Erasmus University Rotterdam

Bachelor Thesis Double Degree programme Econometrics & Economics

The higher order CAPM and Fama & French relationship
A dynamic conditional approach

Author
B.R. Barendrecht
384076

Supervisor
K. Scholtus

Second assessor
J.A Oorschot

Abstract

Ever since the introduction of the Fama and French (1992) factors, the factors have replaced CAPM as the benchmark of asset pricing models. However compared to CAPM it lacked a complete mathematical financial theory. We tried to solve this problem by developing a new theoretical model including higher order co-moments and relating these new risk factors to the Fama & French factors. Using a time-varying estimation method borrowed from Engle (2016) we first reaffirm that dynamic conditional β estimation improves the performance of the Fama & French Factors. When applying the same method to higher order co-moments, we discovered mixed results and care should be taken when estimating the model. We found the Fama & French factors to be priced, as well as the co-skewness. However there is insufficient evidence for the co-kurtosis. Next we found that the Fama & French factors and higher order moments do share a common risk factor as well as a factor which is distinct for both, obtaining this result both between assets and through time.

July 8, 2018

Contents

1 Introduction

In the world of asset management one question might be the most important, is the risk of an asset sufficiently rewarded? To answer this question asset pricing models are usually used. Perhaps the most well known asset pricing model is the capital asset pricing model (CAPM).

The central idea behind the CAPM model is that not all risk is priced. As investors are able to diversify, only the risk common in the market is priced. The standard model requires various assumptions. We will mainly focus on two assumptions, with either of which the CAPM model can be derived: 1) Investors only care about the first two moments, that is average return and volatility of the return. Or 2) the distribution of the returns follow a normal distribution, as this collapses to the first since the normal distribution is completely described by mean and variance alone. However these assumptions are not met in practice, for example the daily returns of stocks show excess kurtosis and negative skewness. Moreover the assumption that investors are not influenced by extreme events is almost impossible to defend.

Multiple extensions to the CAPM have been discussed in previous literature. One of the most well known example is the Fama & French model, introduced in Fama and French (1992). Fama and French (1992) show that not only the market risk is priced, but also a size and book-to-market risk factor. They construct 6 value-weight portfolios by sorting the assets on size and book-to-market ratio. The return of the SMB is the return of the smallest half minus the return of the largest. The HML return is the return of the value stock minus the value of the growth stocks. They than add the additional factors of the portfolio returns to the CAPM model.

A second way to extend the CAPM is addressing the central assumptions of CAPM more directly. That is we do not assume normality or a quadratic utility function. As a result, higher order moments (skewness, kurtosis, ...) enter into the consideration of investors. As a result not only the covariance of an asset and the market matters, but also the co-skewness and co-kurtosis and possibly even higher co-moments. The first determines the common risk that both the market and the asset deliver a negative return simultaneously while the second determines the common risk that both the market and the asset have a extreme return. An example of this is Harvey and Siddique (2000), in this paper they construct a co-skewness measure and show that this risk-factor is indeed priced.

Furthermore in Harvey and Siddique (2000) a few remarks are made between the fama-french model and the higher order co-moments. They show that the variance explained in the returns of assets go up with a similar magnitude when including the Fama-French Factors, co-skewness or both. They also found a cross-sectional relationship between co-skewness and the Fama-French Factor. They showed that when using sorted book-to-market portfolios, the highest excess return portfolio had the lowest co-skewness. This leads the authors to hypothesize that the Fama-French factors may proxy co-skewness. This in turn leads us to hypothesize that this proxy relationship even extends to higher order moments such as kurtosis.

However Hung et al. (2004) found more limited evidence for the prices of higher order moments when compared to Fama & French. Hung et al. (2004) show that higher order co-moments provide little to no benefit to the Fama & French model. They remark that to estimate higher order co-moments correctly, one needs a lot more data than the purely linear relationship of Fama & French. Hung et al. (2004) final thought is that even as they provided some evidence for regarding the higher order co-moments as risk factors, their lack of data leads to recommendation to use daily data instead of monthly data.

Often when estimating asset pricing models an assumption of constant β is made, either trough estimating β using the entire sample period, as for example in the original Fama & French study (Fama and French (1992)) or using rolling window estimation. Engle (2016) remarks that both these methods lead to possible inference problems. The first might be a too restrictive assumption, as over a long period of time companies or even entire industries might completely change. The rolling window assumes that the true β has an equal distribution as the past x days, however when assuming a non-constant β, this assumption is hard to defend.

To summarize we will answer the following questions, *does time-varying estimation improve the estimation of both the Fama-French model and higher moment CAPM models? And are these factors priced?* Finally we will answer: *Are the Fama-French model factors related to these higher moment risk factors and is this relationship not only cross-sectional but as time progresses as well?*

To clarify the last question we will use an example. Engle (2016) remarks that the money industry behaves as a growth stock in 2004-2007 and than reverses to a value stock during a financial crisis. We hypothesize that this shift in the high-minus-low Fama & French factor also correspond with a simultaneous shift in the higher order co-moments risk factors. This would add to the observation in Harvey and Siddique (2000), where this was argued only cross-sectionally within a time period.

First we discuss the theoretical foundation for asset pricing models, while in essence replicating Fama and MacBeth (1973), Harvey and Siddique (2000) and others. We derive that the main theoratical results of Fama and MacBeth (1973) and Harvey and Siddique (2000) are in fact special cases of a more general model. The general model being a model which incorporates all higher moments (HOCAPM).

We also discuss different methods to estimate time-varying effects. Of the various methods considered we choose the DCB method described in Engle (2016) for two main reasons. It obtains a distinct beta for every period, which we need to relate HOCAPM and Fama & French. Secondly the equation to estimate DCB is a conditional version of the HOCAPM estimator.

Next we perform empirical test to test the theory developed against the benchmark of the Fama-French three factor model. When comparing the performance of HOCAPM against Fama & French the HOCAPM model fits the data less well. We reaffirm the results from Fama and French (1992), that indeed size and book-to-market factors are risk factors and furthermore are priced. We reaffirm Harvey and Siddique (2000) that the co-skewness is priced. However, we obtain limited results when pricing co-kurtosis. We also show that even though the relation between HOCAPM and Fama-French risk factors is weak, it seems that the Fama-french risk factors at least partly capture the risk of higher moments. However we cannot exclude that these factors also capture some other type of risk. To conclude, we obtain the insight that if the risk factors are related, the factors are related not only between assets but also through time for a particular asset.

Thus we add to the literature in multiple ways. We first recast CAPM as a general model by including all moments, which at least to our knowledge, has not been done in this way before. Within financial literature it is unclear whether including higher order moments are sufficient to price assets. Furthermore it has been suggested that Fama & French factors proxy higher order co-moments. Our research provides evidence towards to first, however the evidence for the second statement remains partial.

2 Theoretical background higher order CAPM

To include higher order moments in a theoretical framework we essentially follow in the footsteps of Fama and MacBeth (1973), Harvey and Siddique (2000) and Dittmar (2002). However our approach differs in one key aspect we derive the results for all moments simultaneously. That is we start by applying a infinite Taylor expansion on the utility function and follow the appropriate steps to arrive at the extended CAPM which includes all moments.

2.1 Preliminaries

To ease the derivations of the higher order CAPM model, we first state the general rule for calculating moments:

$$\begin{aligned} H_P &= \sum_{i \in I} \lambda_i H_i \\ H_{PP} &= \sum_{i \in I} \sum_{j \in I} \lambda_i \lambda_j H_{i,j} \\ H_{PPP} &= \sum_{i \in I} \sum_{j \in I} \sum_{k \in I} \lambda_i \lambda_j \lambda_k H_{i,j,k} \\ &\dots \end{aligned} \tag{1}$$

where I is the set of all possible assets and/or portfolios. $H_{i,j}$ denotes the second order co-moment between asset i & j and $H_{i,j,k}$ the third order co-moment between these assets. H_P denotes the first moment an complete portfolio.
If we allow for a risk-less asset the above equation still holds. However since $H_{f,f,...}$ is zero for any higher order moment (riskless) and $H_{f,k,...}$ is zero, using that any co-moment between a constant and a random variable is zero we obtain:

$$\begin{aligned} H_R &= \alpha H_f + (1-\alpha) H_P \\ H_{RR} &= (1-\alpha)^2 H_{PP} \\ H_{RRR} &= (1-\alpha)^3 H_{PPP} \\ &\dots \end{aligned} \tag{2}$$

Two things to note: 1) if we drop all moments with order higher than two we simply obtain the results of Sharpe (1964). 2) If we define a risk measure (volatility, skewness, ...) as the appropriate root of these moments (the second root for the second moment and so on) the equations becomes a linear combination. As a result the investor can thus choose any point between the risk less rate and any possible portfolio.

2.2 Pricing with higher order moments

The second step is specifying the utility function U[1]. If we assume the utility function is linear in the moments (and using a Taylor expension of $U = g(R)$ around the risk free rate and taking the expectation gives this result) we obtain:

$$U = \theta_1 H_R + \theta_2 H_{RR} + ... \tag{3}$$

rewriting using equation 1 & 2 and using that the first moment of a constant around the same constant is zero:

$$U = \theta_1(1-\alpha)(\sum_i \lambda_i H_i) + (1-\alpha)^2\theta_2(\sum_i \sum_j \lambda_i \lambda_j H_{i,j}) + ... \tag{4}$$

If we follow the regular CAPM assumption that we can split the problem in first selecting the optimal portfolio and than selection the amount of risk-free rate investment we can rewrite this equation and frame it as an optimization problem with respect to the amount of each specific asset (λ):

$$max \quad U = \theta_1(\sum_i \lambda_i H_i) + \theta_2(\sum_i \sum_j \lambda_i \lambda_j H_{i,j}) + ... \tag{5}$$

The F.O.C. of this problem is given by:

$$\begin{aligned} \frac{\partial U}{\partial \lambda_i} = \theta_1 H_i + 2\theta_2 \lambda_i H_{i,i} + \theta_2 * \sum_{j \neq i} \lambda_j H_{i,j} + \\ 3\theta_3 \lambda_i^2 H_{i,i,i} + 2\theta_3 \lambda_i \sum_{j\neq i} H_{i,i,j} + \theta_3 \sum_{j\neq i} \sum_{k \neq i} H_{i,j,k} + ... \end{aligned} \tag{6}$$

which can be rewritten as:

$$\begin{aligned} \frac{\partial U}{\partial \lambda_i} = \theta_1 H_i + 2\theta_2 \lambda_i H_{i,i} + \theta_2 * H_{i,P} + \\ 3\theta_3 \lambda_i^2 H_{i,i,i} + 2\theta_3 \lambda_i H_{i,i,P} + \theta_3 H_{i,P,P} + ... \end{aligned} \tag{7}$$

Simply put, for an portfolio to be optimal a investor needs to be rewarded (H_i) for both idiosyncratic ($H_{i,i}, H_{i,i,i}, H_{i,i,P}$) and systematic risk ($H_{i,P}, H_{i,P,P}$). However for a perfect divisible and diversified portfolio it holds that:

$$\lim_{\lambda_i \to 0} (\frac{\partial U}{\partial \lambda_i}) = \theta_1 H_i + \theta_2 * H_{i,P} + \theta_3 H_{i,P,P} + ... = 0 \tag{8}$$

Or differently:

$$H_i = -\frac{\theta_2}{\theta_1} * H_{i,P} - \frac{\theta_3}{\theta_1} H_{i,P,P} - ... \tag{9}$$

And hence if an investor is diversified enough, the only risk premium he demands is the price of the systemic risk associated with his current held portfolio.

[1] The author is known with the commonly discussed problems with the utility function. An example of this can be found in Dittmar (2002). However if we take a step back and assume a proper choice of θ describes a useful utility function for a descriptive model, we sidestep this problem. Dittmar (2002) solves these problems by directly specifying and assuming a pricing kernel.

We can compute the total expected compensation for held risk:

$$H_P = \sum_{i \in I} H_i = -\frac{\theta_2}{\theta_1} * \sum_{i \in I} \lambda_i H_{i,P} - \frac{\theta_3}{\theta_1} \sum_{i \in I} \lambda_i H_{i,P,P} - ... \tag{10}$$

Using the insights obtained from Fama and MacBeth (1973), we can calculate H_i as a fraction:

$$\begin{aligned} H_i &= \frac{H_{i,P}}{H_{P,P}} H_{P,P} Pr_2 + \frac{H_{i,P,P}}{H_{P,P,P}} * H_{P,P,P} Pr_3 + ... \\ Pr_z &= \frac{\partial_z H_P}{\partial H_{P^z}} \\ Pr_z &= \frac{\partial_z H_P}{H_{P^z}}. \end{aligned} \tag{11}$$

Where Pr_z can be interpreted as the required price or compensation to hold risk. With ∂_i is meant the expected return increase when holding all risk factors expect i constant, this for ease of notation. H_{i^z} is a shorthand for the z'th moment of the asset or portfolio i. The third equation stems from linearity of (9) Now define $\beta_{i,z}$, the β of asset i relative the z'th order co-moment with the market as:

$$\beta_z = \frac{H_{i,P^z}}{H_{P^z}} \tag{12}$$

The intuition behind the factor β can be explained as follows: It is the amount of systematic risk added by the asset divided by the total amount of risk within a portfolio. Now we can rewrite equation (11) as:

$$H_i = \beta_{i,2}(\partial_2 H_P) + \beta_{i,3}(\partial_3 H_P) + ... \tag{13}$$

Or differently, the expected return of an asset is the sum of all systemic risk within an asset divided by the corresponding total systematic risk times the demanded compensation of the systematic risk. To conclude we have shown that the most principle logic behind CAPM can be extended to Higher Order moments (HOCAPM). If we treat CAPM as a special case of HOCAPM by setting $\theta_i = 0, \forall i > 3$ equation (12) & (13) simply become the β's and pricing equation as specified in Fama and MacBeth (1973) and others.

2.2.1 A note about estimation

To estimate the pricing and risk factors care needs to be taken. In (11) it is silently assumed that the risk factor only has a direct linear effect on the expectation on the portfolio. However this is in general not true, as $(R_p - R_f)$ and $(R_p - R_f)^2$ are functionally related. As a result (11) should be interpreted when keeping all other factors constant. This is effectively what multiple regression does. Thus to estimate the β_i we should estimate the following linear model:

$$(R_i - R_f) = \beta_{i,1}(R_P - R_f) + \beta_2 * (R_P - R_f)^2 + ... \tag{14}$$

with:

$$\hat{\beta} = (\hat{H}_{(P_1,P_2,...)' \otimes (P_1,P_2,...)})^{-1} \hat{H}_{(P_1,P_2,...)' \otimes (i)} \tag{15}$$

also known as the OLS estimator. Here $\otimes$ means the outer product such that $H_{(i,j) \otimes (k,l)}$ means the 2x2 matrix with entries H_{ik}, H_{il}, ...

2.3 Market portfolio

The last step of CAPM is the idea of the efficient frontier. The efficient frontier is ussually defined as given a specified level of risk, the maximum obtainable return. If we modify equation (1) to obtain a optimization problem:

$$\begin{aligned}(H_p, H_{pp}, ...) &\in EF \Leftrightarrow \\ H_p &= max(H_p) \quad s.t. \\ H_P &= \sum_{iinI} \lambda_i H_i \\ H_{PP} &= \sum_{iinI} \sum_{jinI} \lambda_i \lambda_j H_{i,j}^2 \\ H_{PPP} &= \sum_{iinI} \sum_{jinI} \sum_{kinI} \lambda_i \lambda_j \lambda_k H_{i,j,k}^3 \\ &... \end{aligned} \tag{16}$$

Where EF is the set of all portfolios on the efficient frontier. Combined with the risk free investment this looks graphically in 2 dimensions like the well known figure 1. As a result, the only actually efficient portfolio is the portfolio on the intersection of the tangent line and the efficient frontier. As everybody should hold the same portfolio this coincides with the market portfolio, individual preferences simply determine where on this tangent line the individual portfolios lie. However when generalizing to other risk measures, this point which lies on the intersection becomes a $N-2$ dimensional hyperplane.[2] Next this hyperplane is in general non-linear. This leads to some conclusions contrary to CAPM: 1) not everybody has to hold the same portfolio as this hyperplane is not a single point, 2) people trade as their optimal portfolio on this intersection can change not only the price of asset, 3) When estimating beta, one has to use his current held portfolio, as long as it is efficient.

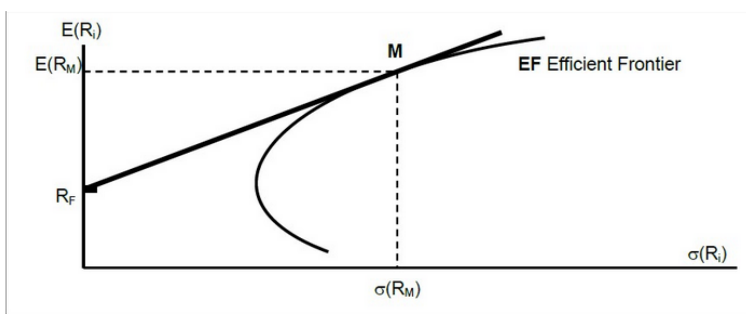


Figure 1: Original CAPM, cropped from Ide (2013).

While 1) & 2) are closer to real world observations, that is there is a massive amount of trade and different investors hold different portfolios, 3) provides an interesting problem. We have no guarantee that the market portfolio even lies on the efficient frontier. As a graphical illustration see figure 2. This represents the 2 dimensional cross section of a random EF. Half of the world holds portfolio A, the other half portfolio B. Thus the market portfolio lies in between these points, not on the EF. However in general with CAPM or even Fama-French factors, one uses an (weighted) portfolio of listed stocks as a proxy for the market portfolio itself. For scientific purposes we can simply replace this by suggesting this portfolio proxies an efficient portfolio instead.

[2]Note that EF frontier and the tangent plane can be described by a function $H_R = f(H_{RR}, H_{RRR}, ...)$, thus N-1 dimensional. The resulting intersection is thus bound by two constraints which means their intersection is an $N-2$ dimensional object. As a result the original CAPM has only one efficient portfolio. A single point.

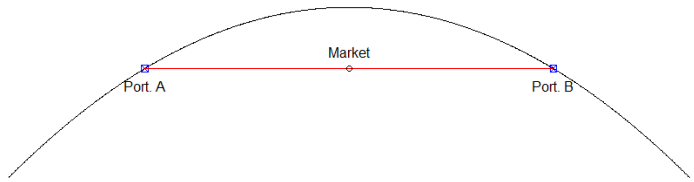


Figure 2: HOCAPM, 2D cross section of EF with Market portfolio included

3 Selecting a time-varying estimation method

When estimating the beta of CAPM and the Fama-French factors one usually assumes a constant beta, however this is generally not necessary. If we assume no transaction cost, which the original CAPM does and we do not deviate from, an investor can update his portfolio every period. Since including other risk factors implies there is no single optimal portfolio for everyone an investor should be interested in changing circumstances to construct his optimal portfolio. However there are a variety of models and estimation methods available to estimate a time-varying β.

The simplest way to accomplish this is a rolling window estimation. That is, we estimate the current beta using the data from the past 60 days for example. Engle (2016) briefly discusses this option and dismisses this. As one implicitly assumes that the beta is as likely to be the same as one of the days within the rolling window. This of course does not have to hold in general.

A second way is to assume structural change as in Quandt (1958). For the CAPM estimation problem this simply means that we assume that the beta's have two or more distinct values through time. While tests for structural change have been developed in Quandt (1960) this does lead to a problem. How many of these structural changes have occurred? While methods can be developed to answer this problem this still means we impose a somewhat arbitrary assumption. Secondly what if the beta returns back to it's original value?

Viewed as a possible extension to structural change models we have regime switching models, which is employed in for example Hamilton (1989). Simply put, the various regimes are specified, after which the process to determine in which regime we actually are is specified. Note that this can be done by using for example an observed state variable (Hansen (2000)) or as in the case of Hamilton (1989) by specifying unobserved regimes and using a Markov process to determine the (likelihood of the) actual regime.

A practical application of a regime switching model is Kon and Jen (1978). The paper concerns the performance of mutual funds managers. That is a priori the author assumes that he manager acts differently (and thus selects portfolios with different beta) when the manager predicted a bull or bear market. As a result we thus have two different regimes. He specified a biniomial process to determine the regime and estimated the two different betas. Kon and Jen (1978) shows a significant non-linearity, that is the assumption of constant beta is severely violated.

There are also Bayesian methods to solve this problem. For example in Sarris (1973). The author tries to estimate the time-varying β while leaving the beta completely free. As a result he tries to estimate T different βs using T observations, which is of course not enough data. To fix this problem he first specifies the prior of all possible paths of β, leaving the starting point and variance undetermined. The author then estimates the starting point and variance using the available data and use the usual Bayesian methods to move from a specified prior to a posterior. The author shows that the estimation method yields desirable properties (minimum variance) if the prior is specified correctly, but these hold only asymptotically if the prior is misspecified.

For our purpose we can also look at ways to test asset pricing models without assuming time-invariance. An example of this is given by Harvey (1989). Here the author first rewrites the CAPM in a conditional form:

$$E(R_{jt}|Z_{t-1}) = \lambda cov[(R_{jt}, R_{mt}|Z_{t-1}] \tag{17}$$

where R_{jt} is the return on asset j at time t, and Z_{t-1} is the available information up and until $t-1$. The asset returns (LHS) are rewritten by Harvey (1989) as:

$$u_t = r_t - Z_{t-1}\delta \tag{18}$$

where u_t is the forecast errors and δ is the coefficient matrix. Secondly the disturbance vector (RHS) is rewritten as:

$$e_t = r_t - \alpha - \lambda(r_t - Z_{t-1}\delta)(r_m - Z_{t-1}\delta_m) \tag{19}$$

Where α is the asset specific intercept, serving as an analogy to Jensen (1969). Stacking both equations together:

$$\epsilon_t = (u_t, e_t) = \begin{pmatrix} r_t - Z_{t-1}\delta \\ r_t - \alpha - \lambda(r_t - Z_{t-1}\delta)(r_m - Z_{t-1}\delta_m) \end{pmatrix} \tag{20}$$

The model can be estimated by GMM with the following moment equation:

$$\begin{gathered} g = vec(\epsilon Z) \\ min_{(\alpha,\lambda)} g'Wg \end{gathered} \tag{21}$$

Where W is the weighting matrix as specified by Hansen (1982).

Last we have DCB as specified by Engle (2016). While being a general model by design, it estimates the beta by looking at the conditional covariance between variables, letting the covariance be specified by a particular model. Engle (2016) uses the DCC method described in Engle (2002) and finds that indeed when applied to Fama & French factors, the betas vary quite substantially over time.

While all discussed methods have particular advantages and disadvantages, for our purposes only Engle (2016) suffices. Because we relate two different sets of risk measures to eachother and wish to show that the Asset pricing models βs are related to eachother, both in the cross section and through time regime switching models simply supply too little. We wish to have a distinct beta for every time period, not for every regime. Harvey (1989) provides a direct method to test asset pricing models in a conditional setting, compared to for example Fama & Mecbeth regressions. However this method does not provide β's, hence we would be unable to relate the obtained estimations from higher order CAPM and Fama & French. Furthermore, if we take a look at equation (15), this is essentially the same equation as (27), where the second is a conditional version. As a result, we will use DCB from Engle (2016) as our preferred method.

However the DCB method has received a negative review since its inception. For example Hollstein and Prokopczuk (2016) discusses various methods to estimate CAPM and Fama & French β's. They included the DCB, OLS as proposed by Fama and MacBeth (1973) and methods based upon implied moments obtained from forward looking options. Hollstein and Prokopczuk (2016) shows that the DCB combined with DCC as used in Engle (2016) produce large errors compared to the estimation methods who combine the information from both historical and forward looking data.

4 Methodology & Data

4.1 DCB

4.1.1 Estimating β

To introduce the dynamic conditional beta, we follow the notation of Engle (2016). That is given a set of observables $(y_t, x_t)'$, with t = 1,..,T we try to find $f_{t-1}(y_t|x_t)$. Also known as the conditional distribution of y given x and the past realizations of y and x. Usually we are interested in $E_{t-1}(y_t|x_t)$. Often we assume linearity in x and parameters independent of the information set, however more generally this function can be non-linear or depend on past realizations.

We can formulate this problems as follows, given:

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} |F_{t-1} \sim N(\mu_t, H_t). \tag{22}$$

The conditional distribution is:

$$y_t|x_t, F_{t-1} \sim N(\mu_{y,t} + H_{yx,t}H_{xx,t}^{-1}(x_t - \mu_{x,t}), H_{yy,t} - H_{yx,t}H_{xx,t}^{-1}H_{xy,t}). \tag{23}$$

After which we can obtain the conditional beta

$$\beta_t = H_{xx,t}^{-1}H_{xy,t}. \tag{24}$$

If we drop the subscript t we obtain the unconditional beta which is the same as the OLS estimator.

To estimate the dynamic conditional beta we use maximum likelihood estimation. We replace μ_t and H_t in equation 22 with a possible parametrization $\mu_t(\theta)$ and $H_t(\theta)$:

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} |F_{t-1} \sim N(\mu_t(\theta), H_t(\theta)), \tag{25}$$

note that this model can encapsulate various different models such as VEC or DCC (Engle (2002)). From this equation we easily obtain the maximum likelihood and in turn obtain the MLE as:

$$\begin{aligned} \hat{\theta} = argmax \quad & L(\theta; z_1, ..., z_T) = \\ argmax \quad & -\frac{1}{2}\sum_{t=1}^{T}(log|H_t(\theta)|+(z_t - \mu_t(\theta)'H_t(\theta)^{-1}(z_t - u_t(\theta))), \end{aligned} \tag{26}$$

where $z_t = (y_t, x_t)$. After which the MLE of the dynamic conditional beta becomes:

$$\hat{\beta}_t = H_{xx,t}(\hat{\theta})^{-1}H_{xy,t}(\hat{\theta}). \tag{27}$$

For correct statistical testing and inference on these betas, some assumptions are necessary. However we will leave this to Engle (2016) and assume these are met throughout the rest of this paper.

4.1.2 Testing model performance

To compare the DCB model with a static model Engle (2016) proposes to use the Shwarz information criterion. To test for significance the Voung test is used (Vuong (1989) & Rivers and Vuong (2002)). That is we compute the following statistic:

$$m_t = -.5(log(\frac{h_{f,t}}{h_{g,t}}) + \epsilon_{f,t}^2 + \epsilon_{g,t}^2) - (k_f - k_g)\frac{log(T)}{T} \tag{28}$$

where $h_{f,t}$ is the partial likelihood at time t of model f, $\epsilon_{f,t}$ is the error term at time t and k_f is the total number of parameters in model f and T is the total number of observations. We regress m_t on a constant. If the resulting regression coefficient is positive we prefer model f, if it is negative we prefer model g. Following Engle (2016) as m_t is likely to be autocorrelated we use HAC standard errors to compute the t-statistic. See Rivers and Vuong (2002).

As a second test to test if the β's are constants Engle (2016) proposes the following nested model:

$$y_t = x_t'\zeta + x_t'(\beta_t(\theta) \circ \xi) + \epsilon_t, \tag{29}$$

where $\beta_t(\theta)$ is the DCB obtained from equation 24 and $\circ$ denotes the element by element multiplication as a Hadamard product. Note that $\xi = 0$ corresponds with a fixed or time-invariant beta and $\zeta = 0$ with a full DCB specification. For estimation purposes we replace θ with the MLE estimation of equation 27, $\hat{\theta}$. However as Engle (2016) notes, this leads to the generated regressor problem described in Pagan (1984). A theorem from Wooldridge (2010) shows that the parameter estimates are consistent and inference is supported. Again I will refer to Engle (2016) for a complete description of the required conditions.

4.2 Fama & French

Often the Fama & French model is specified with a time-invariant beta for all three risk factors. However as the DCB allows us to estimate the time-varying effects of these risk factors, we will estimate a time varying version:

$$\begin{aligned} R_{j,t} - R_{f,t} = \alpha_j &+ \beta_{FF,m,j,t}(R_{m,t} - R_{f,t}) \\ &+ \beta_{FF,smb,j,t} R_{smb,t} \\ &+ \beta_{FF,hml,j,t} R_{hml,t} + e_{j,t}, \end{aligned} \tag{30}$$

where $(R_{j,t}, R_{m,t}, R_{smb,t}, R_{hml,t}, R_{f,t})$ are the returns at time t of asset or portfolio j the market return, the small minus big portfolio, the high minus low portfolio and the risk free rate. To estimate the dynamic conditional betas we follow Engle (2016) and use DCC as specified in Engle (2002). To compare the performance of DCB with a static estimation, we estimate a constant beta specification using GJR-GARCH to model serial dependence in the errors. We will compare these two models using the Schwarz information criterion and test for significance using the Vuong test. This is exactly the same as Engle (2016) for the purpose of replication.

As a second test whether the betas are time-varying or constant we will also estimate the nested model, where we include each factor with both a time-invariant and time-variant coefficient as described in equation 29.

To test if the factors are sufficient to explain the returns generated, we test for all assets if the α is equal to zero, as predicted by both the CAPM and Fama and French (1992). In addition to DCB & GJR-GARCH, we will also estimate equation (30) with OLS and report the α t-statistic. Lastly we also include the α t-statistic for the nested model.

4.3 Higher order co-moments

To estimate to effect of the higher order co-moments as risk factors we extend the model of Hung et al. (2004). While they found disappointing evidence, in contrast to Harvey and Siddique (2000), they attribute this mostly to lack of data. Secondly, Harvey and Siddique (2000) first estimate the original CAPM model, that is regressing the asset return only on the market return, and use the error terms to construct the co-skewness instead of the more direct approach of Hung et al. (2004). Thirdly Hung et al. (2004) allows for direct extension with time-varying betas. As last argument see equation (14) & (15) in section 2.2.1, which suggest Hung et al. (2004) is indeed better from a theoretical viewpoint.

Thus we use the following model:

$$\begin{aligned} R_{t,j} - R_{f,t} = \alpha_j &+ \beta_{HOCAPM,1,j,t}(R_{m,t} - R_{f,t}) \\ &+ \beta_{HOCAPM,2,j,t}(R_{m,t} - R_{f,t})^2 \\ &+ \beta_{HOCAPM,3,j,t}(R_{m,t} - R_{f,t})^3. \end{aligned} \tag{31}$$

Since we are comparing this model with the Fama & French model we use exactly the same estimation methods and statistical test. First we use DCB as described in Engle (2016) and again we use DCC as described in Engle (2002). Again we estimate a GJR-GARCH model and compare these models using Schwarz information criterion and a Voung test. We again estimate the nested DCB model for a second test of DCB. To complete the testing of the models we again test for non-significant α's, and include an estimation with OLS.

4.4 The price of risk

To investigate if these factors are priced, we will use Fama & Macbeth regressions and equation (13). Originally positioned in Fama and MacBeth (1973) as a two step method, that is first regress the various asset returns on a market return, after which the obtained beta's are regressed period-by-period on the stock return. The first step is replaced by the the DCB-method, obtaining T β's per portfolio instead of 1 per portfolio. We can replace the beta in each period by dynamic conditional beta obtained, thus obtaining the following second regression step:

$$R_{j,t} - R_{f,t} = \gamma_{0,t} + \gamma_{1,t}\hat{\beta}_{x,1,j,t} + \gamma_{2,t}\hat{\beta}_{x,2,j,t} + \gamma_{3,t}\hat{\beta}_{x,3,j,t} + \nu_{t,j}, \tag{32}$$

where $\beta_{x,i,j,t}$ corresponds to a specific risk factor beta from an asset at time t, either from the Fama & French model or the higher order CAPM model. To determine if these factors are priced, it must hold that $E(\gamma_{k,t}) \neq 0$, however they may vary trough time. Secondly we adopt the up/down differential adopted in Hung et al. (2004) but originally positioned in Pettengill et al. (1995). That is we divide the period in up and down periods, that is if the return over a month was positive we label it as an up month and vice versa. After this we will average the results over the respective periods. To determine significance we will use a bootstrap method. We re-sample with repetition all obtained prices by the various regression in (32) 20000 times, to obtain a empirical distribution. The p-value is than the fraction of this distribution more extreme than the actual value obtained.

Secondly we will assume that the price is constant in (32) that is we obtain the following regression:

$$\begin{aligned} R_{j,t} - R_{f,t} = \gamma_0 + \gamma_1^- D^- \hat{\beta}_{x,1,j,t} + \gamma_1^+ D^+ \hat{\beta}_{x,1,j,t} + \\ \gamma_2^- D^- \hat{\beta}_{x,2,j,t} + \gamma_2^+ D^+ \hat{\beta}_{x,2,j,t} + \\ \gamma_3^- D^- \hat{\beta}_{x,3,j,t} + \gamma_3^+ D^+ \hat{\beta}_{x,3,j,t} + \\ \nu_{t,j}, \end{aligned} \tag{33}$$

where $D^{\binom{+}{-}}$ specifies if the total return was positive or negative in a month. To avoid problems with possible outliers, we will estimate the regression with the Huber loss function Huber and Ronchetti (1981), avoiding the negative effects of large outliers.

4.5 Relation between higher order co-moments and Fama & French

To draw a full comparison between higher order moments and the Fama & French factors we wish to answer if the Fama & French factors somehow proxy or are related to Higher order moments. We will take two step to achieve this. First we show that the Fama & French Factors returns are correlated with Higher order moments of the market. Secondly we perform a regression between the obtained β's from DCB of the Fama & French model and the Higher order moments.

4.5.1 Relation between factors

To show that the factors are correlated we use the following seemingly unrelated regression (SUR).

$$\begin{aligned} (R_M - R_f)^2 =& \varphi_{21} + \varphi_{22}(R_M - R_f) + \varphi_{23} R_{HML} + \varphi_{24} R_{HML} + \epsilon_1 \\ (R_M - R_f)^3 =& \varphi_{31} + \varphi_{32}(R_M - R_f) + \varphi_{33} R_{HML} + \varphi_{34} R_{HML} + \epsilon_2 \end{aligned} \tag{34}$$

We use the iterated Feasible generalized least squares (FGLS) estimator to estimate the model.

4.5.2 Relation between factor β's

After using the methods described in section 4.2 and 4.3 we obtain two time-varying β vectors for each company. Let them be denoted by $\beta_{FF,t,j}$ and $\beta_{HOCAPM,t,j}$. We will make the assumption that the effect between the fama & French betas and the higher order co-moments β's are constant trough time:

$$\begin{aligned} \beta_{HOCAPM,t,1} &= \alpha + A\beta_{FF,t,1} + \psi_{1,t} \\ \beta_{HOCAPM,t,2} &= \alpha + A\beta_{FF,t,2} + \psi_{2,t} \\ &\dots \\ \beta_{HOCAPM,t,N} &= \alpha + A\beta_{FF,t,N} + \psi_{N,t} \end{aligned} \tag{35}$$

This is essentially a version of SUR with cross-restrictions. Hence this model will be estimated with FGLS. Within this model both the cross-sectional and time-varying effects of the β's are captured with the same coefficient A. A possible alternative specification is:

$$\begin{aligned} \beta_{HOCAPM,t,1} &= \alpha + B\overline{\beta_{FF,1}} + C_1(\beta_{FF,t,1} - \overline{\beta_{FF,1}}) + \psi_{1,t} \\ \beta_{HOCAPM,t,2} &= \alpha + B\overline{\beta_{FF,2}} + C_2(\beta_{FF,t,2} - \overline{\beta_{FF,2}}) + \psi_{1,t} \\ &\dots \\ \beta_{HOCAPM,t,N} &= \alpha + B\overline{\beta_{FF,N}} + C_N(\beta_{FF,t,N} - \overline{\beta_{FF,N}}) + \psi_{1,t}, \end{aligned} \tag{36}$$

where $\overline{\beta}_{FF,1}$ denotes the average beta over the whole time period. In this model B is the matrix which corresponds with the cross-sectional relation between the Fama & French factors and the higher order co-moments, while C_j corresponds to the time-varying effect. Secondly we allow that the effect of the time-varying betas are different for each asset, possibly correcting for a possible non-linear effect, however we will also estimate the model with the restriction $C_1 = C_2 = ... = C_N$. Also model 36 reduces to model 35 for $B = C_1 = ... = C_N$.

4.6 Data

The data that will be used is from Ken French's website, French (2018). The data period selected is from 1963-2011 and concerns daily data, as in Engle (2016). The dataset consist of the daily returns from each of the 12 portfolios selected by Engle (2016), the risk free rate on each day and the return of the market portfolio minus the risk free rate, the SMB portfolio and the HML portfolio.

5 Results

5.1 Fama & French

Table 1 shows that the DCB outperforms static GARCH estimation consistently. Note only has the DCB a lower Schwarz criteria for all assets, the non-nested VOUNG test clearly shows the Dynamic β to be superior to GARCH for every portfolio. This is similar to the results of Engle (2016), although with slightly different numbers. Possible revisions of the data used in the construction of the Fama & French portfolios could be too blame. We conclude essentially the same as Engle (2016), using dynamic conditional β improves the fit of the Fama & French model.

Table 1: Schwarz Information criterion & Vuong test for constant vs. varying betas Fama & French

Name	Schwarz DCB	Schwarz GARCH	VUONG TEST
BusEq	1.42	1.71	10.09
Chems	0.92	1.12	10.60
Durbl	1.66	1.88	10.87
Enrgy	1.79	2.10	8.46
Hlth	1.18	1.41	9.24
Manuf	0.13	0.35	9.76
Money	0.52	0.83	8.14
NoDur	0.51	0.84	8.08
Other	0.48	0.63	10.27
Shops	0.93	1.14	9.59
Telcm	1.55	1.77	11.25
Utils	0.73	1.02	6.44

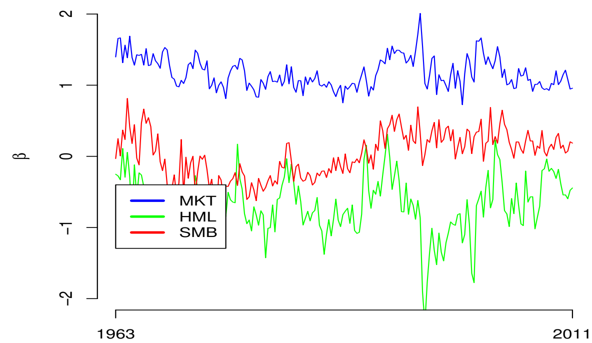
Notes to Table: The Table shows the result when comparing dynamic conditional β and GJR-GARCH models. For the Schwarz criterion lower is better while for the t-values of the VOUNG test, higher is better.

The results of the nested model test is reported in Table 2. All of the t-values of the dynamic part of the nested model are highly significant. Thus both fixed and completely dynamic hypotheses are rejected every time. This again provides significant evidence that dynamic conditional β estimation improves the model fit over static estimation and adds explanatory power to the regression. Again this replicates the original paper of Engle (2016), which shows and concluded the same.

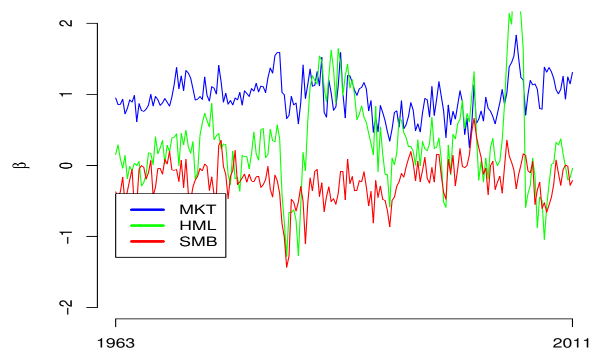
Table 3 shows the result when considering mis-pricing. The DCB t-statistic is closer to zero for almost all portfolios compared to the static estimation while modelling the regression errors as GJR-GARCH. As the pricing model predicts that the α should be zero, the DCB estimation improves the pricing model validity.

If we look at figure 3a, 3b & 3c we see the primary reason for the results. Even considering most of the movement appears as random noise clear trends can be distinguished. First while the business equipment portfolio (3a) seems to have the lowest overall movement, first the SMB β slowly declines after which it start to slowly raise again over the remainder of the period. Secondly the energy portfolio (3b) seems to have two major spikes, first halfway during the period and secondly just before the financial crises of 2007. Lastly Money (3c has the clearest differences. Again similar to Engle (2016), first the money portfolio behaves as a growth stock in 2004-2007, but reverses to a value stock during the financial crisis.

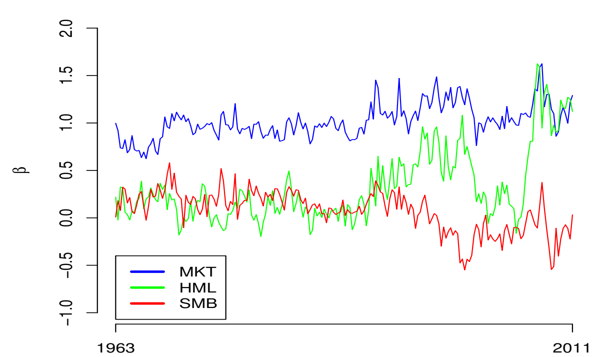
Figure 3: FF-factors βs



(a) Quarterly averages of FF-factors β's using DCB estimation for Businesses equipment



(b) Quarterly averages of FF-factors β's using DCB estimation for Energy



(c) Quarterly averages of FF-factors β's using DCB estimation for Money

Table 2: Nested DCB model, Fama & French

Name	FIX MKT	FIX SMB	FIX HML	DCB MKT	DCB SMB	DCB HML
BusEq	14.66	-4.20	-8.53	25.35	21.92	28.95
Chems	14.89	-9.36	-2.56	21.70	9.00	19.98
Durbl	15.60	1.38	8.49	16.56	19.24	25.40
Enrgy	12.23	-8.18	1.45	25.97	11.80	31.64
Hlth	11.15	-8.90	-6.79	23.98	9.94	24.99
Manuf	18.10	12.78	8.98	17.68	16.67	26.87
Money	13.92	10.64	11.15	39.14	27.34	42.08
NoDur	12.71	-0.98	-0.87	33.75	28.53	28.85
Other	10.79	23.88	5.45	10.34	18.27	23.09
Shops	13.05	5.72	-4.28	19.84	11.43	22.56
Telcm	11.81	-10.41	4.92	20.50	6.41	25.88
Utils	12.02	2.29	7.08	33.10	17.10	29.47

Notes to Table: The Table shows the result when comparing dynamic conditional beta and and static estimation. The values reported are the t-statistic of the coefficients for the nested model where each risk factor is included with a static parameter and a time-varying parameter obtained from DCB.

Table 3: Test for a zero α, FF-factor model

Name	TSTAT DCB	TSTAT NEST	TSTAT GARCH	TSTAT OLS
BusEq	1.73	1.25	1.17	1.98
Chems	-0.32	0.03	0.63	0.92
Durbl	-2.30	-2.31	-2.65	-2.37
Enrgy	0.23	1.03	2.61	1.19
Hlth	2.28	4.22	4.90	3.11
Manuf	-0.33	-2.80	-2.16	-0.33
Money	0.31	-2.55	-2.88	-3.54
NoDur	2.01	2.87	4.22	3.49
Other	-0.58	-1.50	-2.60	-2.67
Shops	0.72	1.50	1.97	1.53
Telcm	-1.08	-0.48	0.40	-0.32
Utils	0.18	-2.98	-2.60	0.03

Notes to Table: The Table shows the results of the t-test for 4 models. OLS with HAC errors, GARCH estimation, DCB estimation and a the nested model. The t-values are for the constants within these models except for DCB. Here the pricing error is computed by removing the effects of the risk portfolios after which the t-values are computed using HAC consistent errors by regressing these values on a constant. Significant t-values show significant mis-pricing after correction for risk factors. Closer to zero is better.

5.2 Higher order co-moments

When estimating the higher order Co-moments β's the differences between DCB and GARCH disappear. In half of the cases the Schwarz information criterion is higher and furthermore the differences are quite small. However the Vuong test tells us these difference are significant. The main reason for this is discovered when looking at figure 4, while the market β varies as much as in Fama & French, the β when looked at higher moments stay relatively the same. The β varies relatively little compared to the case of Fama & French. However looking at 4 we see that for the money portfolio the average beta is quite often above zero, suggesting co-skewness and co-kurtosis factors are indeed important for pricing while the contrary might be expected if these factor are not important.

Table 4: Schwarz Information criterion & Vuong test for constant vs. varying betas HOCAPM

Name	Schwarz DCB	Schwarz GARCH	VUONG TEST
BusEq	1.84	1.87	10.16
Chems	1.12	1.14	7.46
Durbl	1.92	1.92	10.72
Enrgy	2.12	2.13	7.79
Hlth	1.54	1.51	-10.48
Manuf	0.39	0.41	7.48
Money	0.90	0.93	-5.78
NoDur	0.86	0.84	-10.20
Other	0.86	0.82	-5.92
Shops	1.17	1.15	-5.09
Telcm	1.82	1.82	2.47
Utils	1.09	1.10	-0.68

Notes to Table: The Table shows the result when comparing dynamic conditional beta and GJR-GARCH models. For the Schwartz criterion lower prefers DCB while for the t-values of the VOUNG test, higher prefers DCB.

Table 5 even more clearly shows that the higher order moment are less time-variant. While for the Fama & French factors all t-values for the nested model are high and always above the crital value, the results for the higher order moments are mixed. While the market factor is still significantly time-varying, the higher order moments differ greatly between portfolios. In practice, one should take care and perform test if the β of higher order moments are indeed time-varying. If not, a mixed DCB model as described in Engle (2016) might be more appropriate.

When looking at the significance of the mis-pricing (table 6), the DCB model performs somewhat better than the static GARCH model. On average the t-statistics are smaller than that of the GARCH model.Thus if one would only care about the absence of mis-pricing, the dynamic approach might be the better on. However when looking at the results compared to the OLS estimation we see that when DCB outperforms GJR-GARCH OLS outperforms GJR-GARCH as well. While in the FF-factors estimation GJR-GARCH outperformed OLS for most portfolios, OLS does perform better when taking the higher order moments as risk factors. A possible explanation is that after the OLS-regression, the residuals of the higher order moments have lower skewness (f.e. 0.02% vs 0.6% for Energy portfolio) and the squared residuals have a lower autocorrelation (f.e. 0.20 vs 0.25). Thus the the data is closer to the actual assumption of OLS, letting it perform slightly better than in the case of using FF-factors. However DCB does perform reasonably well against OLS regardless.

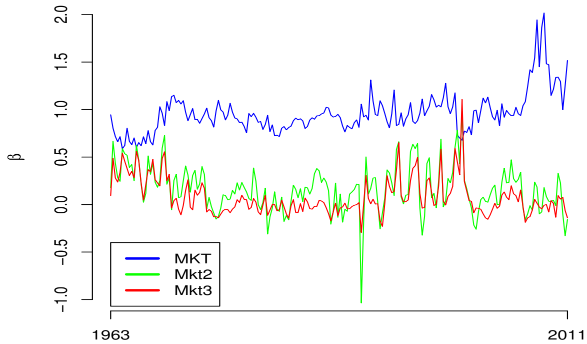


Figure 4: Quarterly averages of HOCAPM-factors β's using DCB estimation for Money

Table 5: Nested DCB model, HOCAPM

Name	FIXMKT	FIXMKT2	FIXMKT3	DCBMKT	DCBMKT2	DCBMKT3
BusEq	7.79	5.27	6.00	33.06	3.55	4.95
Chems	12.59	1.86	0.36	28.56	0.70	-1.98
Durbl	11.42	0.38	2.81	22.58	-0.20	0.73
Enrgy	13.79	2.08	-3.75	24.79	0.56	-1.18
Hlth	7.44	3.21	-2.32	34.44	4.55	0.22
Manuf	11.54	-0.07	-1.36	28.27	-0.99	0.45
Money	10.78	3.00	0.74	34.95	5.40	3.14
NoDur	11.11	0.14	-2.60	37.27	1.11	0.05
Other	9.88	-5.55	-2.64	29.80	0.68	0.81
Shops	11.19	1.68	0.43	25.37	2.70	-0.18
Telcm	9.30	3.87	4.37	25.93	4.52	2.50
Utils	12.16	2.42	-0.88	27.87	1.94	2.74

Notes to Table: The Table shows the result when comparing dynamic conditional beta and and static estimation. The values reported are the t-statistic of the coefficients for the nested model where each risk factor is included with a static parameter and a time-varying parameter obtained from DCB.

Finally comparing the Fama-French factors to the higher order moments, we first look at the model fit. In most cases the Schwarz information criterion is higher for the HOCAPM model, signifying it fits the data less. However when looking at mis-pricing of the models, α of the models this pattern is not visible. Looking a the models in 8 of the 12 cases the HOCAPM model has less significant pricing errors. Furthermore HOCAPM has only 1 significant mis-pricing while Fama & French has 3. However a part of this effects stems for the higher variance of the residuals after taking the risk factors into account, the second part being a slightly lower mis pricing average (the mean of the absolute pricing error is 0.040% vs 0.045% over all portfolios). Thus it is not possible to conclude HOCAPM is better at correctly pricing the assets, nor the other way around.

Table 6: Test for a zero α, HOCAPM model

Name	TSTAT DCB	TSTAT NEST	TSTAT GARCH	TSTAT OLS
BusEq	-1.17	-4.00	-1.92	-0.96
Chems	-0.15	-0.63	-0.28	1.60
Durbl	-2.05	-2.82	-1.87	-0.12
Enrgy	1.75	3.19	3.63	1.02
Hlth	0.79	2.86	3.79	2.31
Manuf	0.15	-0.41	-0.27	1.77
Money	-0.18	-0.40	0.63	-0.84
NoDur	1.89	4.10	4.77	3.52
Other	1.01	1.63	0.81	0.84
Shops	0.32	0.46	1.46	1.28
Telcm	-0.75	-1.64	-0.11	-1.56
Utils	1.19	-0.25	-0.31	-0.09

Notes to Table: The Table shows the results of the t-test for 4 models. OLS with HAC errors, GARCH estimation, DCB estimation and a the nested model. The t-values are for the constants within these models except for DCB. Here the pricing error is computed by removing the effects of the risk portfolios after which the t-values are computed using HAC consistent errors by regressing these values on a constant. Significant t-values show significant mis-pricing after correction for risk factors. Closer to zero is better.

5.3 The price of risk

Looking at table 7 we see mostly significant results for the Fama-French factor model. However the result for the HOCAPM model are less promising. While the Market risk factor is indeed priced in both up and down markets as expected the higher order co-moments do not seem to be priced. The second step in the Fama-Macbeth regression is a regression of only 12 observations. This might have lead to the massive variance and even more massive outliers when looking at the price in each period, creating unreliability in conclusions inferred from the results. Thus looking at table 8, which considered constant price, we see that most of the considered risk factors are significantly priced. The only exception is the co-kurtosis risk factor. The difference between the tables thus mostly from the lack of portfolios considered and the problems in using only 12 data points in the second step of Fama-Macbeth regressions.

Table 7: Pricing of the risk measures, Fama-Macbeth

HOCAPM					FF				
	up	p-value	down	p-value		up	p-value	down	p-value
intercept	0.07	0.001	0.01	0.63	intercept	0.05	0.001	-0.01	0.40
Mkt	0.09	0.001	-0.18	0.001	Mkt	0.10	0.001	-0.16	0.001
Mkt2	-0.01	0.95	0.79	0.052	HML	-0.02	0.003	0.05	0.001
Mkt3	-0.06	0.94	-10.94	0.115	SMB	0.08	0.001	-0.06	0.001

Notes to table: Results for the second step of Fama & Macbeth regressions, using Dynamically estimated betas, differentiated between up and down markets. The p-value is obtained by bootstrap, while the value is the average over the whole time-period.

Again looking at table 8 we see that the HOCAPM risk measured are priced as expected. First in an up market, when the co-skewness divided by the total skewness of the market changes with an absolute 1%, the average return drops by 0.004% on average. As expected the covariance and co-kurtosis are negatively priced in an up-market, thus investors dislike the common price movements between an portfolio and the market and dislike simultaneously extreme events. The co-skewness has a negative price, thus investor appreciate assets which move in the same direction as the market. Furthermore all prices are in a down market of opposite sign as positioned in Pettengill et al. (1995). To conclude we have significant evidence that the FF-factors are priced when using time-varying perspective. Secondly we have some evidence that higher order moments are indeed priced, with only the co-kurtosis factor having no overwhelming evidence.

Table 8: Pricing of the risk measures, constant prices regression

HOCAPM					FF				
	up	p-value	down	p-value		up	p-value	down	p-value
intercept	—	—	0.06	0.001	intercept	—	—	0.05	0.001
Mkt	0.29	0.001	-0.20	0.001	Mkt	0.09	0.001	-0.19	0.001
Mkt2	-0.45	0.001	0.23	0.022	HML	-0.03	0.001	0.04	0.001
Mkt3	0.28	0.053	-0.65	0.001	SMB	0.05	0.001	-0.05	0.001

Notes to table: Results obtained from the robust regression using a Huber and Ronchetti (1981) loss function. The return of the 12 portfolios are directly regressed on the dynamically estimated betas with the assumption of constant price. The p-values are computed using HAC-errors. Only a single intercept was included.

5.4 Relation between higher order co-moments and Fama & French

5.4.1 Relation between factors

Table 9: Direct SUR-model for higher order market moments and Fama & French Factors

Equation	parameter	value	t-value	R^2
Mkt^2	Intercept	1.02	26.77	1.5%
	Market	-0.53	-13.49	
	HML	-0.12	-1.56	
	SMB	-0.46	-5.86	
Mkt^3	Intercept	-0.92	-2.05	13.9%
	Market	19.27	41.49	
	HML	-8.13	-9.45	
	SMB	5.83	6.38	

Estimation results of SUR-model when regressing higher order market moments on the Fama & French portfolios. Market has been included to account for the linear effects the market returns has on its square and cube. The estimated value, t-value and R^2 per equation are reported. R^2 is relative to variance after including the market return alone, and thus represents the improvement when including Fama & French Factors.

When regressing the Market factors directly on the Fama & French Factors we obtain the results from table 9. First, a return of the SMB portfolio of 1% leads to a expected distance from the risk free rate of 0.75% when considering the market, down from 1%. Secondly if a particular day had a SMB return in excess of 0.2%, the market would be on average, non-negatively skewed. However the explained variance is low. Suggesting that instead of a clear relationship between the factors, unrelated effects seem to be more important. While this reduces the evidence that they relate to the same underlying risk.

5.4.2 Relation between factor β's

The result of the simplest model to test the relation between the Fama-French risk factors and that of higher order moments is reported in table 10. While the risk-factors are related significantly, the value of the coefficient is small, an increase of 1 in the HML beta leads to only a decrease of 0.03 in the beta of co-skewness. However as the squared and cubic market factor is of a completely different scale than the HML & SMB factor this could be forgiven. More problematic is the low R^2. Based on the R^2 we can not conclude that the observations from the original Fama and French (1992) paper come from the higher order co-moments. Put differently, within this regression, there is no significant evidence that the Fama & French factors and the HOCAPM factors are different expressions of the same risk. Although they seem to share a small common component, they seem to explain different sources of risk entirely.

Table 10: Relationship between FF-β and HOCAPM-β

Model	parameter	value	R2
Mkt	Intercept	-0.05	79%
	Market	1.04	
	HML	-0.24	
	SMB	-0.09	
Mkt2	Intercept	-0.01	2.5%
	Market	0.02	
	HML	-0.03	
	SMB	-0.02	
Mkt3	Intercept	0.03	5.7%
	Market	-0.02	
	HML	0.01	
	SMB	-0.01	

Notes to table: Results for relationship between the risk factors. The model estimated is a simply SUR model with the HOCAPM risk factors regressed on the Fama & French risk factors. The t-values are unreported as they all all above 10. The R^2 is reported per equation.

When looking at the models which allow for varying effects in time equation (36) and between different portfolios, the results are more promising (Table 11). While the coefficients are still low in both cases (the second case unreported), the explained variance increases for the second model and increase even more for the third. This leads to two conclusion. First, the relationship between higher order moments and Fama-French factors is most likely non-linear and seem to differ when looking at different companies or a selected company through time. Secondly while the relationship between the factors is still weak, the Fama-French factors capture some effects of the higher order moments. However as the R^2 in the unrestricted model is well below 50% for the second and third equation, the Fama-French factors most likely also capture other sources of risk.

Table 11: Relationship between FF-β and HOCAPM-β

	Restricted Model					Unrestricted Model
Equation	parameter	value	parameter	value	R^2	R^2
Mkt	Intercept	-0.09			79%	82%
	Mean Market	1.07	Market	1.02		
	Mean HML	-0.22	HML	-0.24		
	Mean SMB	-0.03	SMB	-0.11		
Mkt^2	Intercept	-0.07			9.5%	19%
	Mean Market	0.08	Market	-0.014		
	Mean HML	-0.012	HML	0.01		
	Mean SMB	-0.053	SMB	-0.01		
Mkt^3	Intercept	-0.017			12.3%	27%
	Mean Market	0.026	Market	-0.05		
	Mean HML	0.006	HML	0.022		
	Mean SMB	-0.037	SMB	-0.001		

Results for the SUR model where the HOCAPM risk factors are regressed on the mean of each company and the difference at each time period. The restricted model reports the result if the relationship between the risk factors is the identical for each company, the unrestricted model allows these to vary. The t-values are unreported as they all all above 10. The last columns reports only the R^2 for the unrestricted version, coefficient estimates are left out.

If we look at the results in we take a reverse look, estimating Fama & French betas on the HOCAPM betas (Table 12 & 13) we obtain similar results. While the estimated value are close to what one might expect if these factors are related (An increase in co-skewness beta leads to a 1.77x decrease in HML beta) the explained variance is still low, with a maximum of 28.7% when we look at the HML or SMB factor. We arrive at the same conclusion, the Fama & French factors seem to explain a different type or non-diversifiable risk than co-skewness or co-kurtosis.

Table 12: Relationship between FF-β and HOCAPM-β

Model	parameter	value	R2
Market	Intercept	0.29	65.3%
	Market	0.71	
	Mkt^2	-0.03	
	Mkt^3	1.26	
HML	Intercept	0.44	10.5%
	Market	-0.42	
	Mkt^2	-1.77	
	Mkt^3	2.10	
SMB	Intercept	-0.009	1.5%
	Market	-0.02	
	Mkt^2	-0.35	
	Mkt^3	-0.51	

Notes to table: Results for relationship between the risk factors. The model estimated is a SUR model with the Fama & French risk factors regressed on the HOCAPM risk factors. The t-values are unreported as they all all above 10. The R^2 is reported per equation.

Table 13: Relationship between FF-β and HOCAPM-β

	Restricted Model					Unrestricted Model
Equation	parameter	value	parameter	value	R^2	R^2
Mkt	Intercept	0.21			66.2%	69.0%
	Mean Market	0.78	Market	0.70		
	Mean Mkt^2	-1.45	Mkt^2	0.09		
	Mean Mkt^3	5.72	Mkt^3	1.03		
HML	Intercept	0.53			21.0%	28.7%
	Mean Market	-0.64	Market	-0.11		
	Mean Mkt^2	7.74	Mkt^2	-0.98		
	Mean Mkt^3	23.93	Mkt^3	1.96		
SMB	Intercept	0.01			9.2%	14.0%
	Mean Market	0.08	Market	-0.08		
	Mean Mkt^2	2.68	Mkt^2	-0.49		
	Mean Mkt^3	-12.91	Mkt^3	-0.41		

Results for the SUR model where the estimated betas of the Fama & French risk factors are regressed on the mean HOCAPM beta of each company and the difference between the mean and the estimated beta at each time period. The restricted model reports the result if the relationship between the risk factors is the identical for each company, the unrestricted model allows these to vary. The t-values are unreported as they all all above 10. The last columns reports only the R^2 for the unrestricted version, coefficient estimates are left out.

6 Conclusion

In this paper, we have tried to compare and relate two competing asset pricing models. The original Fama & French and using higher order moments, expanding CAPM with co-skewness and co-kurtosis risk factors.

First we have redeveloped the old CAPM to encompass higher order co-moments. We have shown that as long as investors care about these co-moments through their utility function, these risk factors are theoretically priced and can be estimated almost identically to the original market factor in CAPM.

Next we looked into various ways to estimate time-varying effects. While most of these methods are excellent, for our use dynamic conditional β estimation is preferred, although it may differ if one has different objectives. The main reason why we choose DCB is simple. It provides a single estimation for each time-period, distinct from all other time-periods included in the research.

We applied DCB on the Fama & French factors and replicated the findings of Engle (2016). Using dynamic estimation improves the model fit significantly and reduces excess return obtained from the model. However when we applied dynamic estimation to our model which included higher order co-moments, the effects became less clear. The higher order co-moments risk factors seem to vary less, in some cases static estimation might even be better. However when comparing pricing errors the models didn't perform worse than static estimation.

When comparing the Fama & French factors and the higher order co-moments risk factors we conclude that Fama & French fits the data better. The Schwartz information criterion was lower for this model throughout all considered portfolios, however the results were mixed when looking at the excess return. As a results, we suggest to use Fama & French combined with DCB in practice.

After obtaining the beta estimation for various risk factors, we determined their price. We observe that both Fama & French and higher order co-moments factors are mostly priced, with only co-kurtosis holding out in up-markets. That is we reaffirm the results from Fama and French (1992) and furthermore show that there is some evidence for extending the CAPM with higher order moments.

At last we turned our attention to the relationship between these factors. First we tries to relate the returns of the Fama & French portfolios directly with the squared and cubic returns of the market portfolio. While there is some significant correlation between them, for example when HML returns are high, the market has a tendency to skew more negatively, the idiosyncratic returns are too high too draw any conclusions.

When relating the factor β's themselves, by regressing first the HOCAPM factors to the FF-factors and than reverse, we obtain similar results. While they do have a common component, that is they tend to react to same underlying risk. The shared component is too low to say that the Fama & French Factors proxy these higher order moments. That is, the results from Fama and French (1992) do not come from the fact the author simply stumbled upon two factors which mirror higher order moments.

From the same regressions we obtain the simple result that if the HOCAPM and FF-factors are related, this applies both in between asset and for the same asset through time.

For further research we have multiple suggestions. First we left a theoretical implication untested. That is if we wish to estimate the abnormal returns we need to use a efficient portfolio. These portfolios can differ for different investors. While in the original CAPM model there is only one, namely the market portfolio, allowing investors to not only select on return and volatility but also higher moments. For example one investor might tolerate higher amounts of negative skewness for the same return and volatility. We thus suggest finding various portfolios which are indeed efficient and look how estimating with various portfolios in place of the market return, obtained from French (2018), influence the results.

Another limitation is that we only used 12 already aggregated portfolios. This gave problems especially when determining the price of the risk factors. As a suggestion we suggest to repeat this exercise with more portfolios to obtain better pricing of risk factors.

It also could be interesting to look at combinations of Fama & French factors and higher order moments. Our research suggest that while the factors share a common component, they do reflect a different type or risk. Thus while beyond the scope and aim of this paper, including higher moments factors might improve the Fama & French model.

References

Dittmar, R. F. (2002). Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. *The Journal of Finance*, 57(1):369–403.

Engle, R. (2002). Dynamic conditional correlation. *Journal of Business & Economic Statistics*, 20(3):339–350.

Engle, R. F. (2016). Dynamic conditional beta. *Journal of Financial Econometrics*, 14(4):643–667.

Fama, E. F. and French, K. R. (1992). The cross-section of expected stock returns. *the Journal of Finance*, 47(2):427–465.

Fama, E. F. and MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of political economy*, 81(3):607–636.

French, K. R. (2018). Current research returns. http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society*, pages 357–384.

Hansen, B. E. (2000). Sample splitting and threshold estimation. *Econometrica*, 68(3):575–603.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, pages 1029–1054.

Harvey, C. R. (1989). Time-varying conditional covariances in tests of asset pricing models. *Journal of Financial Economics*, 24(2):289–317.

Harvey, C. R. and Siddique, A. (2000). Conditional skewness in asset pricing tests. *The Journal of Finance*, 55(3):1263–1295.

Hollstein, F. and Prokopczuk, M. (2016). Estimating beta. *Journal of Financial and Quantitative Analysis*, 51(4):1437–1466.

Huber, P. and Ronchetti, E. (1981). Robust statistics, ser. *Wiley Series in Probability and Mathematical Statistics. New York, NY, USA: Wiley-IEEE*, 52:54.

Hung, D. C.-H., Shackleton, M., and Xu, X. (2004). Capm, higher co-moment and factor models of uk stock returns. *Journal of Business Finance & Accounting*, 31(1-2):87–112.

Ide, M. (2013). 'absurd' comments about capm masquerading as research. https://mk0valuewalkgcar7lmc.kinstacdn.com/wp-content/uploads/2014/12/CAPM-2.jpg. [Online; accessed June 3, 2018].

Jensen, M. C. (1969). Risk, the pricing of capital assets, and the evaluation of investment portfolios. *The Journal of Business*, 42(2):167–247.

Kon, S. J. and Jen, F. C. (1978). Estimation of time-varying systematic risk and performance for mutual fund portfolios: an application of switching regression. *The Journal of Finance*, 33(2):457–475.

Pagan, A. (1984). Econometric issues in the analysis of regressions with generated regressors. *International Economic Review*, pages 221–247.

Pettengill, G. N., Sundaram, S., and Mathur, I. (1995). The conditional relation between beta and returns. *Journal of Financial and quantitative Analysis*, 30(1):101–116.

Quandt, R. E. (1958). The estimation of the parameters of a linear regression system obeying two separate regimes. *Journal of the american statistical association*, 53(284):873–880.

Quandt, R. E. (1960). Tests of the hypothesis that a linear regression system obeys two separate regimes. *Journal of the American statistical Association*, 55(290):324–330.

Rivers, D. and Vuong, Q. (2002). Model selection tests for nonlinear dynamic models. *The Econometrics Journal*, 5(1):1–39.

Sarris, A. H. (1973). A bayesian approach to estimation of time-varying regression coefficients. In *Annals of Economic and Social Measurement, Volume 2, number 4*, pages 501–523. NBER.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, 19(3):425–442.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica: Journal of the Econometric Society*, pages 307–333.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.