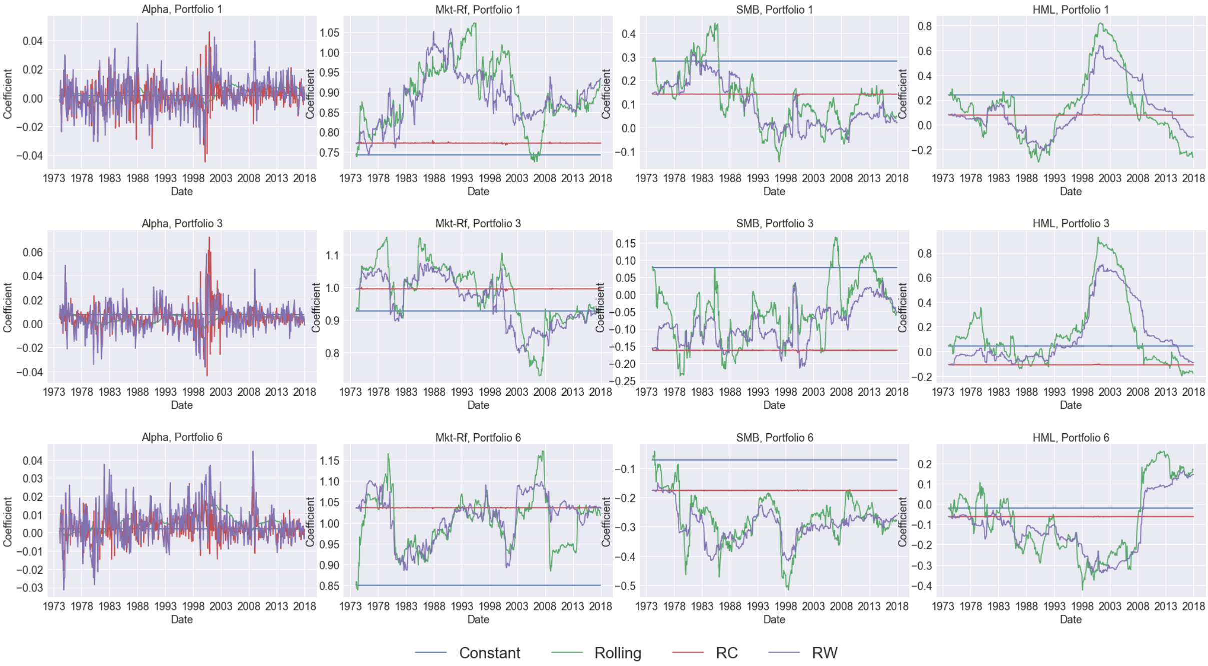
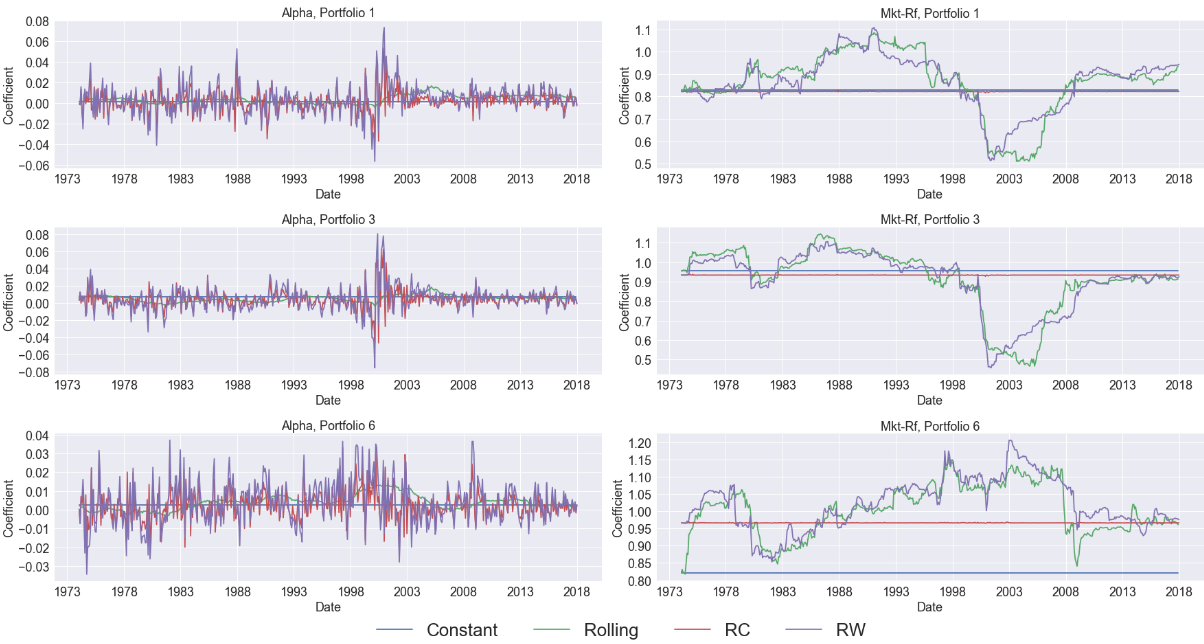


Figure 13: Time Series of Coefficients for CAPM, 1974 - 2018



4.2 Coefficient Evolution

Next we plotted the time series of our coefficients in Figures 11 to 13, on each plot we show the coefficient of one factor for 4 different models for portfolios 1,3 and 6. First thing to note is that the blue line is always flat, this is because the constant model does not have changing coefficients, hence the evolution is simply a flat line. This line serves as a yardstick and shows just how much the other coefficients can deviate from its original value. Similarly, the Kalman Filter with random coefficient (red line) has very small volatility as well. This is because by definition of RC, the coefficients vary around a constant value, and the only way it can vary is through the small noise term, which is estimated from observed values. The reason why constant and rolling method have a different starting point from the two Kalman Filter methods is because the two Kalman Filters' inputs requires a covariance of historical coefficients, and the mean value over this period is taken as the initial guess of parameters while the other two methods have the currently value as the starting value. Another observation is that the Kalman Filter RW is in general more smooth than the rolling regression, this is because it is using all the information available, implicitly through past estimates while the rolling method only uses the past 5 years of data. This has the benefit of filtering out some of the noise in the time series. We can also see patterns in the exposure that confirms our intuition, by looking at the size factor for example. Portfolio 6 (highest traded volume) tends to have negative exposure to SMB while portfolio 1 has in general positive exposure. Highest traded volume stocks are the large stocks in the market and hence they have negative exposure to SMB (long small short big) factor. In the middle, portfolio 3's exposure to size seems to have reversed sign from negative to positive in 1999 and 2000. This could have been a result of the increasing awareness of liquidity premium mentioned before, and people are not willing to take the risk. As a result, the traded volume has shifted towards large stocks as a whole.

4.3 Coefficient Correlation

Next we calculated the correlation between coefficient estimated from different models for each portfolio and for each factor. Here we presented the results for portfolio 1,3,6 only in Tables 5 to 23. Looking at the correlations, at first glance there are no significant negative correlations, which means that the estimation from different models do not contradict each other. Another observation is that the correlation between RW and rolling regression is high for majority of the factors. Again confirming that across different models the results are consistent. Overall the different methods show consistency among themselves and confirms our intuition about relationship between the factors and the return characteristics of the portfolios.

Portfolio 1, Correlation of Factors Between Models, 1974 - 2018

	Constant	Rolling	RC	RW
Constant	1.0	0.000	0.000	0.000
Rolling	0.0	1.000	0.120	0.285
RC	0.0	0.120	1.000	0.774
RW	0.0	0.285	0.774	1.000

Table 5: Alpha

	Constant	Rolling	RC	RW
Constant	1.0	-0.000	0.000	0.000
Rolling	-0.0	1.000	-0.062	0.766
RC	0.0	-0.062	1.000	0.022
RW	0.0	0.766	0.022	1.000

Table 6: Mkt-Rf

	Constant	Rolling	RC	RW
Constant	1.0	-0.000	0.000	-0.000
Rolling	-0.0	1.000	0.063	0.776
RC	0.0	0.063	1.000	0.090
RW	-0.0	0.776	0.090	1.000

Table 7: SMB

	Constant	Rolling	RC	RW
Constant	1.0	-0.000	-0.000	-0.000
Rolling	-0.0	1.000	0.287	0.744
RC	-0.0	0.287	1.000	0.251
RW	-0.0	0.744	0.251	1.000

Table 8: HML

	Constant	Rolling	RC	RW
Constant	1.0	0.000	0.000	-0.000
Rolling	0.0	1.000	0.228	0.692
RC	0.0	0.228	1.000	0.144
RW	-0.0	0.692	0.144	1.000

Table 9: RMW

	Constant	Rolling	RC	RW
Constant	1.0	-0.000	0.000	-0.000
Rolling	-0.0	1.000	0.101	0.414
RC	0.0	0.101	1.000	0.204
RW	-0.0	0.414	0.204	1.000

Table 10: CMA

Portfolio 3, Correlation of Factors Between Models, 1974 - 2018

	Constant	Rolling	RC	RW
Constant	1.0	-0.000	-0.000	0.000
Rolling	-0.0	1.000	0.038	0.176
RC	-0.0	0.038	1.000	0.721
RW	0.0	0.176	0.721	1.000

Table 11: Alpha

	Constant	Rolling	RC	RW
Constant	1.0	-0.000	-0.000	-0.000
Rolling	-0.0	1.000	-0.134	0.759
RC	-0.0	-0.134	1.000	0.098
RW	-0.0	0.759	0.098	1.000

Table 12: Mkt-Rf

	Constant	Rolling	RC	RW
Constant	1.0	-0.000	0.000	0.000
Rolling	-0.0	1.000	0.023	0.737
RC	0.0	0.023	1.000	0.001
RW	0.0	0.737	0.001	1.000

Table 13: SMB

	Constant	Rolling	RC	RW
Constant	1.0	0.000	0.000	0.000
Rolling	0.0	1.000	0.257	0.685
RC	0.0	0.257	1.000	0.327
RW	0.0	0.685	0.327	1.000

Table 14: HML

	Constant	Rolling	RC	RW
Constant	1.0	0.000	0.000	0.000
Rolling	0.0	1.000	0.384	0.591
RC	0.0	0.384	1.000	0.231
RW	0.0	0.591	0.231	1.000

Table 15: RMW

	Constant	Rolling	RC	RW
Constant	1.0	-0.000	-0.000	0.000
Rolling	-0.0	1.000	0.215	0.674
RC	-0.0	0.215	1.000	0.154
RW	0.0	0.674	0.154	1.000

Table 16: CMA

Portfolio 6, Correlation of Factors Between Models, 1974 - 2018

	Constant	Rolling	RC	RW
Constant	0.0	0.000	0.000	0.000
Rolling	0.0	1.000	0.005	0.287
RC	0.0	0.005	1.000	0.626
RW	0.0	0.287	0.626	1.000

Table 17: Alpha

	Constant	Rolling	RC	RW
Constant	1.0	-0.000	0.000	-0.000
Rolling	-0.0	1.000	0.070	0.613
RC	0.0	0.070	1.000	0.045
RW	-0.0	0.613	0.045	1.000

Table 18: Mkt-Rf

click to scroll output; double click to hide

	Constant	Rolling	RC	RW
Constant	1.0	-0.000	-0.000	-0.000
Rolling	-0.0	1.000	-0.174	-0.236
RC	-0.0	-0.174	1.000	-0.045
RW	-0.0	-0.236	-0.045	1.000

Table 19: SMB

	Constant	Rolling	RC	RW
Constant	1.0	-0.000	-0.000	-0.000
Rolling	-0.0	1.000	0.289	0.414
RC	-0.0	0.289	1.000	0.234
RW	-0.0	0.414	0.234	1.000

Table 20: HML

	Constant	Rolling	RC	RW
Constant	1.0	0.000	0.000	0.000
Rolling	0.0	1.000	0.107	0.818
RC	0.0	0.107	1.000	0.092
RW	0.0	0.818	0.092	1.000

Table 21: RMW

	Constant	Rolling	RC	RW
Constant	1.0	-0.00	0.000	0.000
Rolling	-0.0	1.00	0.030	-0.070
RC	0.0	0.03	1.000	0.193
RW	0.0	-0.07	0.193	1.000

Table 22: CMA

Asset Pricing Aspect Analysis

5 Asset Pricing Aspect Analysis

5.1 Explained Variance

In this section we compare the models in their ability to capture the portfolio returns' characteristics, in particular explained variance and the error made over time. First we summarized the VR1 and VR2 as defined before in Table 24:

Table 23: VR1 and VR2, 1978 - 2018

		Portfolio 1		Portfolio 2		Portfolio 3		Portfolio 4		Portfolio 5		Portfolio 6	
		VR1	VR2	VR1	VR2	VR1	VR2	VR1	VR2	VR1	VR2	VR1	VR2
CAPM	Constant	0.73	0.23	0.99	0.21	0.89	0.21	0.67	0.19	0.87	0.15	0.58	0.15
	Rolling	0.81	0.20	0.86	0.17	0.85	0.18	0.84	0.16	0.86	0.14	0.85	0.11
	RC	0.80	0.06	0.91	0.05	0.87	0.05	0.83	0.05	0.89	0.04	0.88	0.03
	RW	0.89	0.04	0.93	0.03	0.91	0.04	0.91	0.03	0.94	0.03	0.94	0.02
3 Factor	Constant	0.64	0.25	0.95	0.22	0.84	0.21	0.74	0.21	0.70	0.18	0.62	0.13
	Rolling	0.86	0.14	0.91	0.11	0.91	0.11	0.91	0.10	0.92	0.09	0.89	0.08
	RC	0.77	0.06	0.91	0.06	0.93	0.06	0.88	0.05	0.86	0.06	0.94	0.02
	RW	0.92	0.03	0.96	0.02	0.95	0.03	0.95	0.02	0.96	0.02	0.95	0.02
5 Factor	Constant	0.64	0.26	0.92	0.18	0.85	0.22	0.83	0.28	0.72	0.18	0.66	0.42
	Rolling	0.86	0.12	0.91	0.09	0.91	0.09	0.92	0.09	0.92	0.08	0.88	0.07
	RC	0.78	0.05	0.91	0.05	0.92	0.05	0.90	0.06	0.73	0.04	0.85	0.07
	RW	0.93	0.03	0.97	0.02	0.96	0.02	0.95	0.02	0.96	0.02	0.94	0.02

Comparing the 4 models, we observe that from constant to rolling to RC to RW, the VR1 is in generally increasing and VR2 is decreasing, which suggests that our RW model has the strongest explanation power in terms of explained variance. This observation is consistent across CAPM, 3 factor and 5 factor models. Taking a deeper dive into the table, we see that the rolling method and RC has very similar VR1 values, better than constant model and worse than RW. This can be a result of the relatively large window size (5 years) of the rolling regression. A long rolling window implies that every period there will be very little variation and hence resembles the case of a fixed offset plus a stochastic noise term.

5.2 Fitting Errors

Next we plotted the cumulative squared error made by the model for portfolio 1,3 and 6 in Figures 14 to 16. Plots for portfolios 2 and 5 are attached in appendix. Looking at the cumulative errors we can see that the error made decreases from the constant model to RC to rolling regression to RW with rolling and RW having very similar errors. This is consistent with the fact that the coefficients from the two models have high correlation and track each other closely in Figure 11 to 13. The error analysis again suggesting that the Kalman Filter with RW state equations is the best model out of the four.

Figure 14: Cumulative Squared Errors, Portfolio 1

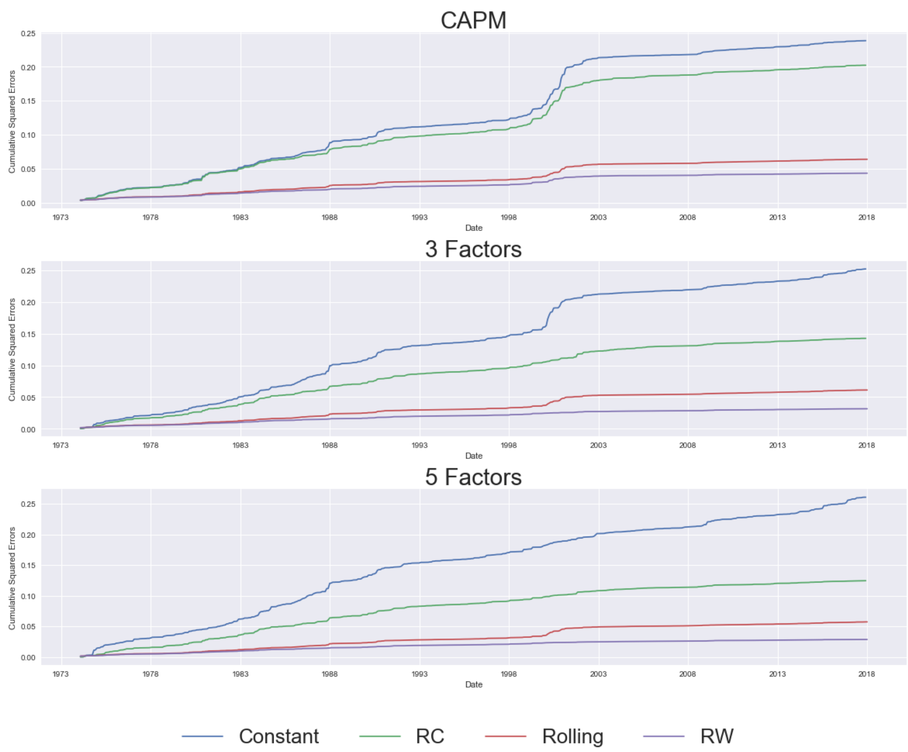


Figure 15: Cumulative Squared Errors, Portfolio 3

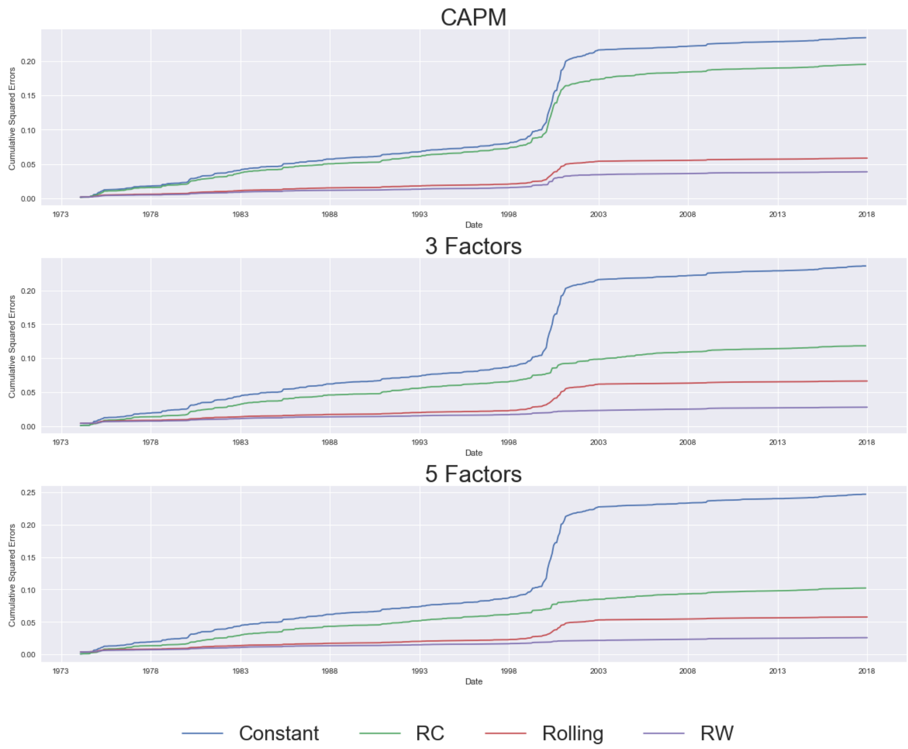
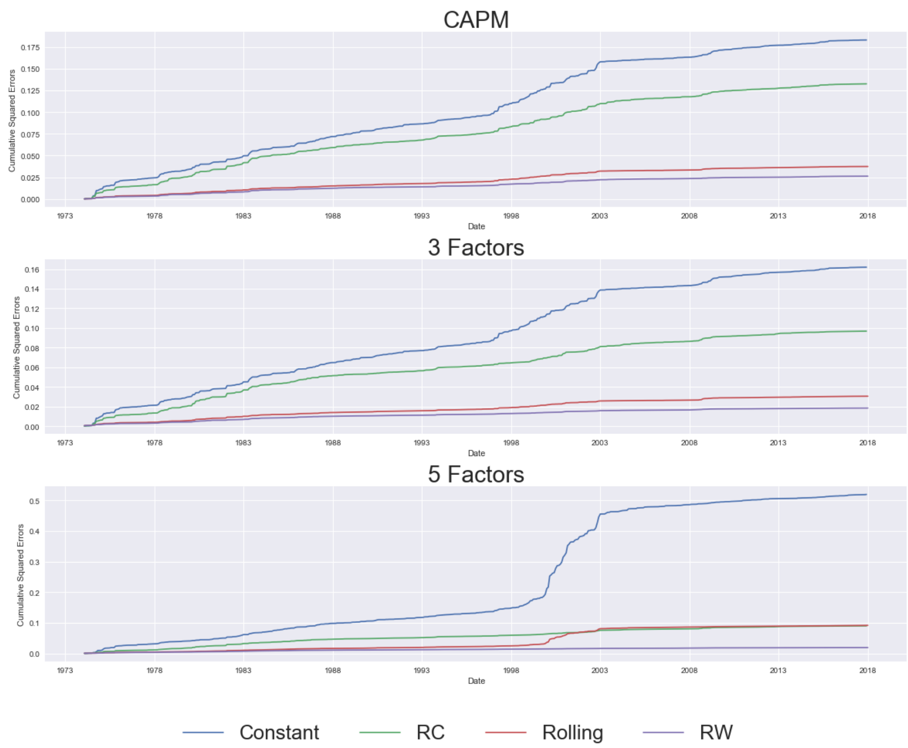


Figure 16: Cumulative Squared Errors, Portfolio 6



Another interesting observation is that there is a sharp increase in the errors during 1999 to 2000. This period corresponds to the ’internet bubble’, where a lot of technology stocks that did not exist before took the center of the stage in the financial markets. This lead to the creation of a new industry (which is proven to be the dominant industry in the market in the next few decades until now because currently most of the 10 biggest capitalization stocks in the S&P500 are technology stocks). Such structural shifts in the market lead to the potential creation of new factors and hence using the 5 factors defined before leads to a large increase in estimation errors. This reasoning is further confirmed if we compare the 5 factor models across portfolios 1,3 and 6. In portfolio 6 (which is the most traded and hence larger stocks), the 5 factor model inhibited such sharp increase in the cumulative errors because of these big technology stocks, however if we look at portfolio 1, the least traded stocks do not include these new joiners from the technology sector and hence the increase in error is relatively smooth.

5.3 Cross Sectional Analysis

Since we have observed such structural break in the market, we decided to split the time into two periods in our next analysis, from 1978 to 1998, and from 1998 to 2017. Within each of these two periods, we do a cross-sectional regression for a 5-factor model and correspondingly for 3-factor models and the CAPM of the following form:

$$\mathrm{R}_{i,t} = \alpha_{i,t} + \beta_{i,1,t}\mathrm{f}_{1,t} + \beta_{i,2,t}\mathrm{f}_{2,t} + \beta_{i,3,t}\mathrm{f}_{3,t} + \beta_{i,4,t}\mathrm{f}_{4,t} + \beta_{i,5,t}\mathrm{f}_{5,t} + \epsilon i, t \tag{5.1}$$

where $i = 1, ..., 200$ forms our asset universe of 200 stocks. One such cross-sectional regression is done at each time t and the result is a time series of implied factor returns and an intercept. If our model is good, the average factor returns should be non negative because inverters expect a risk premium for bearing the factor risk in their portfolio, and the alpha should be small. However we note here that the large α reject the validity of the model and efficient market hypothesis at the same time because of the joint hypothesis issue. The results are presented in Tables 24 for RW Kalman Filter, and the other methods are attached in appendix however they are similar. We can see that the median values of the factor exposures are in general positive and skewed to the right. However the joint hypothesis again comes into play because when the factor risk premium is negative, the reason can be that the model is inadequate or the factor had negative risk premium over that period.

Table 24: Cross-Sectional Regression, RW

		Before 1998			After 1998		
		25 %tile	median	75 %tile	25 %tile	median	75 %tile
CAPM	Intercept	-0.011	0.006	0.027	-0.011	0.008	0.027
	Mkt-Rf	-0.027	0.003	0.041	-0.028	0.001	0.034
3 Factor	Intercept	-0.011	0.009	0.023	-0.010	0.008	0.022
	Mkt-Rf	-0.030	0.004	0.036	-0.027	0.002	0.029
	SMB	-0.014	0.001	0.016	-0.021	0.001	0.025
	HML	-0.019	0.000	0.014	-0.016	0.007	0.021
5 Factor	Intercept	-0.011	0.010	0.023	-0.011	0.008	0.024
	Mkt-Rf	-0.029	0.008	0.038	-0.029	0.001	0.027
	SMB	-0.015	0.002	0.017	-0.021	0.001	0.025
	HML	-0.021	0.001	0.016	-0.015	0.006	0.019
	RMW	-0.008	-0.000	0.011	-0.022	-0.002	0.017
	CMA	-0.015	-0.000	0.010	-0.013	0.005	0.018

Portfolio Management Aspect Analysis

6 Portfolio Management Aspect Analysis

Factor models are changing the way asset managers construct their investment portfolios by improving their risk analysis. Since factor exposures have a direct impact on the volatility of a portfolio, the prediction power of the models can be tested by evaluating the performance of portfolios constructed using these coefficients.

Thus, to compare our time-varying models, we employ two classical methods of portfolio construction: minimizing variance and maximizing risk adjusted return. The idea is to calculate the portfolio weights by running appropriate optimizations (minimum variance optimization and maximum mean-variance optimization) using the covariance matrix estimation given by each of our four models as input (which implicitly incorporates our coefficient estimates). We then analyze the performance of the resulting portfolios and compare the criterion of construction between the different models. For example, for the minimum variance portfolio, we investigate the volatility of constructed portfolios in the next period, since the goals is to minimize variance, a good model should produce portfolios of smaller volatility.

For each model (CAPM, 3 Fama-French Factors and 5 Fama-French Factors model), we apply the 4 beta estimation methods (constant OLS, rolling OLS, Kalman-Filter Random Walk & Kalman-Filter with Random Coefficient) to calculate the covariance matrix of the asset returns that we use to compute our portfolio optimization. Moreover, to construct these portfolios we consider a universe of 200 stocks from the S&P500 index (that has 'survived') between 1988 and 2017 for reasons described earlier. The data we use is of monthly frequency.

6.1 Minimum Variance Portfolio

The Minimum Variance Portfolio consists of a portfolio of individually risky assets that, when viewed as a whole, results in the lowest expected level of risk measured as standard deviation of portfolio returns. Thus, we select our weights such that the variance of our portfolio is minimized as follows:

$$\begin{aligned} \underset{\boldsymbol{w}}{\text{minimize}} \quad & \boldsymbol{w}\boldsymbol{C}\boldsymbol{w}^T \\ \text{subject to} \quad & \sum_{i=1}^{N} w_i = 1 \end{aligned} \tag{6.1}$$

where N is the number of assets, $\boldsymbol{w} = (w_1, w_2, ..., w_N)$ is a vector of asset weights and $\boldsymbol{C}$ is the estimated covariance matrix of the assets' returns. We estimate the Covariance Matrix of Asset Returns by:

$$\boldsymbol{R} = \boldsymbol{\alpha} + \boldsymbol{\beta}\boldsymbol{F} + \boldsymbol{\epsilon}$$

$$Cov(\boldsymbol{R}) = \boldsymbol{\beta}\boldsymbol{\Omega}\boldsymbol{\beta}^T + \boldsymbol{D}$$

where

$$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, ...\alpha_N)^T$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_{1,1} & \beta_{1,2} & \dots & \beta_{1,k} \\ \beta_{2,1} & \beta_{2,2} & \dots & \beta_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{N,1} & \beta_{N,2} & \dots & \beta_{N,k} \end{bmatrix}$$

$$\boldsymbol{D} = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_N^2 \end{bmatrix}$$

with k being the number of factors and $\boldsymbol{D}$ being the covariance of the residuals.

The vectors of coefficients are given by our estimation methods (constant OLS, rolling OLS, Kalman-Filter RW and Kalman-Filter RC). We compute both the covariance matrix of the factor returns and the covariance matrix of the residuals (we assume the residuals are uncorrelated and thus take a diagonal form) using exponential weights with a factor of decay corresponding to a half-life of 6-months. The exponential weighting allows us to grant more importance to more recent observations. We compute the minimum variance portfolio using the four methods with CAPM, 3 Fama-French and 5 Fama-French Model. The results are presented below :

CAPM - Statistics of performance of the 4 portfolios & Market-RF, 1988 - 2017

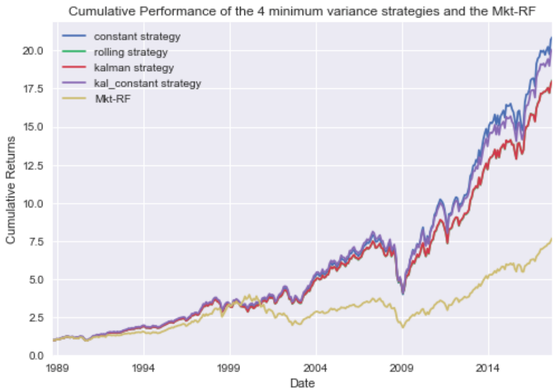


Figure 17: Cumulative Returns

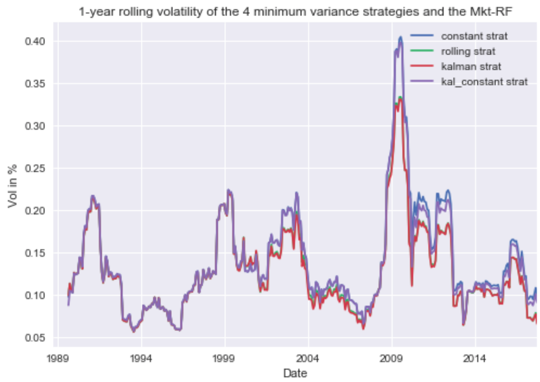


Figure 18: 1-year rolling volatility

Table 25: Summary table

	constant strategy	kal_constant strategy	kalman strategy	rolling strategy
Average return (p.a)	0.116306	0.114615	0.109118	0.109333
Volatility (p.a)	0.150611	0.148195	0.136120	0.137026
Sharpe Ratio	0.772231	0.773409	0.801628	0.797897

3 Fama-French Factors - Statistics of performance of the 4 portfolios & Market-RF, 1988 - 2017

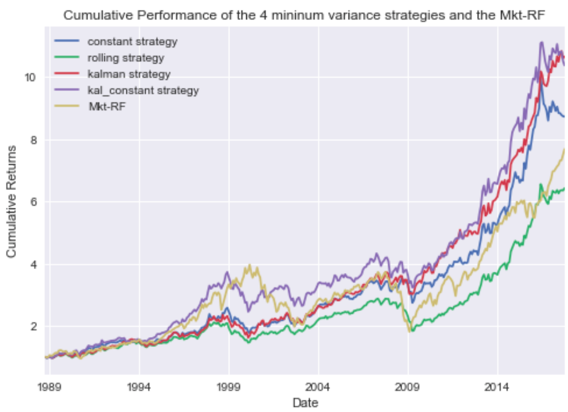


Figure 19: Cumulative Returns

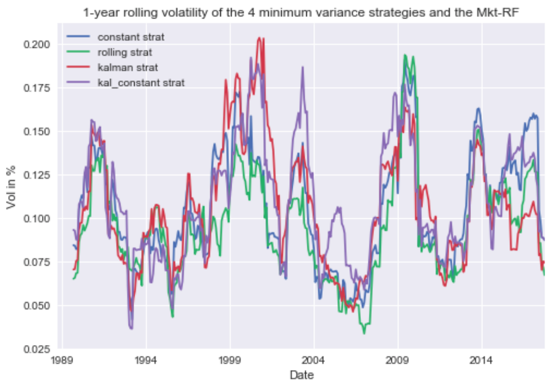


Figure 20: 1-year rolling volatility

Table 26: Summary table

	constant strategy	kal_constant strategy	kalman strategy	rolling strategy
Average return (p.a)	0.080604	0.087218	0.087611	0.069170
Volatility (p.a)	0.108881	0.114257	0.109450	0.100237
Sharpe Ratio	0.740297	0.763354	0.800462	0.690062

5 Fama-French Factors - Statistics of performance of the 4 portfolios & Market-RF, 1988 - 2017

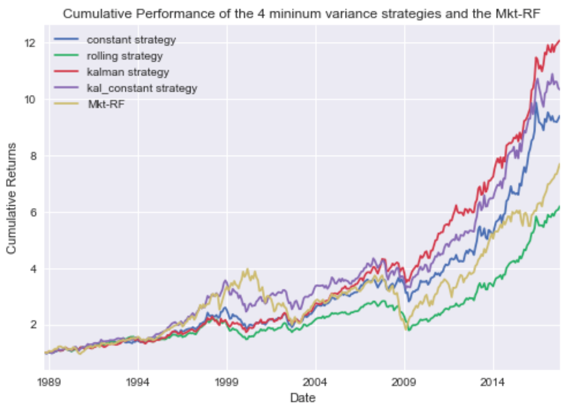


Figure 21: Cumulative Returns

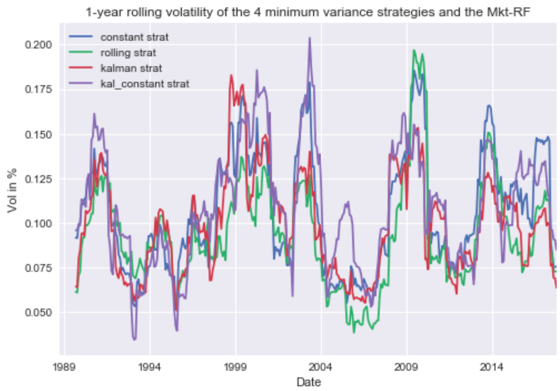


Figure 22: 1-year rolling volatility

Table 27: Summary table

	constant strategy	kal_constant strategy	kalman strategy	rolling strategy
Average return (p.a)	0.083119	0.086811	0.090982	0.067673
Volatility (p.a)	0.108960	0.112367	0.101044	0.098587
Sharpe Ratio	0.762841	0.772561	0.900418	0.686430

First of all, all constructed portfolios achieved very good performances and in general out-performed the market.

While studying the results we should keep in mind that our main interest here lies with the realized volatility of the portfolios we constructed instead of the returns. Indeed, a portfolio with a low risk-level would confirm the robustness of our covariance matrix estimation. For the 3 models, the 4 methods seems to produce relatively low risk-level portfolios given we are using an asset universe that consists of purely equities. Taking a closer look at the 1-year rolling volatility graphs and the summary tables, we can observe that the lowest risk portfolios is produced by the Kalman-Filter RW and Rolling-OLS methods. For the CAPM model, we can observe that the Kalman-Filter RW estimation produces the portfolio with the lowest volatility (13.6%) and beats the Rolling-OLS method (13.7%). Nevertheless, for the 3 FF-factors and 5 FF-factors models, even if the volatilities remain very close, the Rolling-OLS beats the Kalman-Filter. It seems that the Rolling-OLS was better at incorporating more factors . This phenomena can be explained by the diagonal transition matrix that we assumed for the Kalman model. Indeed, we do not allow for correlation between the coefficients whereas the Rolling-OLS grants us more flexibility. The coefficients obviously exhibit some correlations (for example SMB and HML) that the diagonal transition matrix fails to capture. In conclusion, Kalman-Filter RW and Rolling-OLS still are tested to be very good method of estimation and present good prediction power.

We also printed the average returns and Sharpe ratios, but they are less relevant for our purpose since the goal of the model was to minimize variance.

6.2 Mean Variance Portfolio with Momentum Signals

The Mean-Variance Portfolio is constructed by maximizing the expected returns of the portfolios while penalizing for risk. Thus, we solve the following optimization problem :

$$\begin{aligned} \underset{\boldsymbol{w}}{\text{maximize}} \quad & \boldsymbol{w}\mathbb{E}(\boldsymbol{R}) - \gamma \boldsymbol{w}\boldsymbol{C}\boldsymbol{w}^T \\ \text{subject to} \quad & \sum_{i=1}^{N} w_i = 1, \qquad w_i \geqslant 0 \quad \forall i \end{aligned} \tag{6.2}$$

with γ being the parameter of risk aversion.

We use the same covariance matrix estimations as previous part. In addition to that, we need a new input : the expected returns as a signal. We decide not to use our factor models to compute this signal as they tend to have very poor prediction power. Instead, we decide to use a simple momentum signal (keeping in mind that the goal here is to test the performance of the covariance matrix estimation using our beta estimates instead of computing the best strategy). To compute the momentum signal at time t, we use the average returns between month t-13 and month t-2. The reason for taking the last 13 months less the most recent month is that it is well known in momentum strategies that the short term reversal behavior undermines the momentum strategies' performance. Hence we leave out the most recent month in constructing our momentum signals.

$$\mathbb{E}_t(\widehat{\boldsymbol{R}}) = \frac{\sum_{k=2}^{13} R_{i,t-k}}{12} \tag{6.3}$$

with $R_{i,t}$ being the return of asset i at time t.

We compute for the four methods for the CAPM, 3 Fama-French and 5 Fama-French Model. The results are presented below :

CAPM - Statistics of performance of the 4 portfolios & Market-RF, 1988 - 2017

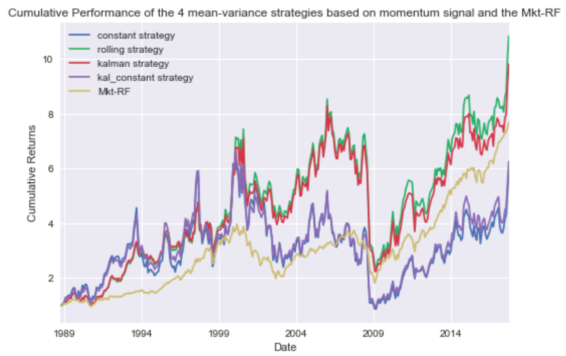


Figure 23: Cumulative Returns

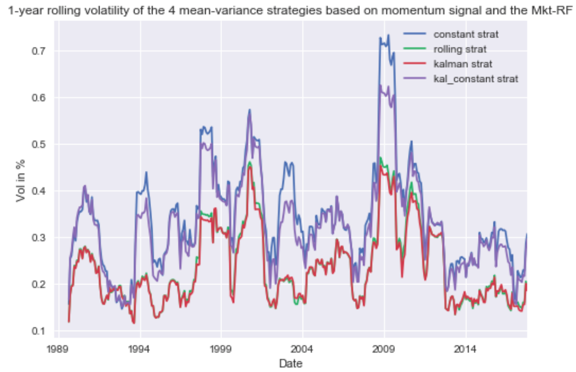


Figure 24: 1-year rolling volatility

Figure 25: Summary table

	constant strategy	kal_constant strategy	kalman strategy	rolling strategy
Average return (p.a)	0.127375	0.120843	0.109159	0.113562
Volatility (p.a)	0.355118	0.333991	0.243272	0.246976
Sharpe Ratio	0.358682	0.361815	0.448711	0.459811

3 Fama-French Factors - Statistics of performance of the 4 portfolios & Market-RF, 1988 - 2017

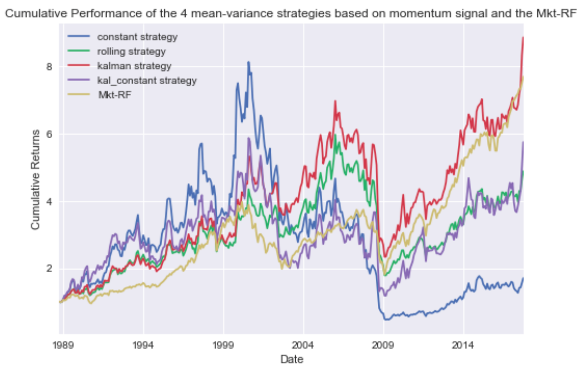


Figure 26: Cumulative Returns

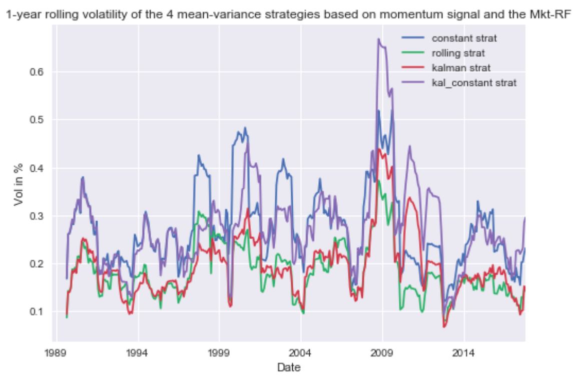


Figure 27: 1-year rolling volatility

Figure 28: Cumulative Squared Errors, Portfolio 1

	constant strategy	kal_constant strategy	kalman strategy	rolling strategy
Average return (p.a)	0.061315	0.103791	0.096045	0.072875
Volatility (p.a)	0.289932	0.289678	0.202450	0.190308
Sharpe Ratio	0.211480	0.358298	0.474414	0.382935

5 Fama-French Factors - Statistics of performance of the 4 portfolios & Market-RF, 1988 - 2017

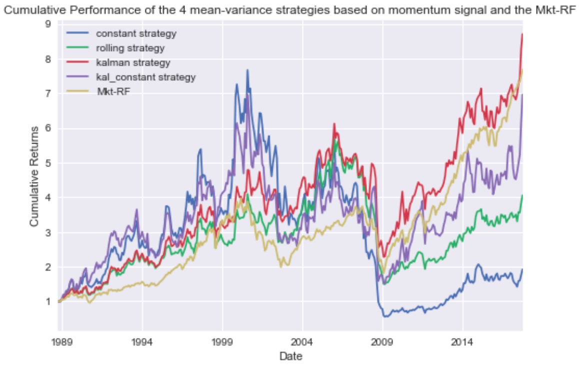


Figure 29: Cumulative Returns

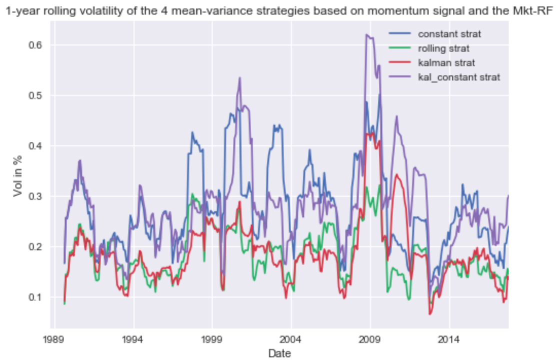


Figure 30: 1-year rolling volatility

Figure 31: Cumulative Squared Errors, Portfolio 1

	constant strategy	kal_constant strategy	kalman strategy	rolling strategy
Average return (p.a)	0.065122	0.109994	0.094029	0.066393
Volatility (p.a)	0.289898	0.289413	0.195236	0.190067
Sharpe Ratio	0.224639	0.380060	0.481619	0.349314

First, we notice that here the portfolios returns are much more noisy. This is due to the fact that momentum signals tends to trade frequently.

One interesting point is that once again we can observe the previous pattern. It seems indeed that we can distinguish two groups : the two constant models on one hand and the KF RW and Rolling-OLS on the other hand which produce the portfolios with the lowest realized volatility for each of the 3 considered factor models. In addition the KF RW beats the Rolling-OLS for the CAPM model (24.3% vs 24.7%) whereas the Rolling-OLS beats the K-F for the 3 FF-model and 5 FF-model which seems to confirm previous hypothesis that rolling regression is better when more factors are included. Once again, it seems that Kalman-Filter and Rolling-OLS present good prediction power.

Conclusion

7 Conclusion

The four models (constant coefficient, rolling regression, Kalman Filter with random walk state equation and Kalman Filter with random coefficient state equation) and three methods (CAPM, 3 factor, 5 factor) are tested thoroughly in our paper. To our knowledge this is the first time that 5 factor models has been tested using the Kalman Filter and the results are promising. Fist of all all methods show consistency in the basic model statistics, and coincides with our economic intuition about the market. In addition, in terms of explained variance and the errors made, the Kalman Filter with random walk was the best model in the four, closely followed by the rolling regression model. And under a portfolio management perspective, the RW Kalman Filter has also proven to be in general the model that produced the best portfolio performances and seems to present very good prediction power. Finally, one important aspect that we discovered during our analysis is that it could be interesting to customize the the Kalman Filter's state equations in a way which the coefficients are inter-dependent. We have observed evidence for this dependency in our analysis and this would be a fruitful direction in terms of future research.

Appendix

8 Appendix

Figure 32: Cumulative Squared Errors, Portfolio 2

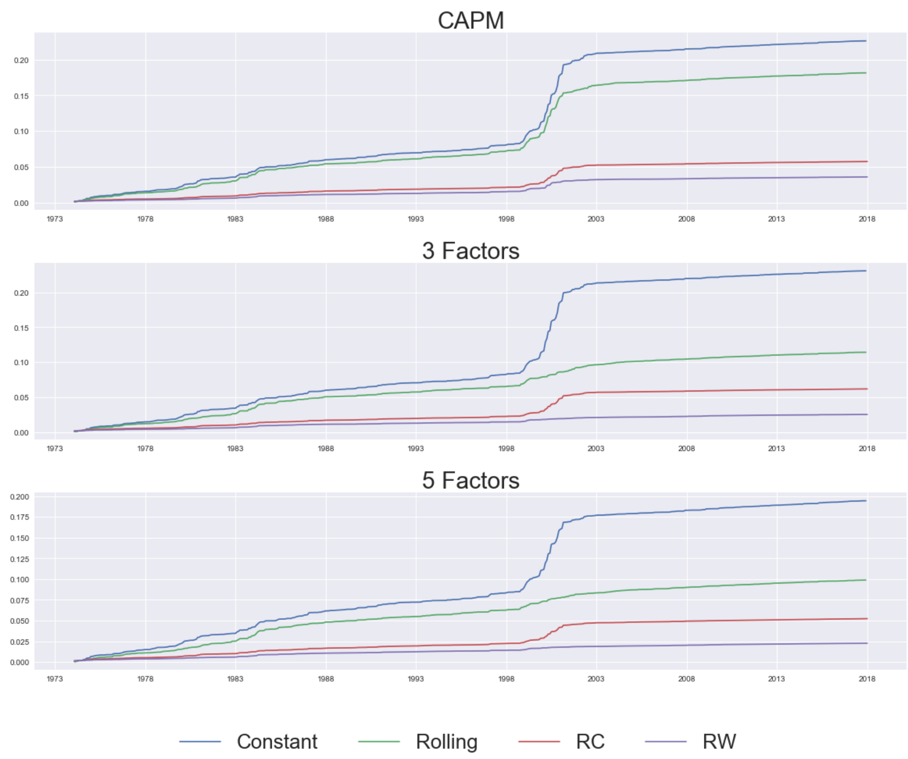


Table 28: Cross-Sectional Regression, Constant

		Before 1998			After 1998		
		25 %tile	**median**	**75 %tile**	**25 %tile**	**median**	**75 %tile**
CAPM	**Intercept**	-0.021	0.004	0.024	-0.014	0.007	0.030
	Mkt-Rf	-0.026	0.010	0.037	-0.027	0.003	0.031
3 Factor	**Intercept**	-0.015	0.006	0.023	-0.013	0.008	0.035
	Mkt-Rf	-0.020	0.009	0.031	-0.024	0.001	0.025
	SMB	-0.009	0.000	0.014	-0.012	0.001	0.014
	HML	-0.017	-0.003	0.012	-0.016	-0.001	0.013
5 Factor	**Intercept**	-0.011	0.005	0.024	-0.016	0.009	0.037
	Mkt-Rf	-0.017	0.006	0.029	-0.024	0.000	0.024
	SMB	-0.007	0.002	0.012	-0.013	0.000	0.014
	HML	-0.017	-0.003	0.012	-0.016	-0.001	0.013
	RMW	-0.005	0.001	0.006	-0.007	-0.000	0.006
	CMA	-0.010	-0.002	0.007	-0.012	-0.000	0.009

Table 29: Cross-Sectional Regression, Rolling

click to scroll output; double click to hide

		Before 1998			After 1998		
		25 %tile	**median**	**75 %tile**	**25 %tile**	**median**	**75 %tile**
CAPM	**Intercept**	-0.015	0.007	0.023	-0.013	0.009	0.029
	Mkt-Rf	-0.027	0.001	0.037	-0.024	-0.001	0.033
3 Factor	**Intercept**	-0.007	0.005	0.020	-0.007	0.007	0.022
	Mkt-Rf	-0.026	0.005	0.036	-0.025	0.001	0.027
	SMB	-0.018	0.000	0.018	-0.017	0.003	0.024
	HML	-0.013	-0.001	0.015	-0.015	0.002	0.019
5 Factor	**Intercept**	-0.008	0.005	0.023	-0.004	0.008	0.022
	Mkt-Rf	-0.024	0.007	0.037	-0.026	0.000	0.028
	SMB	-0.016	-0.000	0.017	-0.017	0.001	0.024
	HML	-0.013	-0.000	0.017	-0.016	0.003	0.018
	RMW	-0.008	-0.001	0.010	-0.016	-0.002	0.013
	CMA	-0.012	-0.001	0.012	-0.013	-0.001	0.015

Figure 33: Cumulative Squared Errors, Portfolio 5

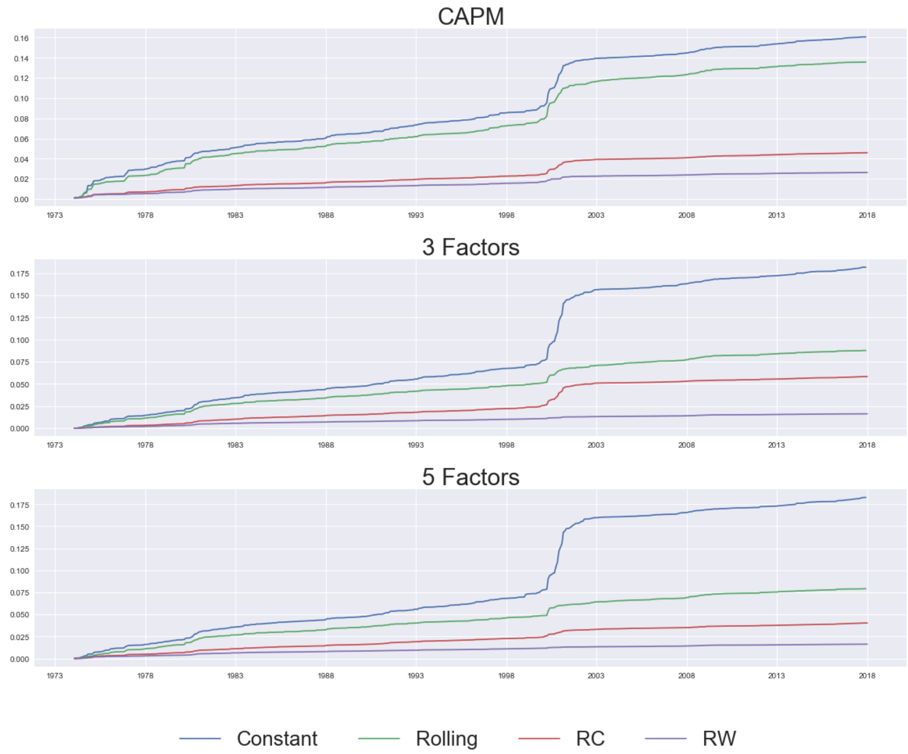


Table 30: Cross-Sectional Regression, RC

		Before 1998			After 1998		
		25 %tile	**median**	**75 %tile**	**25 %tile**	**median**	**75 %tile**
CAPM	**Intercept**	-0.008	0.007	0.023	-0.013	0.007	0.030
	Mkt-Rf	-0.019	0.006	0.026	-0.022	0.004	0.030
3 Factor	**Intercept**	-0.006	0.007	0.027	-0.015	0.009	0.028
	Mkt-Rf	-0.018	0.003	0.021	-0.023	0.002	0.027
	SMB	-0.009	0.000	0.013	-0.011	0.000	0.014
	HML	-0.017	-0.001	0.010	-0.014	-0.001	0.015
5 Factor	**Intercept**	-0.005	0.008	0.025	-0.017	0.008	0.028
	Mkt-Rf	-0.017	0.005	0.021	-0.020	0.003	0.027
	SMB	-0.009	0.001	0.013	-0.011	0.001	0.014
	HML	-0.016	-0.001	0.011	-0.015	-0.002	0.014
	RMW	-0.004	0.003	0.009	-0.009	0.001	0.009
	CMA	-0.011	-0.001	0.008	-0.010	-0.002	0.009

Bibliography

[1] Ferson, W.E. and Harvey, C.R., 1999. Conditioning variables and the cross section of stock returns. The Journal of Finance, 54(4), pp.1325-1360.

[2] Fama, E.F. and French, K.R., 1997. Industry costs of equity. Journal of financial economics, 43(2), pp.153-193.

[3] Diebold, F.X. and Rudebusch, G.D., 1994. Measuring business cycles: A modern perspective (No. w4643). National Bureau of Economic Research.

[4] Tong, H. and Lim, K.S., 2009. Threshold autoregression, limit cycles and cyclical data. In Exploration Of A Nonlinear World: An Appreciation of Howell Tong's Contributions to Statistics (pp. 9-56).

[5] Lettau, M. and Van Nieuwerburgh, S., 2007. Reconciling the return predictability evidence: The review of financial studies: Reconciling the return predictability evidence. The Review of Financial Studies, 21(4), pp.1607-1652.

[6] Adrian, T. and Franzoni, F., 2009. Learning about beta: Time-varying factor loadings, expected returns, and the conditional CAPM. Journal of Empirical Finance, 16(4), pp.537-556.

[7] Huang, P., 2006. Applications of Time-varying-parameter Models to Economics and Finance.

[8] Harvey, C.R., Solnik, B. and Zhou, G., 2002. What determines expected international asset returns?. Annals of Economics and Finance, 3(2), pp.249-98.

[9] Nieto, B., Orbe, S. and Zarraga, A., 2014. Time-varying market beta: does the estimation methodology matter?. SORT-Statistics and Operations Research Transactions, 38(1), pp.13-42.

[10] Ang, A. and Kristensen, D. (2012). Testing conditional factor models. Journal of Financial Economics, 106, 132–156.

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

[12] Jagannathan, R. and Wang, Z. (1996). The conditional CAPM and cross-section of expected returns. Journal of Finance, 51, 3–53.

[13] Lewellen, J. and Nagel, S. (2006). The conditional CAPM does not explain asset-pricing anomalies. Journal of Financial Economics, 82, 289–314.

[14] Li, Y. and Yang, L. (2011). Testing conditional factor models: a nonparametric approach. Journal of Empirical Finance, 18, 975–992.