Department of Economics
SCHOOL OF ECONOMICS
AND MANAGEMENT
LUND UNIVERSITY

Master Essay I:

Portfolio Pricing with Measures of Conditional Skewness and Kurtosis

Supervisor: Prof. Björn Hansson **Author:** Natalia Lelis 870129-T083

Department of Economics, Lund University, Sweden

April 2011

Abstract

On the ground of a highly dynamic economic environment, the necessity for time-varying risk measures emerged. Inclusion of higher-order conditional moments in asset pricing models is a very common topic in recent research articles. The present essay was inspired by the seminal work of Harvey and Siddique (1999), who proposed estimation of time-varying skewness and pricing its explanatory power by a conditional three-moment CAPM. By estimating the first four conditional return moments I confirm previous findings about their high persistence, after which these risk measures are employed in testing the four-moment conditional CAPM. I analyze both time-series and cross-sectional regression results for 25 portfolios formed on different criteria (industry, size, momentum). In the time-series approach, conditional kurtosis is highly correlated with covariance and adds no pricing power. Neither conditional skewness has a well-defined impact in determining return compensation. However, in cross-sectional regressions, kurtosis risk is priced in most of the crises years, but its risk premium has the opposite sign. Investors prefer more kurtosis to less, suggesting that kurtosis is still much underestimated in financial markets during crises. Skewness is still insignificantly priced in cross-sectional CAPM. Altogether the four-moment cross-sectional CAPM performs better than its two-moment counterpart.

Keywords: Asset Pricing, CAPM, Time-varying Moments, Conditional Skewness, Conditional Kurtosis

Supervisor: Björn Hansson, Professor, Department of Economics, Lund University, Sweden

Acknowledgements

I would like to express my gratitude to Professor Björn Hansson for the great patience and valuable suggestions during the elaboration of this essay.

Table of Contents

1. Introduction………………………………..……………………..………………....1
    1.1. Background……………………………………………………………………...1
    1.2. Problem Discussion……………………………………………………………...2
    1.3. Purpose…………………………………………………………………………..3
    1.4. Limitations……………………………………………………………………….3
    1.5. Target Group……………………………………………………………………..4
    1.6. Outline…………………………………………………………………………...4
2. Previous Research…………………………………………………………………….5
    2.1. Modeling Conditional Moments………………………………………………....5
    2.2. Pricing of Conditional Moments..7
3. Methodology (Econometric Models)………………………………………………...11
    3.1. Jarque-Berra Normality Test…………………………………………………….11
    3.2. Engle Test for ARCH/GARCH Effects………………………………………...11
    3.3. GARCH Models………………………………………………………………...12
    3.4. Conditional Skewness and Kurtosis……………………………………………15
    3.5. Factor Pricing Tests……………………………………………………………..16
        3.5.1. Conditional Two-Moment CAPM of Sharpe-Lintner (1964-1965)……………..16
        3.5.2. Harvey and Siddique (2000a) Model……………………………………………17
        3.5.3. Fang and Lai (1997) Model……………………………………………………...17
4. Empirical Results…………………………………………………………………….18
    4.1. Data……………………………………………………………………………..18
    4.2. Summary Statistics and JB Test………………………………………………...19
    4.3. Test for ARCH Effects…………………………………………………………21
    4.4. Estimation of Conditional Variances and Covariances…………………………21
    4.5. Estimation of Conditional Skewness and Kurtosis……………………………..26
    4.6. Pricing of Risk Factors in a Time-Series Regression…………………………...28
    4.7. Cross-Sectional Analysis………………………………………………………..30
5. Conclusion…………………………………………………………………………...32
    5.1. Concluding remarks……………………………………………………………..32
    5.2. Possible Extensions……………………………………………………………...33

References

Appendices

EViews Codes

1. Introduction

The introduction starts with a background for the present study, after which I discuss the posed questions and set a purpose for my work. I also mention the limitations of this paper, the target group and the outline for the whole essay.

1.1. Background

Measuring risk in financial markets remains a very topical issue for analysts, fund managers and different investor groups, becoming particularly relevant during crashes and periods of financial turmoil. Assuming that capital markets are efficient, an asset's return must reflect a compensation for a given level of risk assumed by the investor by holding that particular security. Following this argument different measures have been proposed to quantify riskiness. Starting with the seminal work of Markowitz (1952) entitled "Portfolio Selection", where a portfolio's risk is measured by the second return moment – variance, some years later (1964-1965), Sharpe and Lintner developed a cornerstone asset pricing model – the CAPM, which continues to play a cardinal role in pricing securities even nowadays, although numerous extensions as well as critics of this model emerged. One major weakness of Markowitz' risk measure is the assumption of normally distributed returns, under which the distribution is symmetrical and has a kurtosis of 3. However, a well-known truth about financial series is that returns are most often negatively skewed and exhibiting excess kurtosis. Another drawback is that the second return moment proved insufficient in explaining returns of some asset categories like small caps and momentum portfolios. These facts question the risk measure based only on the second return moment and it stimulated further research regarding how skewness and kurtosis risks are priced in the market. In the present essay I will also analyze higher return moments and look at their explanatory power for return series.

1.2. Problem Discussion

Based on utility theory investors manifest aversion to variance, preference for positive skewness and aversion to excess kurtosis. Given these preferences, from two assets an individual will always choose the one with the lowest variance, the highest skewness and the lowest kurtosis.

An important aspect concerning return moments is the choice between moments themselves and co-moments between an asset and the market. For example, theory argues that variance itself is not the most important when forming portfolios. What is more relevant, is the asset's contribution to the overall risk of a portfolio, i.e. its covariance with a well-diversified portfolio, which shows the asset's systematic risk. Whether an asset's return moves in the same or in opposite directions with the market plays great importance, because in the latter case it could serve as a hedge and would certainly diminish the portfolio's risk.

Another relevant conclusion based on the previous research is that return moments are not necessarily constant over time, but can exhibit "clusters" which are especially noticeable during financial crises. In this respect the development of GARCH models (and their extensions) proved successful in modeling time-varying conditional variance (and even higher moments, e.g. Harvey and Siddique (1999)). These models revealed two variance features: *persistence* (i.e. the tendency of high conditional variance to be followed by high variance and vice versa) and *asymmetry* (negative news decrease returns more than increases due to positive shocks of same size – Nelson (1991); Glosten, Jagannathan, and Runkle (1993); Engle and Ng (1993)).

Asymmetry is a very common feature of different economic variables: stock returns, indices, exchange rates. Negative return skewness implies that negative returns of a magnitude are more probable than positive ones of the same magnitude. By theory, investors are risk-averse and thus prefer positively skewed portfolios to negative asymmetry. So, more left-skewed assets must provide higher expected return to reward extra risk. Also, securities that increase the skewness of a portfolio, must have lower return. Harvey and Siddique (1999) build a model in which skewness is proven to be time-varying just like variance.

Kurtosis or the fourth return moment can be interpreted as the variance of variance. Very often financial series exhibit excess kurtosis (i.e. fat tails), which is also a source of excess risk. So, investors would demand higher compensation on more leptokurtic assets. According to Brooks et al (2002) non-normality in return distributions comes primarily from excess kurtosis rather than skewness. So, extreme returns are more likely to occur than under normal distribution, which leads to underestimating risk when measuring it by variance alone. Necessary to note that allowing for time variation in the fourth moment implicitly presumes its dependence on dynamics in conditional variance (potential high correlation between conditional second and fourth moments).

1.3. Purpose

Since the interest in researching conditional return moments has increased lately, in the present essay I will also estimate conditional moments and co-moments between an asset and the market in order to quantify risk. My goal is to test the conditional four-moment CAPM, accounting for conditional covariance, co-skewness and co-kurtosis. The novelty of my study is that the model is tested on a different group of assets, and namely the portfolios formed by Fama-French criteria, and not just country indices or exchange rates as performed in most of the past studies. I also contribute to past research by a slightly different methodology in estimating conditional moments. For the second moments I use the BEKK model, after which I combine it with Leon, Rubio and Serna (2004) model to estimate higher moments. As stated in Harvey and Siddique (1999), results are highly dependent on the applied methodology, which is an additional motivation to try a different methodology in this essay. Thus, I pose two research questions:

1) Determine whether the first four conditional moments are priced *over time* when considering portfolios formed by industry, size and momentum criteria;
2) Determine whether the same risk factors are priced *cross-sectionally*;

1.4. Limitations

The present study is focused on one single market – US, which is considered the most liquid. Thus, portfolios formed by the same criteria but based on a different market might produce different conclusions.

1.5. Target Group

This paper is aimed at different investor groups, as well as students, professors, researchers showing interest in pricing securities.

1.6. Outline

The remainder of the essay is organized as follows:

Part 2 presents the relevant previous research for the topic under consideration.

Part 3 presents the data used in our analysis and the step-by-step methodology followed to attain the results mentioned in the next section.

Part 4 presents the results obtained by applying the proposed econometric models.

Part 5 is designed for conclusions and possible extensions of this essay.

2. Previous Research

First, the reader will be introduced to modeling methodologies for conditional return moments. Next, we consider how the first four moments enter the asset pricing models.

2.1. Modeling Conditional Moments

According to Chunhachinda et al (1997) including moments higher than the second one into investor's portfolio choice influences the selection of the optimal portfolio. Most of the studies mentioned below documented high persistence in conditional variance and higher moments (sum of the ARCH and GARCH terms is close to unity). Another common feature is the estimation of models in steps (nested models) from the simplest specification to the most complex one and using coefficient estimates obtained in the previous step as starting values in the next model.

Several alternatives have been proposed to model the documented asymmetry in asset returns. One approach is to use models for asymmetric variance: (1) *asymmetric GARCH* of Glosten, Jagannathan and Runkle, 1993 (a dummy is used to capture the higher variance produced by negative shocks); (2) Nelson, 1991 etc. An alternative way is to model conditional skewness as proposed by Harvey and Siddique (1999), where skewness is proven to be time-varying just like variance. The traditional GARCH(1,1) is extended by also including conditional skewness into the model and estimating jointly conditional values of mean, variance and skewness. In addition, Harvey and Siddique (1999) consider the interaction between conditional skewness from their model and the asymmetric variance produced by the two models mentioned above (Glosten et al (1993) and Nelson (1991)). Thus, conditional skewness is consistent with asymmetric variance and its inclusion reduces persistence in conditional variance.

Brooks et al (2002) present a similar method for modeling autoregressive conditional kurtosis (GARCHK, t-distributed errors), but without an explicit skewness equation. However, asymmetry is captured through dummy variables (in variance and kurtosis equations) similar to Glosten et al (1993) and Nelson (1991). The asymmetry coefficient in the variance equations is significant for all series, but the same coefficient in kurtosis equation is insignificant in most cases. The article also tests the significance of including variance and kurtosis terms into the mean return equation (GARCHK-M) to test the sign of risk-return tradeoff. Both of them turn out insignificant, but have the intuitive positive sign (more return compensation for higher risk).

León, Rubio and Serna (2004) present a GARCHSK methodology for modeling jointly the second, third and fourth conditional moments. The novelty of this article consists in the fact that it presents a much simpler likelihood function and at the same time captures time variation in both skewness and kurtosis based on Gram-Charlier series expansion of the normal density function, while the previous two articles accounted only for one time-varying moment higher than variance.

Jondeau and Rockinger (2003) studies stock index and exchange rate returns. Similar to the previous article, this paper also includes both conditional skewness and kurtosis into the model, concluding that conditional kurtosis is less persistent than skewness. Authors also document *cross-sectional* variability in skewness and kurtosis (extreme observations occur simultaneously in different markets). For all exchange rates there turns to be the same dynamics of asymmetry coefficient, while kurtosis coefficient is constant. For stock indices only some series have the same evolution of kurtosis coefficient, while others experience a complex evolution determined by large re-occuring economic events (crashes happen at the same time in different markets).

Table 1 below summarizes the findings of the articles mentioned above.

Article	**Year**	**Method**	**Data**	**Results**
Harvey and Siddique	1999	GARCHS(1,1,1)-M with explicit modeling of conditional skewness; Assumed non-central t-distributed errors	Daily, weekly, monthly stock returns US, Germany, Japan, Chile, Mexico, Taiwan, Thailand (excess returns) series length: 1969-1997; 1975-1997; 1980-1997	Asymmetric variance is equivalent with conditional skewness; Including conditional skewness reduces persistence in conditional variance; Time dynamics of moments also depends on data frequency (daily/monthly), seasonality (January effect, day-of-the-week effect) and aggregation of stocks into portfolios.
Brooks et al	2002	GARCHK(1,1,1) and GARCHK(1,1,1)-M with explicit modeling of	Four daily stock/bond returns (US and UK) 1990-2000	Conditional kurtosis is positively, but not significantly related to returns The response of kurtosis to good and

		conditional kurtosis, t-distributed errors		bad news is not significantly asymmetric Kurtosis decreases with time- and cross-sectional aggregation in returns
Jondeau & Rockinger	2003	GARCH Hansen's generalized t-distribution of errors Monte Carlo simulations to test model's validity	Five daily stock-index returns and four exchange rates 1971-1999 US, Germany, Canada	Skewness is very persistent, but kurtosis is less persistent Cross-sectional variability in third and fourth moments documented
León, Rubio & Serna	2004	GARCHSK(1,1,1)	Five daily stock indices (US, Germany, Spain, Mexico) and exchange rates 1990-2003	Evidence of time-varying skewness and kurtosis; Models allowing for conditional third and fourth moments outperform those based on conditional variance alone Skewness and kurtosis are less persistent than variance (lower coefficients)

Table 1. Models for conditional return moments

2.2. Pricing of Conditional Moments

After development of Sharpe-Lintner CAPM (1965), the first work to include higher moments into asset pricing (and namely skewness) belongs to Kraus and Litzenberger (1976). Assuming that investors have cubic utility function of wealth, there should be aversion to variance and preference for positive skewness. The derived three-moment CAPM implies that in equilibrium an asset's excess return equals the sum of two products: (1) market beta times the market risk premium (price of beta risk) and (2) systematic skewness (asset gamma, i.e. $\gamma = \frac{cosk_{i,m}}{skew_m}$) times a skewness premium (price of gamma risk). Prices of risk are the same for all investors due to common beliefs. Kraus and Litzenberger (1976) conclude that "the prediction of a significant price of systematic skewness is confirmed (and the price has the predicted sign) and the prediction of a zero intercept for the security market line in excess return space is not rejected".

Bollerslev, Engle and Wooldridge (1988) present a *conditional* version of covariance-based CAPM, estimated by a multivariate GARCH-M. The method is applied to bills, bond and stock returns. Authors conclude that conditional covariances are an important determinant of time-varying risk premia. But there should exist additional variables to explain variation in returns.

Harvey and Siddique (2000a) extend the three-moment CAPM of Kraus and Litzenberger (1976), by using conditional skewness instead of unconditional third moment. Due to low explanatory power of the standard Sharpe-Lintner CAPM when working with cross-sectional returns, Fama and French (1993) proposed two additional factors: size and book-to-market ratio, which proved to capture very well cross-sectional variation in returns. However, Harvey and Siddique (2000a) conclude that conditional skewness adds explanatory power to cross-sectional returns even when size and book-to-market factors are considered. Authors also note that results depend much on the used method, data, precision of asset beta computation (estimation risk). In addition, skewness is connected to momentum effect (low return momentum portfolios have higher asymmetry).

Fang and Lai (1997) presents a four-moment CAPM, where kurtosis risk is also priced. Authors find positive risk premiums for conditional skewness and conditional kurtosis.

Momentum, or "price continuation", was commonly documented in finance (*over time* - Jegadeesh and Titman, 1993, 2001, *across countries* - Rouwenhorst, 1998; Griffin et al., 2003, *across industries* - Moskowitz and Grinblatt, 1999). Momentum persists even when including the market risk (Jegadeesh and Titman, 1993), size and book-to-market value (Fama and French, 1996), and macro-factors (Griffin et al., 2003).

Similar to Harvey and Siddique (2000a), Fuertes, Miffre and Tan (2009) relate momentum effect to non-normality risks (coskewness, cokurtosis). Examining different momentum trading strategies, authors conclude that risks vary over business cycles consistent with risk aversion. Authors use the same factor construction procedure as Fama and French (1993), creating skewness- and kurtosis-mimicking portfolios. Although non-normality risks partly explain momentum returns, a large portion is still unexplained. That opposes the market efficiency hypothesis where return is a compensation for risk, leaving space for behaviorist views considering that momentum comes from a slow market response to news (incl. arbitrage limitations). Fuertes, Miffre and Tan (2009) complements Harvey and Siddique (2000) by confirming that winner returns are more negatively skewed than loser returns. Also winners have higher positive kurtosis than losers. So, "the market compensates investors with higher returns for exposure to the negative skewness and leptokurtosis of momentum returns"[1]. These findings are also consistent with a recent literature suggesting that higher moments matter in theory but may play a relatively small role in practice.

Smith (2007) also studies pricing of conditional coskewness in cross-sectional stock returns. The model is tested on 17 industry portfolios and 25 Fama-French portfolios based on market capitalization and book-to-market ratios (all taken from Kenneth French data library). Estimation is carried out using GMM and instrumental variables. The article concludes that for positive conditional market skewness investors sacrifice 7.87% annual return per unit of gamma, but require only 1.80% premium when the market has negative skewness. These results are consistent with Harvey and Siddique (2000a), where the average annual coskewness premium is 3.6%.

Table 2 below summarizes the findings of the above-mentioned articles in pricing conditional moments.

Article	**Year**	**Method**	**Data**	**Results**
Kraus & Litzenberger	1976	3-moment CAPM (OLS)	Monthly portfolio returns of NYSE stocks ranked by betas and gammas 1926-1935	The price of systematic skewness is confirmed to be significant and has the predicted positive sign; The prediction of a zero intercept for the security market line in excess return space is not rejected.
Bollerslev, Engle & Wooldridge	1988	Multivariate GARCH-M (2-moment CAPM)	Quarterly US returns on bills, bonds and stocks 1959-1984	Conditional covariances are important in explaining time variation in risk premia, but additional risk factors must exist
Fang & Lai	1997	4-moment CAPM (OLS and OLS with	Monthly portfolio returns (assets sorted by betas)	Positive risk premiums for conditional skewness and kurtosis

[1] Fuertes, Ana-Maria , Miffre, Joëlle and Tan, Wooi-Hou (2009)

		instruments)	1969-1988	
Harvey & Siddique	2000a	3-moment conditional CAPM (OLS)	Monthly US stock returns and their portfolios by different criteria: industry, size, book-to-market ratios and momentum 1963-1993	Conditional coskewness is important in cross-sectional asset pricing even after including size, book-to-market factors; Momentum effect is related to systematic skewness Measuring ex ante skewness is difficult
Smith	2007	2-moment and 3-moment conditional CAPM with instrumental variables Estimation by GMM	Monthly portfolio returns (17 industry portfolios, 25 portfolios based on size and book-to-market ratios) 1963-1997	When the market is negatively skewed, investors demand 1.8% premium for extra risk; when market is positively skewed they give up 7.87% annually; The 2-moment CAPM is rejected, but the 3-moment CAPM cannot be rejected
Fuertes, Miffre & Tan	2009	Fama & French 3-factor model complemented with skewness and kurtosis mimicking portfolios	Monthly momentum portfolios (on US stocks) 1973-2004	Skewness and kurtosis mimicking portfolios partly explain momentum returns, but a large part is still attributed to behavioral views Risks vary over the business cycle, consistent with risk aversion

Table 2. Pricing of conditional moments

3. Methodology (Econometric Models)

In this part of the essay I present the step-by-step procedure applied to obtain the estimation results provided in chapter 4.

3.1. Jarque-Berra Normality Test

To motivate the explicit modeling of the first four return moments, I first test for the presence of asymmetry and excess kurtosis in the chosen portfolios using the Jarque-Berra test:

$$JB = \frac{T}{6} * skew^2 + \frac{T}{24} * (kurt - 3)^2 \sim \chi^2(2) \qquad T - no.of\ observations$$

$$skew = \frac{\mu_3}{\sigma^3} = \frac{E[(x_i - E[x_i])^3]}{\sigma^3} \qquad kurt = \frac{\mu_4}{\sigma^4} = \frac{E[(x_i - E[x_i])^4]}{\sigma^4}$$

H_0: $skew \sim N\left(0; \frac{6}{T}\right)$ and $kurt \sim N(0; \frac{24}{T})$ which implies **normality**

H_1: $\mathbf{non-normality}$

Considering a 5% significance level: $\chi^2(2) = 5.991$; so, the null is rejected if $JB > \chi^2(2)$.

3.2. Engle Test for ARCH/GARCH Effects

The next step is to perform the Engle (1982) test for the presence of ARCH/GARCH effects in our return series, in order to make sure that we indeed need a conditional-heteroskedasticity model. I apply the test to raw (excess) returns, i.e. I run a regression on a constant only, and then regress its squared errors on their past *k* lags (I assumed k=5):

$$\boldsymbol{r_{M,t} = \alpha_0 + \eta_{M,t}} \qquad \text{and} \qquad \boldsymbol{\eta_{M,t}^{\ 2} = \sum_{i=1}^{k} \alpha_i \eta_{M,t-i}^{\ 2} + \varepsilon_t}$$

So, we set the hypotheses: H_0: $all\ \alpha_i = 0$ ($\boldsymbol{no\ ARCH\ effect}$)

$$H_1:\ at\ least\ one\ coefficient\ \alpha_i = 0$$

Extracting the R^2 measure from the residual regression, the test-statistic is TR^2, which follows a chi-squared distribution with k degrees of freedom.

3.3. GARCH models

Introduced in 1986 by Bollerslev and Taylor, these specifications allow conditional variance to depend on its previous lags and lagged past residuals. Usually a GARCH(1,1) is sufficient for explaining all the variation in conditional volatility, so I also focus on this parsimonious model, but in a bivariate setting, because univariate specifications model each asset apart from all the other ones, which is not quite realistic taking into account "volatility spillovers" occurring in financial markets[2]. Considering the bivariate GARCH-M proposed by Harvey and Siddique (1999), it primarily focuses on modeling conditional variance and skewness, but assumes a constant kurtosis. However, I would like to also estimate series of conditional fourth moments. Second, this model implies maximizing a quite complicated likelihood function. Third, I would like to focus on conditional covariance instead of variance measure, because according to CAPM this is the true measure of idiosyncratic (i.e. non-diversifiable) risk typical of a particular security, which should be priced in equilibrium. Due to these reasons, I decided to keep the bi-variate GARCH methodology, but to estimate a BEKK model instead, in order to get an explicit equation for conditional covariance. Later I use another model to find conditional third and fourth moments. Below I build three similar BEKK models and choose the most appropriate for my data.

The reason I selected a BEKK compared to a VECH model is its advantage of restricting the variance-covariance matrix to be positive definite, which is important for it in order to be invertible. The simplest form of BEKK specification proposed by Engle and Kroner (1995) is:

$$\boldsymbol{H_t} = \boldsymbol{WW'} + \boldsymbol{AH_{t-1}A'} + \boldsymbol{B\Xi_{t-1}\Xi'_{t-1}B'}$$

Where $\boldsymbol{A}$ and $\boldsymbol{B}$ are (2x2) matrices of parameters and $\boldsymbol{W}$ is a lower triangular (2x2) matrix, $\boldsymbol{H_t}$ is a (2x2) conditional variance-covariance matrix, $\boldsymbol{\Xi_t}$ is a (2x1) disturbance vector where

[2] Brooks (2008), "Introductory Econometrics for finance", 2nd ed.

$\boldsymbol{\Xi_t \sim N(0; H_t|\Omega_{t-1})}$. So, in the full bivariate BEKK, the equations for variances and covariances are given by:

$$h_{11t} = c_{11} + a_{11}h_{11t-1} + a_{12}h_{22t-1} + a_{13}h_{12t-1} + b_{11}u_{1t-1}^2 + b_{12}u_{2t-1}^2 + b_{13}u_{1t-1}u_{2t-1}$$

$$h_{22t} = c_{21} + a_{21}h_{11t-1} + a_{22}h_{22t-1} + a_{23}h_{12t-1} + b_{21}u_{1t-1}^2 + b_{22}u_{2t-1}^2 + b_{23}u_{1t-1}u_{2t-1}$$

$$h_{12t} = c_{31} + a_{31}h_{11t-1} + a_{32}h_{22t-1} + a_{33}h_{12t-1} + b_{31}u_{1t-1}^2 + b_{32}u_{2t-1}^2 + b_{33}u_{1t-1}u_{2t-1}$$

However, the main disadvantage is the large number of coefficients to estimate. For that reason, we simplify our model by assuming a diagonal BEKK, for which each variance and covariance depends only on its own past lags and lagged disturbances. So, the ***first GARCH*** I estimate is the simple diagonal BEKK[3]:

$$\boldsymbol{r_{i.t} = \alpha_{0,i} + \eta_{i,t}}$$

$$\boldsymbol{r_{M.t} = \alpha_{0,M} + \eta_{M,t}}$$

$$\boldsymbol{h_{i.t} = \beta_{0,i} + \beta_{1,i}h_{i,t-1} + \beta_{2,i}\eta_{i,t-1}{}^2}$$

$$\boldsymbol{h_{M.t} = \beta_{0,M} + \beta_{1,M}h_{M,t-1} + \beta_{2,M}\eta_{M,t-1}{}^2}$$

$$\boldsymbol{h_{im.t} = \beta_{0,iM} + \beta_{1,iM}h_{im,t-1} + \beta_{2,iM}\eta_{i,t-1}\eta_{M,t-1}}$$

For the second and third models I add some extra terms, based on the model proposed by Harvey and Siddique:

GARCH 2 (add excess market return in the mean equation):

$$\boldsymbol{r_{i.t} = \alpha_{0,i} + \alpha_{1,i}r_{M,t} + \eta_{i,t}}$$

$$\boldsymbol{r_{M.t} = \alpha_{0,M} + \eta_{M,t}}$$

$$\boldsymbol{h_{i.t} = \beta_{0,i} + \beta_{1,i}h_{i,t-1} + \beta_{2,i}\eta_{i,t-1}{}^2}$$

$$\boldsymbol{h_{M.t} = \beta_{0,M} + \beta_{1,M}h_{M,t-1} + \beta_{2,M}\eta_{M,t-1}{}^2}$$

$$\boldsymbol{h_{im.t} = \beta_{0,iM} + \beta_{1,iM}h_{im,t-1} + \beta_{2,iM}\eta_{i,t-1}\eta_{M,t-1}}$$

[3] In initial specifications I also tried to include the terms $\delta_1 h_{M,t} + \delta_2 \eta_{i,t-1}\eta_{M,t-1}$ in all the equations of asset variance (similar to Harvey and Siddique, 1999), but due to high correlation between variables they always turn out insignificant

GARCH 3 (add variances in mean eq.): $\boldsymbol{r_{i.t} = \alpha_{0,i} + \alpha_{2,i} h_{i,t-1} + \eta_{i,t}}$

$$\boldsymbol{r_{M.t} = \alpha_{0,M} + \alpha_{2,M} h_{M,t-1} + \eta_{M,t}}$$

$$\boldsymbol{h_{i.t} = \beta_{0,i} + \beta_{1,i} h_{i,t-1} + \beta_{2,i} \eta_{i,t-1}{}^2}$$

$$\boldsymbol{h_{M.t} = \beta_{0,M} + \beta_{1,M} h_{M,t-1} + \beta_{2,M} \eta_{M,t-1}{}^2}$$

$$\boldsymbol{h_{im.t} = \beta_{0,iM} + \beta_{1,iM} h_{im,t-1} + \beta_{2,iM} \eta_{i,t-1} \eta_{M,t-1}}$$

For stationarity, in each model I restrict the variance and covariance equation coefficients: $\boldsymbol{\beta_{1,i}}$, $\boldsymbol{\beta_{1,M}}$, $\boldsymbol{\beta_{2,i}}$, $\boldsymbol{\beta_{2,M}}$, $\boldsymbol{\beta_{1,iM}}$ and $\boldsymbol{\beta_{2,iM}}$ to lie between (0;1).

For simplicity and comparability reasons I assume conditionally normal distribution of errors in all the models below. As these three models are not built-in in EViews (EViews 6 supports the diagonal BEKK only for variance equations, but not for covariance equation also), I wrote special programs for their estimation, whose codes are provided in the appendix.

Since GARCH models are non-linear, we can no longer use OLS to estimate them and need to apply the maximum likelihood technique. So, we need to build a log-likelihood function and maximize it with respect to each unknown parameter. As known, this function can be highly non-linear and it is cardinal to choose adequate starting values for our parameters in order to reach the global maximum of our likelihood function and not just a local extreme point. In this respect I use several steps. First, I estimate all the corresponding regressions by OLS and use the computed coefficients as parameter initial values under similar *univariate* GARCH models. Finally, we use the coefficients estimated from the univariate GARCH specifications as starting values for estimation of all the bi-variate models. This nested-model procedure also serves as a diagnostic test for the estimated parameters. The initial values for conditional variance and covariance are set to their unconditional values.

Assuming conditionally normal distribution of residuals, the log-likelihood function for all the bivariate models from above has the form:

$$L(\theta) = -\frac{TN}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}(\log|H_t| + \boldsymbol{\Xi_t'} H_t{}^{-1} \boldsymbol{\Xi_t})$$

where $\mathrm{N} = 2$; $|H_t|$ is the determinant of H_t, and θ are all the unknown parameters.

To conclude, the series we obtain from the models above include: conditional portfolio variances and conditional covariances between each portfolio and the market.

3.4. Conditional Skewness and Conditional Kurtosis

The next step is to jointly estimate conditional third and fourth moments. For this purpose, I apply the methodology proposed by León, Rubio and Serna (2003), because it has an easier likelihood function than Harvey and Siddique (1999) and it also allows for time-varying kurtosis, not just conditional skewness:

$$r_t = E_{t-1}(r_t) + \varepsilon_t \qquad \varepsilon_t \sim (0, h_t | \Omega_{t-1})$$

$$\varepsilon_t = h^{\frac{1}{2}} \eta_t \qquad \eta_t \sim (0; 1) \qquad E_{t-1}({\eta_t}^3) = s_t \qquad E_{t-1}({\eta_t}^4) = k_t$$

$$h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 h_{t-1}$$

$$s_t = \gamma_0 + \gamma_1 \eta_{t-1}^3 + \gamma_2 s_{t-1}$$

$$k_t = \delta_0 + \delta_1 \eta_{t-1}^4 + \delta_2 k_{t-1}$$

To get stationary variance, skewness and kurtosis (also positive variance and kurtosis), some constraints are set: $\mathbf{0} < \boldsymbol{\beta_1}, \boldsymbol{\beta_2}, (\boldsymbol{\beta_1} + \boldsymbol{\beta_2}) < 1$; $\mathbf{-1} < \boldsymbol{\gamma_1}, \boldsymbol{\gamma_2}, (\boldsymbol{\gamma_1} + \boldsymbol{\gamma_2}) < 1$ and $\mathbf{0} < \boldsymbol{\delta_1}, \boldsymbol{\delta_2}, (\boldsymbol{\delta_1} + \boldsymbol{\delta_2}) < 1$.

This model is estimated using a Gram-Charlier series expansion of the normal density function for the standardized errors, which is truncated at the fourth moment:

- pdf: $f(\eta_t | I_{t-1}) = \Phi(\eta_t) \psi(\eta_t)^2 / \Gamma_t$

where $\psi(\eta_t) = 1 + \frac{s_t}{3!}(\eta_t^3 - 3\eta_t) + \frac{k_t - 3}{4!}(\eta_t^4 - 6{\eta_t}^2 + 3)$ and $\Gamma_t = 1 + \frac{s_t}{3!} + \frac{(k_t - 3)^2}{4!}$

- $log - likelihood\ function$: $l_t = -\frac{1}{2} ln h_t - \frac{1}{2} {\eta_t}^2 + \ln(\psi(\eta_t)^2) - \ln(\Gamma_t)$

So, the likelihood function is similar to the normal distribution, but in addition it also has two adjustment terms for time-varying skewness and kurtosis.

To conclude, this last model will provide series for conditional skewness and kurtosis for each portfolio $\boldsymbol{i}$.

3.5. Factor Pricing Tests

To test whether the first four return moments are priced in the market I use the CAPM models summarized below.

3.5.1. Conditional Two-Moment CAPM of Sharpe-Lintner (1964-1965)

$$E_{t-1}(r_{i,t}) = \beta_{i,t-1}E_{t-1}(r_{m,t}) \qquad \beta_{i,t-1} = \frac{cov_{t-1}(r_{i,t}, r_{m,t})}{var(r_{m,t})}$$

To estimate this relation in a *time-series regression* (excess asset returns against the market excess return), we require that the intercept should be zero, the betas - significant, and the market risk premium – be the same for all assets. In a *cross-sectional regression* (excess returns against betas), the slope (i.e. market risk premium), should be significantly different from zero.

However literature (Campbell et al 1997) suggests that even the *conditional* CAPM with time-varying covariances, betas and the market variance explains insufficiently stock returns. Moreover, the mean–variance CAPM require normally distributed returns or quadratic utility. It is well-known that returns are non-normal, while quadratic utility has the drawback of increasing absolute risk aversion. Consequently, nonlinear asset pricing models (e.g. the three-moment CAPM etc.) perform better. However, there is one critique of even higher-moment CAPM. Post, Levy, and van Vliet (2008) conclude that when risk aversion restriction is set, the implied cubic utility function has an inverted "S" shape, which means that optimization might lead to a global minimum rather than the needed maximum.

3.5.2. Harvey and Siddique (2000a) Model

They extend the three-moment CAPM of Kraus and Litzenberger (1976) to its conditional form:

$$E_{t-1}(r_{i,t}) = \beta_{i,t-1}\mu_{1,t} + \gamma_{i,t-1}\mu_{2,t} \qquad \beta_{i,t-1} = \frac{cov_{t-1}(r_{i,t}, r_{m,t})}{var(r_{m,t})} \qquad \gamma_{i,t-1} = \frac{cosk_{t-1}(r_{i,t}, r_{m,t})}{skew(r_{m,t})}$$

$$skew_{t-1}(r_{m,t}) = E_{t-1}[\left(r_{m,t} - E_{t-1}(r_{m,t})\right)^3]$$

$$cosk_{t-1}(r_{i,t}, r_{m,t}) = E_{t-1}[\left(r_{i,t} - E_{t-1}(r_{i,t})\right)\left(r_{m,t} - E_{t-1}(r_{m,t})\right)^2]$$

$\mu_{1,t}$ is the price of beta risk and $\mu_{2,t}$ is the price of gamma risk. Since investors prefer positive skewness, the sign of $\mu_{2,t}$ should be opposite to the sign of the conditional *market* skewness $skew(r_{m,t})$, because in equilibrium want to sacrifice return for positive skewness, but would require a premium for negatively skewed returns.

3.5.3. Fang and Lai (1997) Model

$$r_i = b_1\beta_i + b_2\gamma_i + b_3\delta_i \qquad \beta_i = \frac{cov(r_i, r_m)}{var(r_m)} \qquad \gamma_i = \frac{cosk(r_i, {r_m}^2)}{skew(r_m)} \qquad \delta_i = \frac{cokurt(r_i, {r_m}^3)}{kurt(r_m)}$$

r_i is expected excess return on asset *i*; β_i is the systematic beta, γ_i is the systematic skewness and δ_i is the systematic kurtosis of asset *i*; b_1, b_2 and b_3 are the market risk premiums corresponding to these risks (according to utility theory $b_2 < 0$; $b_1, b_3 > 0$).

4. Empirical Results

Initially the reader is introduced to the data sample used in this essay, after which I illustrate the results obtained by applying the method described in the previous chapter.

4.1. Data

In the present essay I will focus on several aggregated return series, based on all NYSE, AMEX, and NASDAQ stocks:

- 10 US Industry Portfolios;
- 5 quintile-based Size Portfolios;
- 10 Momentum Portfolios (value-weighted returns for 10 prior-return portfolios (from (t-12) to (t-2)).

All data is collected with *monthly* frequency over the period: January 1970 – December 2010. I do not consider higher frequency in order to simplify analysis and to avoid disturbing effects from the possible day-of-the-week effect (Foster and Viswanathan (1993) conclude that Mondays have higher volatility and trading costs).

I also consider a return series to represent the market portfolio. In order to ensure data comparability, all the variables mentioned above are extracted from the same data source, and namely the Kenneth French data library[4].

My choice of these particular portfolios is motivated by their representativeness for the economy as a whole (the industry portfolios) and some CAPM deficiencies when it comes for explaining "the smallest market-capitalized deciles and returns from specific strategies such as ones based on momentum"[5].

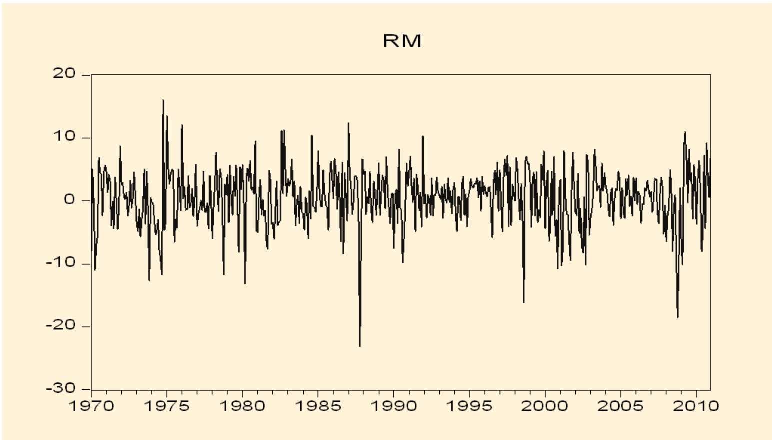
Since all the series are already in returns form, there is no need for compounding. However, for each series I use not the returns themselves, but the excess portfolio returns over the risk-free rate (one-month T-Bill rate from Ibbotson and Associates).

[4] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
[5] Harvey and Siddique (2000a)

4.2. Summary Statistics and JB test

As a first step, I analyze return series in order to conclude that a GARCH methodology is indeed appropriate. As an example, the market excess returns are considered here, while table 1 in appendix provides similar summary statistics for the other 25 portfolios.

Below I plot the market excess return along the time axis. As can be noted, there is indeed much variability in this variable over the last 41 years (492 monthly observations) and its variance can hardly be approximated by a constant. Intuitively, a GARCH methodology would work better in this setting.



Further, we look at the summary statistics for this variable and formally test whether it is normally distributed, by applying the joint Jarque-Berra test of normality[6]:

$$skew = \frac{\mu_3}{\sigma^3} = \frac{E[(x_i - E[x_i])^3]}{\sigma^3} = \frac{-58.9744}{103.8980} = -0.568 \qquad negative\ skewness$$

$$kurt = \frac{\mu_4}{\sigma^4} = \frac{E[(x_i - E[x_i])^4]}{\sigma^4} = \frac{2374.3383}{488.4384} = 4.861 > 3 \qquad leptokurtosis$$

$$JB = \frac{T}{6} * skew^2 + \frac{T}{24} * (kurt - 3)^2 \ \sim \chi^2(2) \qquad T - no. of\ observations$$

$$JB = 82 * 0.3226 + 20.5 * 3.4633 = 26.4532 + 70.9977 = 97.45$$

[6] Note: small differences between my Excel calculation and EViews values don't affect the general conclusion and are due to approximation and adjustment for degrees of freedom; asymptotically results must coincide

H_0: $skew \sim N\left(0; \frac{6}{T}\right)$ and $kurt \sim N(0; \frac{24}{T})$ which implies **normality**

H_1: **non − normality**

Considering a 5% significance level, $\chi^2(2) = 5.991$ and obviously the null is strongly rejected at this significance. So, market returns are not normally distributed, being characterized by the common features of negative skewness and excess kurtosis.

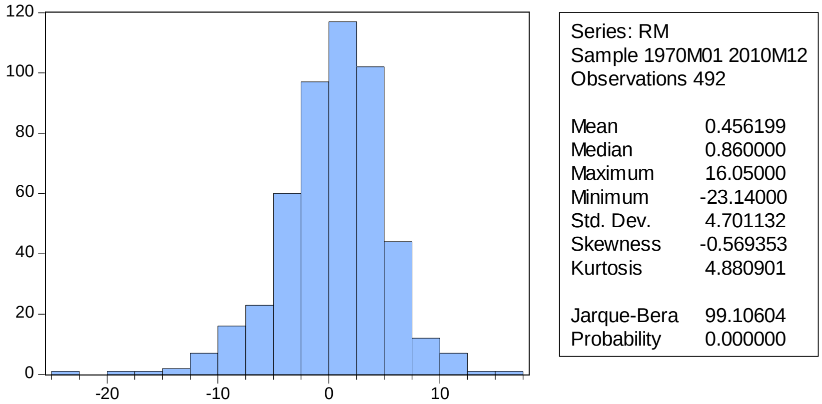


Table 1 in appendix shows that the normality hypothesis is also strongly rejected for all the 25 portfolios. As can be noted, most series have negative skewness (except three industry and three momentum portfolios). Moreover, all the series exhibit excess kurtosis.

Another conclusion from table 1 is that standard deviation (or variance) alone is not enough to describe the risk-return relationship. For example, *Durables* portfolio has the same mean return as the market, but its standard deviation is much higher (6.54 compared to 4.70 for the market). Also, *Energy* has a higher return than *Durables*, but a lower standard deviation. *Hitec* has the lowest mean excess return, but the highest risk, judging by variance. As a matter of fact, accounting for the 3rd and 4th moments in asset pricing might be a reasonable intuition behind the methodology presented in the previous section. However, it is still questionable whether this is sufficient to fully explain the overall risk of an asset. For example, comparing the market portfolio with *Utilities*, we note that they have the same mean return, but Utilities have lower risk judging by all three measures: Variance, Skewness and Kurtosis.

4.3. Test for ARCH Effects

As mentioned in the methodology section, I apply Engle test on raw excess return series. Table 2 in appendix shows results for k=5 residual lags. Assuming a 5% significance level, the corresponding chi-squared value is: $\boldsymbol{\chi^2(5) = 11.07}$. Thus, at five lags the null (of homoskedasticity) is rejected for most excess return series (except 8 portfolios). Further, keeping the same significance level and trying other lag values, I also detected ARCH effects in four out of these portfolios: *Shops* (four lags, p-value=0.043), *P6* (one lag, p-val=0.020), *P7* (one lag, p-val=0.036) and *P9* (one lag, p-val=0.035). However, the remaining four excess portfolio returns (*Qnt2, Qnt3, Qnt4, P8*) do not exhibit error autocorrelation. So, their variance could be roughly approximated by a constant. However, in my analysis below I decided to apply the GARCH methodology even on these returns, in order to check the model's applicability to a more general setting (because when we have a very large number of assets, it might be cumbersome to check for ARCH effects in each series).

4.4. Extimation of Conditional Variances and Covariances

When estimating models, I include the whole period under consideration (January 1970 – December 2010) and do not leave the last years for forecasting as usual, because, first, I want to produce series of conditional covariances over the whole period and, second, the last years incorporate the effect of the global financial crisis and might be relevant in asset pricing using higher moments.

As mentioned above, it is very important to choose proper starting values for coefficients when maximizing the log-likelihood function. So, I estimate the models in stages: first, I estimate the mean equations alone using OLS, then use the coefficients as starting values for the univariate GARCH estimation, and finally use the obtained results as starting values for the bivariate GARCH. Table 3 in appendix shows the estimates found for all the three models (insignificant coefficients at 5% level are highlighted), as well as the corresponding log-likelihood value. Initially, I will not look closely to interpretation of coefficients, but on their overall significance, in order to motivate the choice of the best GARCH specification among the three proposed models. So, for now I consider how many insignificant coefficients appear under each GARCH, and what is their relevance for depicting the studied portfolios. As can be noted, each model generates quite consistent conclusions across all portfolios. Thus, almost all coefficients for the *first GARCH* are highly statistically significant (under 1 %), and only two portfolios (*Durables* and *Low*) have insignificant intercepts. Also we can note that even the four portfolios which did not exhibit ARCH effects under Engle test (*Qnt2, Qnt3, Qnt4, P8*) have highly significant coefficients. It follows that this modeling procedure works even for assets, whose variance can be approximated by a constant.

Intercepts estimated through the GARCH can be interpreted as the long-run mean return on each portfolio. So, we can compare them to mean returns presented in table 1. Since these figures are quite close to each other, the first model seems to describe well the portfolio returns. However, by also looking at the intercepts in the second mean equation, we conclude that the model slightly overestimates the long-run market return, because almost all intercepts are higher than 0.46.

In order to make results more clear, for the second and the third model specifications, I do not present p-values, but just highlight the insignificant coefficients under 5% level. This is because below I will motivate my choice of the first model as the best to describe the data. However, these missing p-values are available upon request.

Looking at the results of the *second GARCH*, we conclude that including the market excess return into the first mean equation makes most of its intercepts insignificant at 5% level, while the estimated coefficient for the market return turns out highly significant for most portfolios (consistent with CAPM theory, because regressing excess asset returns on the market excess return, the intercept shows mispricing and should be insignificant in a time-series regression). Here also, the intercept in the second mean equation slightly overestimates the mean market return. Additionally, omega(2) is insignificant for most portfolios, and taking account that it determines the intercept of covariance equation, there seems to be no long-run trend component in covariance series.

Making an overall comparison between coefficient values for the first two models, we conclude that there is not much change in them. As mentioned, the portfolio's intercept becomes insignificant, being totally captured by the market excess return factor. Also due to insignificance of omega(2), it has an increasing effect on omega(3). So, in conclusion, the variance and covariance equations do not change much across these two models.

In the third model, I include conditional variances into the mean equations (building a GARCH-M similar to Harvey and Siddique). As can be seen, their coefficients turn out insignificant for most of the portfolios, which implies very weak explanatory power for these factors. This result might be attributed to *aggregation criteria*, because Harvey and Siddique (1999) mention that this factor, together with data *frequency* and *seasonality*, could have impact on results. As many insignificant estimates produce noise in the model, I leave out this last GARCH and focus only on the first two ones.

In order to choose the best specification, I use the *likelihood ratio* (LR) test, because GARCH 1 can be considered a restricted version of GARCH 2, when the market coefficient is set to zero. So, we can test whether the imposed restriction is supported by the data. The hypotheses we set are:

$$H_0: \boldsymbol{\alpha_{1,i}} = \mathbf{0} \text{ (coefficient of market return)} \qquad H_1: \boldsymbol{\alpha_{1,i}} \neq \mathbf{0}$$

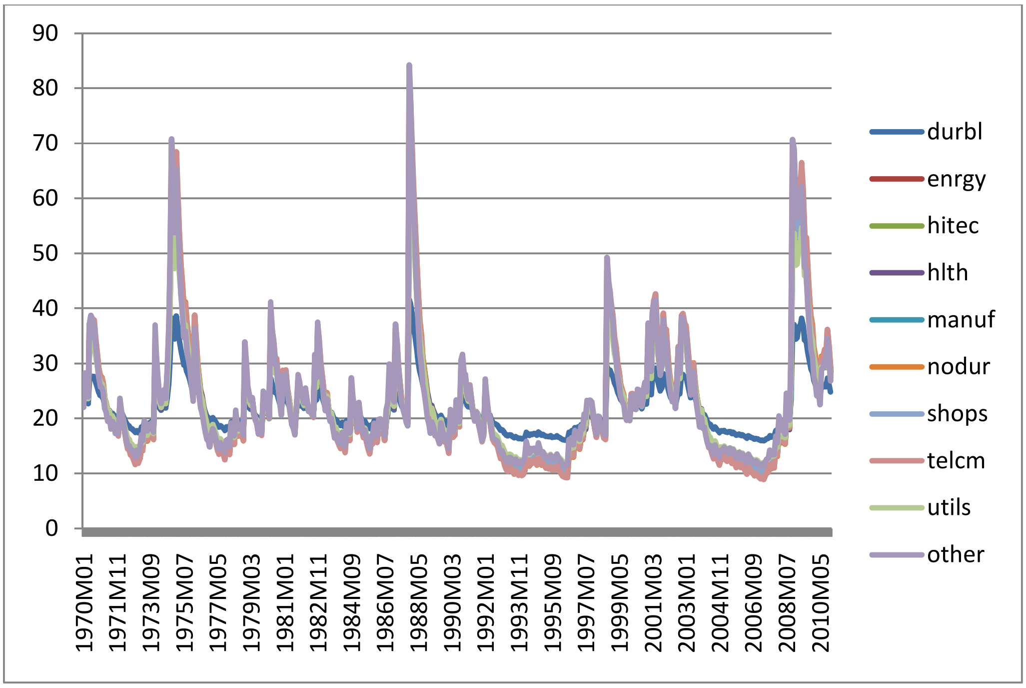
$$LR = -2(logL^* - logL) \quad \sim \quad \chi^2(1)$$

We have one restriction, because only one coefficient in the first mean equation is set to zero. So, considering a 5% significance level, I compare the computed LR value to $\chi^2(1) = 3.841$. The test results are illustrated in appendix (table 4). Thus, for most portfolios (except 4 ones: *Other, Hi20, P3, High*) the null cannot be rejected, implying that the restriction is supported by the data[7]. So, I decided to choose the most parsimonious model, i.e. the first GARCH.

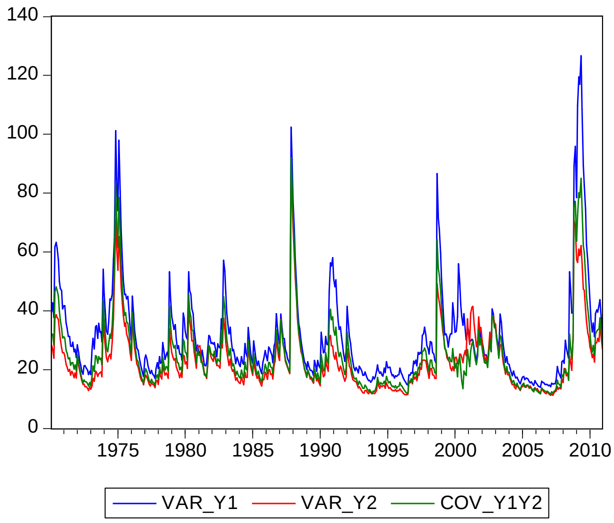
Next, we look closer at coefficients of this chosen model. Table 5 in appendix presents its final GARCH coefficients, which follow directly from the BEKK's multiplication rules. In the final columns (columns 2,3,4 counting from the right of the table) I computed the sum of ARCH and GARCH coefficients for each variance-covariance equation, in order to show that the model is not "explosive" (stationarity), because each sum is less than one. Consistent with conclusions in Harvey and Siddique (1999), I find high persistence in conditional variance (and covariance),

[7] In some cases I got negative LR, which can be explained by the fact that sometimes the distribution of the test statistic is complicated, being a combination of chi-square

because the coefficients $\boldsymbol{\beta_{1,i}}$, $\boldsymbol{\beta_{1,M}}$, $\boldsymbol{\beta_{1,iM}}$ are greater than 0.75 for most portfolios. For only one portfolio (*Low*) these coefficients are substantially lower, indicating weaker volatility persistence. Theoretically, the coefficients $\boldsymbol{\alpha_{0,M}}$, $\boldsymbol{\beta_{0,M}}$, $\boldsymbol{\beta_{1,M}}$ and $\boldsymbol{\beta_{2,M}}$ must be the same across all portfolios, because they describe the same mean and variance equations for market excess return, but combined with different portfolios. However, looking at their values in table 5, we can note differences although for every portfolio estimation starts with the same initial values coming from the univariate GARCH. For this reason I plotted all conditional market variances resulting from different portfolios within the same graph to compare their behavior over time. The graph below shows results (Note: all the plotted series are variances for market portfolio, but the names of series in the legend simply denote the GARCH model to which they belong).



As can be noted all the lines nearly coincide, which is a normal result and proves comparability for our models. The line that departs mostly from the others is “Durbl”, and for this reason this market variance is not quite reliable for further testing. So, to further represent market variance I chose the portfolio “Other”, because as shown below it has most similarities with the market return variability, and that is why I consider it more trustworthy.



The value of $\boldsymbol{\beta_{1,M} = 0.8062}$ for portfolio "Other" shows that market volatility in one period consists of 81% of its value from the previous period (high persistency). The remaining 19% represent news entering the market. So, agents primarily ground their expectations on the historical data, which is quite convenient for forecasting future volatility.

Now, I consider coefficients $\boldsymbol{\beta_{1,i}}$ and $\boldsymbol{\beta_{1,iM}}$, which represent in each portfolio's variance and covariance respectively. The values of these estimates also show high dependence on past information for most portfolios (over 70%). Consistent with previous research, some momentum returns are most difficult to explain and as a consequence they are characterized by the lowest coefficients: *Low*, *P3*, *P7* and *High*.

Another conclusion resulting from table 5 is that $\boldsymbol{\beta_{0,i}}$, $\boldsymbol{\beta_{0,M}}$ and $\boldsymbol{\beta_{0,iM}}$ (representing intercepts in variance-covariance equations) are all positive (and as mentioned before highly significant). These positive values are intuitive, because they suggest that there always exist a long-run trend in variances and covariances, i.e. there is some unavoidable positive risk in each portfolio. However, the absolute value of these coefficients cannot be always used to compare risk across portfolios. We can only judge about the riskiness of an asset after also taking account of previously analyzed persistency coefficients: $\boldsymbol{\beta_{1,i}}$, $\boldsymbol{\beta_{1,M}}$, $\boldsymbol{\beta_{1,iM}}$. For example, portfolio *Low* has highest intercepts, but lowest persistency coefficients. Thus, the overall risk (represented by variance and covariance) might happen to be comparable to other industry portfolios (and not necessarily higher). This is an obvious result due to the large number of coefficients estimated under the BEKK model, and also imposing the additional restriction to have a positive definite variance-covariance matrix. Thus, I consider a better way to compare riskiness of different portfolios is to benchmark their mean returns (i.e. intercepts $\boldsymbol{\alpha_{0,i}}$) against the market mean return $\boldsymbol{\alpha_{0,M}}$**,** because the highest return is expected to be associated with higher risk. As mentioned above, equations for market excess return differ slightly from each other, but that does not impact the overall market volatility. So, to compare portfolios, in the last column of table 5, I compute the ratio between each portfolio intercept and the corresponding intercept for market excess return. If the ratio is higher than unity, that portfolio is riskier than the market, otherwise (if ratio is less than one) it is less risky. These ratios are also comparable across portfolios, because the market mean return is a common benchmark for all of them. In addition, comparing them to Mean returns from table 1, we can note that they suggest similar conclusions across portfolios.

So, analyzing these ratios, we conclude that judging by mean returns the riskiest portfolios are: *High, P8, Energy, Qnt2* and *Qnt3*. The lowest risk is typical of portfolio *Low,* which also has insignificant intercept in our GARCH model, while in table 1 it was the only asset to have a negative mean return (also possibly not statistically different from zero).

The next step is to estimate another model, which will produce series of conditional skewness and kurtosis.

4.5. Estimation of Conditional Skewness and Kurtosis

As mentioned above, I use León, Rubio and Serna (2003) model to compute series of conditional third and fourth moments:

$$r_t = \omega + \varepsilon_t \qquad \varepsilon_t \sim (0, h_t | \Omega_{t-1})$$

$$\varepsilon_t = h^{\frac{1}{2}} \eta_t \qquad \eta_t \sim (0; 1) \qquad E_{t-1}({\eta_t}^3) = s_t \qquad E_{t-1}({\eta_t}^4) = k_t$$

$$h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 h_{t-1}$$

$$s_t = \gamma_0 + \gamma_1 \eta_{t-1}^3 + \gamma_2 s_{t-1}$$

$$k_t = \delta_0 + \delta_1 \eta_{t-1}^4 + \delta_2 k_{t-1}$$

As can be noted, conditional variance is computed based on usual residuals from the model, but the conditional third and fourth moments are estimated using standardized errors. To ensure comparability between this model and the previously estimated BEKK, I set the variance equation coefficients to their estimated values using the BEKK model. Initial skewness and kurtosis are set to their unconditional sample values over the first three years (Jan 1970 – Dec 1972), because skewness and kurtosis change a lot during the studied period and it would be more precise to use their values over a smaller sample.

To maximize the log-likelihood function, which is very non-linear, I tried different initial values for coefficients (with 0.1 increments) and the ones providing maximum likelihood value are shown in appendix (table 6), together with the corresponding estimation results. As before p-values are shown in brackets. Imposing the necessary restrictions on parameters, I get a stationary model.

Looking at the skewness equation's intercepts, we conclude that for seven portfolios they turn out insignificant at 5% level, which implies no long-run trend component in skewness for the portfolios: *Enrgy, Hlth, Qnt4, Low, P4, P6* and *P9*. So, these assets' skewness is only influenced by past period's asymmetry and news in the current period. Also, two garch terms (skewness equation for *P7* and kurtosis eq. for *Utils*) are insignificant. For the rest, the coefficient estimates are highly significant.

The last two columns show averages of each estimated series of conditional skewness and kurtosis. Comparing them to the unconditional third and fourth moments from table 1, we can conclude that this model captures quite well the variability of skewness and kurtosis over time and across assets.

Analyzing estimated values for γ_2 and δ_2, we note that there is less persistence in third and fourth moment over time compared to previously estimated variances and covariances. These coefficients range from 0.5 to 0.8, which implies that 50-80% of current period skewness and kurtosis is still due to their values in the past period. Also, coefficients γ_1 and δ_1 showing the skewness and kurtosis part attributable to errors (i.e. innovations to the market) is higher. As expected, all kurtosis intercepts are positive, because they show the long-run component which is independent on previous information. For skewness equation, some intercepts turn negative, while others positive, which is normal taking account that most portfolios exhibit negative skewness in table 1.

4.6. Pricing of Risk Factors in a Time-Series Regression

Before proceeding to test how the first four moments are priced in return series, I examine correlations between conditional covariances, (co)skewness and (co)kurtosis (table 7). First, it should be noted that covariance is more correlated with co-moments (co-skewness, co-kurtosis) rather than conditional moments (skewness, kurtosis). To avoid spurious results, in the time-series regression I test the pricing of skewness and kurtosis instead of co-skewness and co-kurtosis with the market. A second conclusion is that, consistent with previous research, (co)kurtosis exhibits high (but not perfect) correlation with conditional covariance. So, either kurtosis should be excluded from the study, or we can apply principal component analysis to build independent factors out of the conditional second and fourth moments, or we can use the pricing errors from the second and third return moments to regress on kurtosis. To choose among these options, I regressed portfolio returns on conditional kurtosis alone. The results are not presented in appendix, but they show that for most portfolios kurtosis risk is not priced. Significant coefficients (under 10%) appear only for six assets: *Enrgy, Low, P2, P3, P4, P5*. So, by applying principal components we risk including irrelevant information in the model and producing noise. However, excluding this variable from the analysis might skip a potentially relevant risk factor for the six portfolios mentioned above. Thus, I decided to test kurtosis pricing only for these six portfolios, based on errors left after other *statistically significant* factors (i.e. if no coefficient from the previous regression is significant, I regress the initial excess portfolio return on kurtosis alone). The first three columns in table 8 show results from regressing excess returns on conditional covariance and skewness. The next column presents coefficient estimates for kurtosis based on pricing errors from the first model. Then, the last two columns illustrate coefficients from a model based only on covariance.

As can be noted, from all 25 portfolios, after taking account of covariance risk, the fourth return moment has no more explanatory power, except for one asset: *Enrgy*. However, this unique significant coefficient on kurtosis might appear occasionally and not be representative for pricing assets. We should also note that it even does not have the needed sign, because in theory there should be a positive relation between returns and the fourth moment ("fatter tails" imply more risk and require a higher return compensation). Besides that the other five kurtosis coefficients are insignificant at 10% level, their signs are also inconclusive, because some are positive, others negative. According to these results, in a time-series regression kurtosis risk has no pricing power after accounting for the second return moment.

Next, I consider the covariance-based model in the last two columns of table 8. As can be noted this risk is priced for the market returns (in the form of conditional variance), as well as 15 other portfolios. Insignificant coefficients appear for four high-momentum portfolios (*P7, P8, P9, High*), the highest quintile when sorting assets by size (*Hi20*), as well as five industry portfolios (*Energy, Hitech, Health, Telecommunications* and *Utilities*). The results for momentum portfolios are consistent with previous findings that there might appear difficulties in pricing this group of assets. However, for the portfolios formed by size, I obtained quite opposite results, because it is not the lowest-cap assets that are problematic to explain, but on the contrary – the highest quintile. That might be explained intuitively, if we assume that investors associate highest cap companies with more financial health and thus lower risk, no matter of the size of their covariance with the market return. It should be noted that all covariance coefficients have the expected positive sign (the higher risk measured by covariance, the higher must be compensation as return). Interestingly, all intercepts for portfolios where covariance risk is priced, are negative, implying that investors accept more covariance risk than they are compensated for. Additionally, five assets with priced second return moment still have significant intercepts, showing that there is mispricing left to be explained by other risk factors.

Finally, I look at the first three columns in table 8, which show results after additionally including skewness into the model. Here covariance risk is priced for the same portfolios as in the previous regression, which shows consistency in results. Overall, conditional skewness is priced for eight assets and the market return. All these significant coefficients have a positive sign, although by theory gamma should be negative. By also looking to other assets where the skewness coefficient turned out insignificant, we conclude that for only five portfolios the sign is negative. This leads to conclusion that skewness risk, similar to kurtosis is not accounted by investors, because it has a significant slope only randomly, and its sign is not the expected one. As a matter of fact, I come to the same conclusion as Fuertes, Miffre and Tan (2009) (although by a different model), and namely that higher return moments are priced only in theory, while in financial markets return compensation primarily comes for accepting variance (covariance) risks. A second conclusion is that there should exist other risk factors to be priced in the market, because a great part of returns still remains unexplained. This last point is a motivation for further research, in order to identify and test the pricing of some different factors.

4.7. Cross-Sectional Analysis

I also perform a cross-sectional analysis by regressing average yearly returns on yearly conditional betas, gammas and deltas. Since, skewness and kurtosis risks are especially noticeable during crises, I focus my analysis on the two most recent crises: the dot-com bubble (2001, 2002) and the last financial crisis (2007-2009), together with the years following them (2003, 2010). Table 9 in appendix shows the results of the respective regressions (insignificant coefficients at 10% level are highlighted). Initially we consider the correlations between betas, gammas and deltas in the last three columns. As can be seen there are only two highly correlated series corresponding to years 2009 and 2010. To avoid spurious results, when I test the last two regressions I only include three moments in the CAPM tests.

Now, we compare R-squared measures of the conditional two-moment CAPM (the middle three columns) with the same measure for a four-moment CAPM presented in the initial columns. For all years except 2010, the four-moment CAPM has better explanatory power, because it is associated with a higher R-squared. However, the effect is only marginal for 2008 and there is no rise in R-squared at all for 2010. Moreover, for both these years neither gamma nor delta risk is priced. So, in most years the four-moment CAPM is better than its two-moment version.

As expected, beta risk premium is significant (at 10% level) for all years except 2001 (and 2007 in the four-moment CAPM when gamma premium proves more important). Another conclusion is that in periods of crises (2001-2002, 2007-2008) the risk premium for beta is negative or insignificant, while for calm periods it is positive. This result is intuitive taking account that crises are characterized by negative returns for most stocks.

Looking at the gamma risk premiums we note that they are insignificant for all series except year 2007, when the estimated premium is negative. This sign is intuitive taking account that this year was characterized by financial turmoil, which implies a negative risk premium.

A third important conclusion is that kurtosis risk is more important in cross-sectional tests than skewness, because it has more significant coefficients. This finding is consistent with Brooks et al (2002). However, the sign for delta risk premium is negative for all the series which is against the arguments presented by Fang and Lai (1997). It implies that investors behave irrationally and prefer more kurtosis to less. This may be due to the fact that kurtosis risk is still much underestimated in markets during periods of crises, implying an insufficient return compensation for this risk.

5. Conclusion

5.1. Concluding remarks

Consistent with Harvey and Siddique (1999), I find high persistence in conditional variance (and covariance) series. So, agents primarily ground their risk expectations on the historical data, which is quite convenient for forecasting future volatility. The lowest persistence coefficients are typical of some momentum portfolios (*Low, P3, P7* and *High*), which confirm previous findings that these portfolios are more difficult to explain. For conditional third and fourth moments I find less persistence compared to variances. However 50-80% of current period skewness and kurtosis is still due to their values in the past period. Conditional kurtosis is highly correlated with the second return moment, and due to that it brings no new explanatory power into the time-series CAPM tests. Overall, time-series regressions of the four-moment CAPM reveal that conditional skewness and kurtosis are priced only spontaneously and their coefficients don't have the expected sign in most of the cases. Higher conditional moments remain important only in theory, while in financial markets return compensation primarily comes for accepting variance (covariance) risks. I also obtain significant intercepts (measures of mispricing), suggesting that other risk factors with substantial explanatory power should exist, which is a motivation for further research.

Testing the four-moment CAPM cross-sectionally (focus on crises periods), I conclude that compared to time-series regressions, here kurtosis risk is priced in most of the years, but its risk premium has the opposite sign than expected (negative risk premium). Investors prefer more kurtosis to less and accept a lower return compensation for portfolios exhibiting excess kurtosis. This may be due to the fact that kurtosis risk is still much underestimated in markets during periods of crises. Skewness risk is still insignificantly priced in cross-sectional CAPM. But altogether in cross-sectional tests, CAPM versions accounting for the third and fourth return moments perform better than their two-moment counterparts.

5.2. Possible Extensions

As mentioned above, significant intercepts in time-series tests of the conditional CAPM suggest that additional risk factors besides covariance should exist to explain better portfolio returns. So, a potential extension of the present study might focus on testing the significance of a different group of factors. Another extension could consider the out-of-sample performance of the presented models for estimating time-varying moments (this essay considered only the in-sample results).

References

Books

Asgharian, H. (2010), *Empirical Finance (Lecture Notes)*, Lund University

Brooks, Chris (2008), *Introductory Econometrics for Finance*, 2.ed. Cambridge University Press

Campbell, J.Y., Lo, A.W., MacKinlay, A.C. (1997), *The Econometrics of Financial Markets*, Princeton University Press, USA

Articles

Bodurtha, J.N., Mark, N.C. (1991), *Testing the CAPM with Time Varying Risk and Return*, Journal of Finance, Vol. 66:4, pp. 1485–1505

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

Bollerslev, T., Engle, R.F., Wooldridge J.M. (1988), *A Capital Asset Pricing Model with Time-Varying Covariances*, Journal of Political Economy, Vol. 96, No. 1, pp. 116-131

Brooks, Ch., Burke, S.P., Persand, G. (2002), *Autoregressive Conditional Kurtosis*, ISMA Centre Discussion Papers in Finance

Dittmar, R.F. (2002), *Nonlinear Pricing Kernels, Kutosis Preferences and Evidence from the Cross-section of Equity Return*, Journal of Finance, Vol. 51:1, pp. 369-403

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

Fama, E.F., French, K.R. (1992), *The Cross-Section of Expected Stock Returns*, Journal of Finance, Vol. 47, No. 2, pp. 427-465

Fang, H., Lai, T.Y. (1997), *Co-Kurtosis and Capital Asset Pricing*, The Financial Review, Vol. 32, pp. 293–307

Friend, J., Westerfield, R. (1980), *Co-Skewness and Capital Asset Pricing*, Journal Finance, pp. 897–913

Fuertes, A., Miffre, J., Tan, W. (2009), *Momentum profits, nonnormality risks and the business cycle*, Applied Financial Economics, Vol. 19:12, pp. 935-953

Harvey, C.R., Siddique, A. (1999), *Autoregressive Conditional Skewness*, Journal of Financial and Quantitative Analysis, Vol. 34, pp. 456–487

Harvey, C.R., Siddique, A. (2000a), *Conditional Skewness in Asset Pricing Tests*, Journal of Finance, Vol. 55, No. 4, pp. 1263–95

Harvey, C. R., Siddique, A. (2000b), *Time Varying Conditional Skewness and the Market Risk Premium*, Research in Banking and Finance, Vol. 1:1

Jondeau, E. and M. Rockinger (2000), *Conditional Volatility, Skewness and Kurtosis: Existence and Persistence*", Working Paper, HEC School of Management

Kraus, A., Litzenberger, R. (1976), *Skewness Preference and the Valuation of Risky Assets*, Journal of Finance, pp. 1085–1094

León, A., Rubio, G., Serna, G. (2004), *Autoregressive conditional volatility, skewness and kurtosis*, Quarterly Journal of Economics and Finance,

Lim, K. G. (1989), *A New Test of the Three-moment Capital Asset Pricing Model*, Journal of Financial and Quantitative Analysis, Vol. 24, pp. 205–216

Lintner, J. (1966), *The Valuation of Risk Assets and Selection of Risky Investments in Stock*

Portfolio and Capital Budgets, Review of Economics and Statistics, Vol. 47:1, pp. 13–47

Markowitz, H.M. (1952), *The utility of wealth*, Journal of Political Economy, Vol. 60, pp. 151–158

Post, Th., van Vliet, P., Levy H. (2008), *Risk Aversion and Skewness Preference*, Journal of Banking and Finance, Vol. 32

Scott R.C., Horvath, P.A. (1980), *On The Direction of Preference for Moments of Higher Order Than The Variance*, The Journal of Finance, Vol. 35, No. 4

Sears, R.S., Wei, K.C.J. (1985), *Asset Pricing, Higher Moments, and the Market Risk Premium: A Note*, Journal of Finance, Vol. 40, pp 1251-1251

Sharpe, W.F. (1965), *Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk*, Journal of Finance, Vol. 19(3), pp. 425-442

Smith, D.R. (2007), *Conditional Coskewness and Asset Prices*, Journal of Empirical Finance, Vol. 14, pp. 97-119

Internet

Kenneth French data library:

http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Appendix

Table 1. Summary Statistics for Excess Portfolio Returns

The table presents main descriptive statistics (mean, maximum, minimum, standard deviation, unconditional sample skewness and kurtosis, and Jarque-Berra test value) for the studied 25 excess US portfolio returns (10 industry portfolios, 5 quintile-based portfolios formed by size and 10 momentum portfolios), along with the market portfolio.

	Mean	Max	Min	Std dev	Skew	Kurt	JB	p-val
Rm	0.46	16.05	-23.14	4.70	-0.57	4.88	99.11	0
Durables	0.46	42.91	-32.97	6.54	0.12	8.20	556.21	0
Energy	0.72	23.33	-19.10	5.61	0.01	4.28	33.75	0
HiTech	0.44	20.02	-26.54	6.97	-0.20	4.01	24.34	0
Health	0.53	29.07	-21.07	5.11	0.07	5.52	131.04	0
Manufacturing	0.54	17.77	-27.92	5.20	-0.52	5.53	153.22	0
Non-durables	0.65	18.15	-21.63	4.51	-0.32	5.04	93.66	0
Shops	0.56	25.22	-28.91	5.52	-0.30	5.21	107.18	0
Telecommunications	0.48	21.98	-15.97	4.92	-0.16	4.17	30.40	0
Utilities	0.46	18.22	-12.94	4.21	-0.13	3.96	20.52	0
Other	0.45	19.65	-24.28	5.49	-0.48	4.84	87.93	0
Lo20	0.62	27.27	-30.23	6.49	-0.23	5.43	125.72	0
Qnt2	0.66	24.44	-29.84	6.10	-0.54	5.06	110.64	0
Qnt3	0.65	22.01	-27.64	5.61	-0.58	5.07	115.31	0
Qnt4	0.62	19.68	-25.80	5.33	-0.51	4.95	98.75	0
Hi20	0.42	17.57	-20.92	4.51	-0.39	4.57	62.88	0
Low	-0.30	45.76	-26.18	8.55	0.70	7.14	391.09	0
P2	0.26	35.65	-25.00	6.61	0.22	5.88	173.81	0
P3	0.43	34.04	-23.45	5.67	0.31	6.50	258.42	0
P4	0.48	21.49	-19.24	5.06	-0.15	4.95	79.89	0
P5	0.37	20.81	-22.09	4.69	-0.29	5.14	100.97	0
P6	0.45	16.05	-24.38	4.76	-0.42	5.35	127.18	0
P7	0.52	18.39	-24.89	4.57	-0.50	5.74	175.14	0
P8	0.68	18.42	-21.06	4.64	-0.34	4.89	82.40	0
P9	0.68	20.87	-26.87	5.03	-0.61	5.84	195.60	0
High	1.00	22.67	-27.34	6.37	-0.45	4.97	96.27	0

Table 2. Engle Test for ARCH Effects

The table shows results for Engle's test of ARCH effects (in column four I assumed k=5 lags, but for the eight portfolios with insignificant test value (are highlighted), I also tried other lag values, and in case of a lower p-value this is shown in the last column together with the corresponding lag value).

	TR^2	$\chi^2(5)$	p-val	p-val (other k)
Rm	15.43	**11.07**	0.009	
durbl	13.46		0.020	
enrgy	17.80		0.003	
Hitec	69.30		0	
hlth	20.83		0.001	
manuf	11.23		0.047	
nodur	16.26		0.006	
shops	9.88		0.080	0.043 (4 lags)
telcm	51.82		0	
utils	40.27		0	
other	31.98		0	
lo20	11.32		0.045	
qnt2	4.56		0.472	
qnt3	4.00		0.549	
qnt4	8.21		0.145	
hi20	22.42		0	
low	145.18		0	
p2	62.81		0	
p3	35.36		0	
p4	45.05		0	
p5	42.18		0	
p6	9.35		0.096	0.020 (1 lag)
p7	6.29		0.279	0.036 (1 lag)
p8	6.418		0.268	
p9	6.385		0.271	0.035 (1 lag)
high	12.60		0.028	

Table 3. Estimation Results for the Three Proposed GARCH Specifications

The tables below show estimation results for the three proposed BEKK models tested on 25 portfolios. By definition these models set restrictions in order to get positive semi-definite covariance matrix; so, instead of the original coefficients mentioned in the essay text, we estimate others, whose multiplication rules are shown below (after choosing the suitable model, I also present the original coefficients mentioned in the essay):

$$\alpha_{0,i} = mu_1 \qquad \alpha_{0,M} = mu_2 \qquad \beta_{0,i} = \omega_1^2 \qquad \beta_{1,i} = \beta_1^2 \qquad \beta_{2,i} = \alpha_1^2$$

$$\beta_{0,M} = \omega_2^2 + \omega_3^2 \qquad \beta_{1,M} = \beta_2^2 \qquad \beta_{2,M} = \alpha_2^2 \qquad \beta_{0,iM} = \omega_1 * \omega_2 \qquad \beta_{1,iM} = \beta_1 * \beta_2 \qquad \beta_{2,iM} = \alpha_1 * \alpha_2$$

All p-values are from a two-sided test and are shown in brackets, insignificant coefficients at 5% are highlighted to follow easier the conclusions. For models 2 and 3, I do not present p-values, but simply highlight insignificant coefficients, because, as shown below, I chose the first model specification.

GARCH 1	mu_1	$\alpha_{1,i}$	$\alpha_{2,i}$	mu_2	$\alpha_{2,M}$	ω_1	ω_2	ω_3	α_1	α_2	β_1	β_2	logL
	Mean eq. (p-values* in brackets)					***Var-Cov eq. (p-values in brackets)***							
Durables	0.4819 (0.076)	-	-	0.5647 (0.006)	-	1.6532 (0)	1.2497 (0)	0.5027 (0)	0.3481 (0)	0.1967 (0)	0.9082 (0)	0.9362 (0)	-2757.60
Energy	0.8071 (0)	-	-	0.5279 (0.007)	-	1.3103 (0)	0.7373 (0)	0.8144 (0)	0.2612 (0)	0.3140 (0)	0.9362 (0)	0.9235 (0)	-2826.64
HiTech	0.5342 (0.045)	-	-	0.5462 (0.004)	-	1.3897 (0)	0.9124 (0)	0.4993 (0)	0.2865 (0)	0.2978 (0)	0.9357 (0)	0.9301 (0)	-2695.58
Health	0.6066 (0.005)	-	-	0.5391 (0.004)	-	1.3917 (0)	0.8239 (0)	0.6004 (0)	0.3327 (0)	0.2855 (0)	0.9063 (0)	0.9349 (0)	-2694.58
Manufacturing	0.5765 (0.009)	-	-	0.5259 (0.007)	-	1.3262 (0)	1.1256 (0)	0.3652 (0)	0.2874 (0)	0.2952 (0)	0.9272 (0)	0.9237 (0)	-2386.48
Non-durables	0.6890 (0)	-	-	0.5282 (0.006)	-	0.9495 (0)	0.8550 (0)	0.4912 (0)	0.3164 (0)	0.3282 (0)	0.9275 (0)	0.9248 (0)	-2519.55
Shops	0.5971 (0.009)	-	-	0.5642 (0.003)	-	1.2940 (0)	0.9490 (0)	0.5430 (0)	0.3129 (0)	0.3110 (0)	0.9255 (0)	0.9223 (0)	-2623.61
Telecommunications	0.6007 (0.002)	-	-	0.5381 (0.004)	-	0.9532 (0)	0.7765 (0)	0.6674 (0)	0.2845 (0)	0.3385 (0)	0.9393 (0)	0.9197 (0)	-2683.53
Utilities	0.5948 (0)	-	-	0.5096 (0.012)	-	0.9878 (0)	0.7041 (0)	0.9241 (0)	0.2715 (0)	0.2826 (0)	0.9291 (0)	0.9286 (0)	-2705.18
Other	0.6744 (0.004)	-	-	0.6088 (0.002)	-	1.6201 (0)	1.3165 (0)	0.4222 (0)	0.3643 (0)	0.3454 (0)	0.8901 (0)	0.8979 (0)	-2406.27
Lo20	0.7869 (0.007)	-	-	0.7059 (0)	-	1.5922 (0)	1.0637 (0)	0.3780 (0.010)	0.2501 (0)	0.3632 (0)	0.9390 (0)	0.9101 (0)	-2738.57
Qnt2	0.8710 (0.002)	-	-	0.6641 (0.001)	-	1.4695 (0)	1.0611 (0)	0.3524 (0)	0.2516 (0)	0.3228 (0)	0.9401 (0)	0.9213 (0)	-2599.12

Qnt3	0.8688 (0.001)	-	-	0.6502 (0.002)	-	1.3387 (0)	1.1138 (0)	0.3205 (0)	0.2486 (0)	0.2963 (0)	0.9410 (0)	0.9251 (0)	-2451.10
Qnt4	0.6833 (0.004)	-	-	0.5410 (0.009)	-	1.0850 (0)	0.9571 (0)	0.1926 (0)	0.3040 (0)	0.3266 (0)	0.9361 (0)	0.9290 (0)	-2244.60
Hi20	0.5268 (0.009)	-	-	0.5699 (0.007)	-	-1.7437 (0)	-1.7309 (0)	-0.2881 (0)	0.3263 (0)	0.3067 (0)	0.8602 (0)	0.8739 (0)	-2006.14
Low	0.0860 (0.776)	-	-	0.7554 (0.001)	-	3.2590 (0)	3.1406 (0)	1.1078 (0)	0.5192 (0)	0.4043 (0)	0.7503 (0)	0.5863 (0)	-2836.23
P2	0.5910 (0.007)	-	-	0.7177 (0)	-	1.7382 (0)	1.2245 (0)	0.5852 (0)	0.4890 (0)	0.3739 (0)	0.8399 (0)	0.8874 (0)	-2657.33
P3	0.6058 (0.001)	-	-	0.6196 (0)	-	1.6341 (0)	1.2771 (0)	0.4188 (0)	0.4978 (0)	0.3813 (0)	0.8310 (0)	0.8865 (0)	-2538.17
P4	0.5398 (0.004)	-	-	0.5502 (0.004)	-	1.5485 (0)	1.5030 (0)	0.4566 (0)	0.4384 (0)	0.3431 (0)	0.8544 (0)	0.8804 (0)	-2431.54
P5	0.5211 (0.006)	-	-	0.6318 (0.001)	-	1.3581 (0)	1.3336 (0)	-0.4342 (0)	0.3969 (0)	0.3453 (0)	0.8797 (0)	0.8953 (0)	-2359.46
P6	0.5297 (0.007)	-	-	0.5439 (0.005)	-	1.0694 (0)	1.0707 (0)	0.3314 (0)	0.3272 (0)	0.3052 (0)	0.9241 (0)	0.9253 (0)	-2344.05
P7	0.6620 (0.001)	-	-	0.6006 (0.002)	-	1.8674 (0)	1.5716 (0)	0.5172 (0)	0.3282 (0)	0.3711 (0)	0.8601 (0)	0.8646 (0)	-2363.53
P8	0.6833 (0)	-	-	0.4949 (0.009)	-	1.5117 (0)	1.2302 (0)	0.3766 (0)	0.3145 (0)	0.3560 (0)	0.8965 (0)	0.8995 (0)	-2365.00
P9	0.5450 (0.007)	-	-	0.4435 (0.007)	-	1.1913 (0)	1.1177 (0)	0.3364 (0)	0.3220 (0)	0.3848 (0)	0.9226 (0)	0.8956 (0)	-2441.77
High	1.0516 (0)	-	-	0.5736 (0.003)	-	2.5869 (0)	1.4410 (0)	0.7162 (0)	0.4351 (0)	0.3978 (0)	0.8148 (0)	0.8595 (0)	-2675.73

GARCH 2	mu_1	$\alpha_{1,i}$	$\alpha_{2,i}$	mu_2	$\alpha_{2,M}$	ω_1	ω_2	ω_3	α_1	α_2	β_1	β_2	logL
	Mean equation					***Var-Cov equation***							
Durables	0.2113	0.4587	-	0.5589	-	1.1041	1.1306	0.9479	0.3693	0.1974	0.9089	0.9268	-2759.15
Energy	0.5350	0.4862	-	0.5294	-	1.0080	0.3625	1.0475	0.2623	0.3141	0.9381	0.9228	-2826.80
HiTech	-0.1107	1.1760	-	0.5448	-	0.6481	0.0697	1.0481	0.2722	0.3163	0.9413	0.9244	-2695.40
Health	0.2043	0.7410	-	0.5255	-	0.8319	0.1462	0.9387	0.3164	0.2878	0.9169	0.9382	-2695.55
Manufacturing	0.0747	0.9502	-	0.5278	-	0.4408	0.4309	1.1058	0.2872	0.2957	0.9273	0.9234	-2386.15
Non-durables	0.1499	1.0135	-	0.5277	-	0.5422	-0.3088	0.8229	0.3285	0.3210	0.9203	0.9326	-2518.86
Shops	-0.0954	1.2340	-	0.5284	-	0.7082	-0.2952	0.9329	0.2940	0.3341	0.9273	0.9240	-2622.78
Telecommunications	0.0674	0.9659	-	0.5250	-	0.6501	-0.3798	1.0147	0.2755	0.3376	0.9422	0.9170	-2682.79
Utilities	0.1024	0.9774	-	0.5045	-	0.8961	-0.7658	1.0537	0.2748	0.2645	0.9350	0.9238	-2703.70

Other	-0.0289	1.1684	-	0.5634	-	0.5728	-0.2690	1.0367	0.4103	0.3049	0.8606	0.9281	-2403.53
Lo20	-0.6554	2.0691	-	0.6334	-	1.3797	-0.9669	0.7206	0.3392	0.2637	0.9097	0.9305	-2737.91
Qnt2	-0.3583	1.8439	-	0.6383	-	0.9622	-1.0135	0.7760	0.3275	0.2528	0.9167	0.9285	-2598.58
Qnt3	-0.0673	1.4243	-	0.6419	-	0.5236	-0.7199	1.0572	0.2995	0.2697	0.9291	0.9240	-2451.57
Qnt4	0.0059	1.2563	-	0.5300	-	0.2606	-0.7345	1.0379	0.3442	0.2582	0.9284	0.9276	-2243.97
Hi20	-0.0274	0.9688	-	0.5187	-	0.4510	-0.3259	0.8294	0.3702	0.3342	0.7390	0.9284	-2002.45
Low	-0.0494	0.9114	-	0.5734	-	1.8545	0.6593	1.2215	0.6297	0.2466	0.7141	0.9213	-2841.46
P2	0.0533	0.7529	-	0.6995	-	0.9569	0.6956	1.1660	0.5827	0.3292	0.8041	0.9003	-2657.50
P3	0.2673	0.5381	-	0.6129	-	1.0133	1.0232	0.7454	0.5554	0.3461	0.8023	0.9030	-2534.56
P4	0.2597	0.4931	-	0.5265	-	0.9087	1.3006	0.8284	0.4871	0.3100	0.8334	0.8911	-2431.16
P5	-0.0161	0.8526	-	0.5608	-	0.5238	0.2077	1.1642	0.4647	0.3093	0.8415	0.9198	-2359.51
P6	-0.0270	1.0290	-	0.5155	-	0.3502	-0.2190	1.0410	0.3263	0.3126	0.9224	0.9266	-2345.18
P7	0.0007	1.0766	-	0.5943	-	0.6777	-0.4124	1.1154	0.4214	0.2989	0.8295	0.9212	-2364.23
P8	0.0363	1.2815	-	0.4759	-	0.5951	-1.0820	1.0760	0.3834	0.2851	0.8907	0.8992	-2367.43
P9	0.2456	0.6726	-	0.4412	-	0.5669	0.9528	0.7598	0.3000	0.4231	0.9332	0.8810	-2445.90
High	0.5353	0.9005	-	0.5440	-	1.5304	0.6022	1.0069	0.4948	0.3487	0.7619	0.9078	-2673.47

GARCH 3	mu_1	$\alpha_{1,i}$	$\alpha_{2,i}$	mu_2	$\alpha_{2,M}$	ω_1	ω_2	ω_3	α_1	α_2	β_1	β_2	logL
	Mean equation					***Var-Cov equation***							
Durables	-0.5829	-	0.0236	0.1514	0.0276	2.0571	2.5874	0.0028	0.3382	0.3921	0.8901	0.7450	-2752.43
Energy	0.7140	-	0.0048	0.1566	0.0197	1.3182	0.7544	0.8332	0.2622	0.3118	0.9357	0.9224	-2825.64
HiTech	0.3117	-	0.0065	0.4452	0.0066	1.4084	0.9290	0.5045	0.2880	0.2988	0.9345	0.9293	-2695.21
Health	0.9164	-	-0.0113	0.2076	0.0159	1.3828	0.8375	0.6126	0.3368	0.2882	0.9060	0.9329	-2691.85
Manufacturing	0.2155	-	0.0158	0.2156	0.0168	1.3470	1.1468	0.3726	0.2899	0.2966	0.9254	0.9218	-2385.70
Non-durables	0.5269	-	0.0104	0.3647	0.0101	0.9548	0.8645	0.4951	0.3179	0.3288	0.9272	0.9239	-2519.28
Shops	0.1985	-	0.0138	0.6283	-0.0019	1.3026	0.9515	0.5426	0.3126	0.3172	0.9246	0.9219	-2621.76
Telecommunications	0.5357	-	0.0036	0.4708	0.0039	0.9552	0.7792	0.6739	0.2855	0.3364	0.9380	0.9198	-2683.51
Utilities	0.6692	-	-0.0037	0.2487	0.0130	0.9918	0.7188	0.9394	0.2712	0.2813	0.9309	0.9278	-2704.72
Other	0.5284	-	0.0071	0.3346	0.0150	1.6255	1.3320	0.4272	0.3648	0.3458	0.8888	0.8952	-2404.86
Lo20	-0.9200	-	0.0428	0.4909	0.0092	1.6372	1.0343	0.3761	0.2671	0.3736	0.9328	0.9092	-2732.81
Qnt2	-0.9741	-	0.0516	0.3544	0.0147	1.4725	1.0494	0.3407	0.2666	0.3376	0.9361	0.9184	-2593.24
Qnt4	-0.2593	-	0.0358	0.1045	0.0215	1.0870	0.9666	0.1867	0.3062	0.3306	0.9351	0.9274	-2234.85
Hi20	0.0024	-	0.0309	-0.2289	0.0411	-1.8831	-1.9828	-0.3441	0.3427	0.3285	0.8396	0.8399	-1997.91
Low	-0.5512	-	0.0162	0.5633	0.0143	3.2250	3.0531	1.1594	0.5079	0.4062	0.7546	0.5985	-2830.73
P2	0.4656	-	0.0068	0.9416	-0.0068	1.6845	1.2041	0.5765	0.4790	0.3682	0.8459	0.8910	-2554.52
P3	0.6077	-	0.0040	1.0213	-0.0149	1.6471	1.3574	0.3949	0.5049	0.3695	0.8255	0.8852	-2534.32
P4	0.0397	-	0.0302	0.2368	0.0217	1.6125	1.5381	0.4623	0.4420	0.3434	0.8465	0.8774	-2425.96

P5	0.3675	-	0.0104	0.4467	0.0111	1.3527	1.3264	-0.4276	0.3943	0.3401	0.8808	0.8973	-2358.97
P6	0.3958	-	0.0073	0.3503	0.0103	1.0738	1.0774	0.3327	0.3272	0.3048	0.9238	0.9249	-2343.72
P7	0.7253	-	0.0003	0.3392	0.0153	1.7676	1.5534	0.5213	0.3404	0.3806	0.8676	0.8634	-2360.02
P8	0.5386	-	0.0106	0.1109	0.0230	1.4273	1.2138	0.3926	0.3102	0.3506	0.9044	0.9014	-2361.68
P9	0.1048	-	0.0240	-0.0177	0.0291	1.2080	1.1306	0.3394	0.3203	0.3818	0.9223	0.8958	-2438.22
High	1.2453	-	-0.0040	0.3196	0.0132	2.5549	1.4471	0.7247	0.4386	0.3991	0.8164	0.8584	-2673.88

Table 4. LR tests

The table shows results from the likelihood ratio test performed to choose between models 1 and 2, since the first specification is nested in the second GARCH.

Portfolio	*Model 1 versus 2*								
	logL*	logL	LR	$\chi^2(1)$		logL*	logL	LR	$\chi^2(1)$
Durables	-2757.60	-2759.15	-3.1		***Low***	-2836.23	-2841.46	-10.46	
Energy	-2826.64	-2826.80	-0.32		***P2***	-2657.33	-2657.50	-0.34	
HiTech	-2695.58	-2695.40	0.36		***P3***	-2538.17	-2534.56	7.22	
Health	-2694.58	-2695.55	-1.94		***P4***	-2431.54	-2431.16	0.76	
Manufacturing	-2386.48	-2386.15	0.66		***P5***	-2359.46	-2359.51	-0.1	
Non-durables	-2519.55	-2518.86	1.38		***P6***	-2344.05	-2345.18	-2.26	
Shops	-2623.61	-2622.78	1.66		***P7***	-2363.53	-2364.23	-1.4	
Telecommunications	-2683.53	-2682.79	1.48		***P8***	-2365.00	-2367.43	-4.86	
Utilities	-2705.18	-2703.70	2.96	3.841	***P9***	-2441.77	-2445.90	-8.26	
Other	-2406.27	-2403.53	5.48		***High***	-2675.73	-2673.47	4.52	3.841
Lo20	-2738.57	-2737.91	1.32						
Qnt2	-2599.12	-2598.58	1.08						
Qnt3	-2451.10	-2451.57	-0.94						
Qnt4	-2244.60	-2243.97	1.26						
Hi20	-2006.14	-2002.45	7.38	3.841					

Table 5. Original GARCH coefficients

They follow directly from the BEKK multiplication rules applied to coefficients of model 1 (insignificant coefficients under 5% are highlighted as before): $\alpha_{0,i} = mu_1$ $\alpha_{0,M} = mu_2$ $\beta_{0,i} = \omega_1^2$ $\beta_{1,i} = \beta_1^2$ $\beta_{2,i} = \alpha_1^2$

$\beta_{0,M} = \omega_2^2 + \omega_3^2$ $\beta_{1,M} = \beta_2^2$ $\beta_{2,M} = \alpha_2^2$ $\beta_{0,iM} = \omega_1 * \omega_2$ $\beta_{1,iM} = \beta_1 * \beta_2$ $\beta_{2,iM} = \alpha_1 * \alpha_2$

	$\alpha_{0,i}$	$\alpha_{0,M}$	$\beta_{0,i}$	$\beta_{0,M}$	$\beta_{0,iM}$	$\beta_{1,i}$	$\beta_{2,i}$	$\beta_{1,M}$	$\beta_{2,M}$	$\beta_{1,iM}$	$\beta_{2,iM}$	$\beta_{1,i} + \beta_{2,i}$	$\beta_{1,M} + \beta_{2,M}$	$\beta_{1,iM} + \beta_{2,iM}$	$\frac{\alpha_{0,i}}{\alpha_{0,M}}$
Durables	0.4819	0.5647	2.7331	1.8145	2.0660	0.8248	0.1212	0.8765	0.0387	0.8503	0.0685	0.9460	0.9152	0.9187	0.85
Energy	0.8071	0.5279	1.7169	1.2069	0.9661	0.8765	0.0682	0.8529	0.0986	0.8646	0.0820	0.9447	0.9514	0.9466	1.53
HiTech	0.5342	0.5462	1.9313	1.0818	1.2680	0.8755	0.0821	0.8651	0.0887	0.8703	0.0853	0.9576	0.9538	0.9556	0.98
Health	0.6066	0.5391	1.9368	1.0393	1.1466	0.8214	0.1107	0.8740	0.0815	0.8473	0.0950	0.9321	0.9555	0.9423	1.13
Manuf	0.5765	0.5259	1.7588	1.4003	1.4928	0.8597	0.0826	0.8532	0.0871	0.8565	0.0848	0.9423	0.9404	0.9413	1.10
Nodurbl	0.6890	0.5282	0.9016	0.9723	0.8118	0.8603	0.1001	0.8553	0.1077	0.8578	0.1038	0.9604	0.9630	0.9616	1.30
Shops	0.5971	0.5642	1.6744	1.1955	1.2280	0.8566	0.0979	0.8506	0.0967	0.8536	0.0973	0.9545	0.9474	0.9509	1.06
Telecom.	0.6007	0.5381	0.9086	1.0484	0.7402	0.8823	0.0809	0.8458	0.1146	0.8639	0.0963	0.9632	0.9604	0.9602	1.12
Utilities	0.5948	0.5096	0.9757	1.3497	0.6955	0.8632	0.0737	0.8623	0.0799	0.8628	0.0767	0.9369	0.9422	0.9395	1.17
Other	0.6744	0.6088	2.6247	1.9114	2.1329	0.7923	0.1327	0.8062	0.1193	0.7992	0.1258	0.9250	0.9255	0.9251	1.11
Lo20	0.7869	0.7059	2.5351	1.2743	1.6936	0.8817	0.0626	0.8283	0.1319	0.8546	0.0908	0.9443	0.9602	0.9454	1.11
Qnt2	0.8710	0.6641	2.1594	1.2501	1.5593	0.8838	0.0633	0.8488	0.1042	0.8661	0.0812	0.9471	0.9530	0.9473	1.31
Qnt3	0.8688	0.6502	1.7921	1.3433	1.4910	0.8855	0.0618	0.8558	0.0878	0.8705	0.0737	0.9473	0.9436	0.9442	1.34
Qnt4	0.6833	0.5410	1.1772	0.9531	1.0385	0.8763	0.0924	0.8630	0.1067	0.8696	0.0993	0.9687	0.9697	0.9689	1.26
Hi20	0.5268	0.5699	3.0405	3.0790	3.0182	0.7399	0.1065	0.7637	0.0941	0.7517	0.1001	0.8464	0.8578	0.8518	0.92
Low	0.0860	0.7554	10.621	11.091	10.235	0.5630	0.2696	0.3437	0.1635	0.4399	0.2099	0.8325	0.5072	0.6498	0.11
P2	0.5910	0.7177	3.0213	1.8419	2.1284	0.7054	0.2391	0.7875	0.1398	0.7453	0.1828	0.9446	0.9273	0.9282	0.82
P3	0.6058	0.6196	2.6703	1.8064	2.0869	0.6906	0.2478	0.7859	0.1454	0.7367	0.1898	0.9384	0.9313	0.9265	0.98
P4	0.5398	0.5502	2.3979	2.4675	2.3274	0.7300	0.1922	0.7751	0.1177	0.7522	0.1504	0.9222	0.8928	0.9026	0.98
P5	0.5211	0.6318	1.8444	1.9670	1.8112	0.7739	0.1575	0.8016	0.1192	0.7876	0.1370	0.9314	0.9208	0.9246	0.82
P6	0.5297	0.5439	1.1436	1.2562	1.1450	0.8540	0.1071	0.8562	0.0931	0.8551	0.0999	0.9610	0.9493	0.9549	0.97
P7	0.6620	0.6006	3.4872	2.7374	2.9348	0.7398	0.1077	0.7475	0.1377	0.7436	0.1218	0.8475	0.8852	0.8654	1.10
P8	0.6833	0.4949	2.2852	1.6552	1.8597	0.8037	0.0989	0.8091	0.1267	0.8064	0.1120	0.9026	0.9358	0.9184	1.38
P9	0.5450	0.4435	1.4192	1.3624	1.3315	0.8512	0.1037	0.8021	0.1481	0.8263	0.1239	0.9549	0.9502	0.9502	1.23
High	1.0516	0.5736	6.6921	2.5894	3.7277	0.6639	0.1893	0.7387	0.1582	0.7003	0.1731	0.8532	0.8970	0.8734	1.83

Table 6. León, Rubio and Serna (2003) model

This model is used to compute series of conditional skewness and kurtosis, whose estimated coefficients are shown in the table, together with p-values in brackets and initial values of parameters used to optimize the likelihood functions. The last 2 columns show average skewness and kurtosis computed for each obtained series, which can be compared to unconditional values from table 1.

	initial skew	initial kurt	γ_2&δ_2	γ_1& δ_1	Logl	γ_0	γ_2	γ_1	δ_0	δ_2	δ_1	γ_1+γ_2	$\delta_1 + \delta_2$	Avg skew	Avg kurt
			Best initial values			***Skewness eq.***			***Kurtosis eq.***			*Stationarity*			
rm	-0.60	3.01	0.7	0.3, 0.4	-1161.11	0.0821 (0)	0.7583 (0)	0.1879 (0)	0.7248 (0)	0.6373 (0)	0.3188 (0)	0.9462	0.9561	-0.13	4.39
durbl	-0.10	2.23	0.7	0.3	-1306.92	0.0435 (0)	0.6492 (0)	0.2762 (0)	0.8481 (0)	0.6778 (0)	0.2545 (0)	0.9254	0.9323	0.20	9.09
enrgy	-0.09	3.78	0.7	0.3	-1253.53	0.0079 (0.409)	0.7166 (0)	0.2275 (0)	0.7501 (0)	0.7256 (0)	0.1808 (0)	0.9441	0.9064	0.02	5.58
hitec	-0.53	2.90	0.8	0.15	-1336.51	0.0553 (0)	0.8297 (0)	0.1269 (0)	0.5078 (0)	0.7948 (0)	0.1061 (0)	0.9566	0.9009	0.15	4.55
hlth	-0.88	3.83	0.7	0.25	-1256.40	0.0016 (0.616)	0.7120 (0)	0.2140 (0)	0.1966 (0)	0.7088 (0)	0.2235 (0)	0.9260	0.9323	0.04	4.94
manuf	-0.47	2.66	0.6	0.3	-1166.40	-0.0905 (0)	0.8043 (0)	0.1549 (0)	0.9502 (0)	0.6059 (0)	0.1681 (0)	0.9592	0.7740	-0.89	4.78
nodur	-0.28	3.42	0.8	0.15	-1146.62	0.0280 (0)	0.7974 (0)	0.1408 (0)	0.3554 (0)	0.7813 (0)	0.1277 (0)	0.9382	0.9090	-0.10	4.58
shops	-0.34	2.80	0.8	0.2	-1257.75	0.0349 (0)	0.7890 (0)	0.1622 (0)	0.4042 (0)	0.7927 (0)	0.1548 (0)	0.9512	0.9475	-0.08	5.82
telcm	0.06	2.71	0.6	0.4	-1170.05	0.0820 (0)	0.5449 (0)	0.3835 (0)	1.1752 (0)	0.5275 (0.046)	0.3873 (0)	0.9284	0.9148	0.03	5.82
utils	0.32	2.18	0.6	0.4	-1078.82	-0.0823 (0)	0.6122 (0)	0.3515 (0)	1.0448 (0)	0.5248 (0.231)	0.3588 (0)	0.9637	0.8836	-0.34	5.19
other	-0.48	2.64	0.6	0.3	-1184.86	0.1397 (0)	0.5305 (0)	0.3024 (0)	0.9867 (0)	0.5796 (0)	0.2077 (0)	0.8329	0.7873	-0.03	4.74
Lo20	0.21	3.09	0.7	0.3	-1375.70	0.0436 (0)	0.6932 (0)	0.2507 (0)	0.7661 (0)	0.7253 (0)	0.1879 (0)	0.9439	0.9132	-0.06	6.47
Qnt2	-0.23	3.27	0.7	0.25	-1371.60	0.1212 (0)	0.7080 (0)	0.2103 (0)	0.3311 (0)	0.7178 (0)	0.1880 (0)	0.9183	0.9058	0.01	4.54
Qnt3	-0.42	2.93	0.6	0.4	-1306.60	-0.0347 (0)	0.6065 (0)	0.3141 (0)	0.5907 (0)	0.5979 (0)	0.2804 (0)	0.9206	0.8783	-0.57	5.01
Qnt4	-0.36	2.70	0.8	0.2	-1282.80	0.0163 (0.103)	0.7865 (0)	0.1825 (0)	0.2692 (0)	0.7866 (0)	0.1847 (0)	0.9690	0.9713	-0.38	5.56

Hi20	-0.63	3.02	0.7	0.25	-1138.30	-0.0451 (0)	0.6936 (0)	0.2682 (0)	0.6298 (0)	0.6620 (0)	0.2582 (0)	0.9618	0.9202	-0.51	5.38
Low	0.39	3.28	0.7	0.3	-1327.26	-0.0274 (0.168)	0.6261 (0)	0.2897 (0)	1.4837 (0)	0.5922 (0)	0.1827 (0)	0.9158	0.7749	0.45	6.81
P2	0.02	2.37	0.8	0.2	-1314.71	-0.0429 (0)	0.8467 (0)	0.1119 (0)	0.6568 (0)	0.7846 (0)	0.1275 (0)	0.9586	0.9121	-0.13	6.50
P3	0.10	2.36	0.8	0.2	-1241.45	0.0179 (0)	0.8443 (0)	0.0758 (0)	0.4351 (0)	0.8132 (0)	0.1186 (0)	0.9201	0.9318	0.25	6.43
P4	-0.15	4.60	0.7	0.3	-1170.63	0.0096 (0.546)	0.6623 (0)	0.2252 (0)	1.0557 (0)	0.6457 (0)	0.2325 (0)	0.8875	0.8782	-0.09	6.29
P5	-0.37	2.64	0.8	0.2	-1176.98	0.0482 (0)	0.7997 (0)	0.1446 (0)	0.3586 (0)	0.7977 (0)	0.1445 (0)	0.9443	0.9422	0.01	5.44
P6	-0.36	2.66	0.6	0.4	-1180.25	0.0172 (0.083)	0.5834 (0)	0.3739 (0)	0.7113 (0)	0.5656 (0)	0.3363 (0)	0.9573	0.9019	-0.35	5.83
P7	-0.51	2.59	0.6	0.4	-1163.73	-0.1205 (0)	0.5140 (0.229)	0.4297 (0)	0.6826 (0)	0.5548 (0)	0.3670 (0)	0.9437	0.9218	-0.71	6.28
P8	-0.46	2.74	0.6	0.4	-1174.02	-0.0524 (0)	0.6753 (0)	0.2826 (0)	0.6676 (0)	0.6343 (0)	0.2889 (0)	0.9579	0.9232	-0.48	5.70
P9	-0.72	3.09	0.7	0.3	-1210.82	0.0148 (0.116)	0.5967 (0)	0.2637 (0)	0.8796 (0)	0.6995 (0)	0.2131 (0)	0.8604	0.9126	-0.38	7.05
High	-0.61	3.52	0.7	0.3	-1321.98	-0.0723 (0)	0.6024 (0)	0.2703 (0)	0.4030 (0)	0.6889 (0)	0.2406 (0)	0.8727	0.9295	-0.50	5.14

Table 7. Correlations between moments

The table shows computed correlation coefficients between series of conditional covariance, (co)skewness and (co)kurtosis for each portfolio.

	durbl	enrgy	hitec	hlth	manuf	nodur	shops	telcm	utils	other	lo20	qnt2	
corr(cov,skew)	0.13	-0.26	-0.53	0.17	-0.56	-0.37	-0.39	-0.09	0.02	-0.39	-0.33	-0.44	
corr(cov,kurt)	0.69	0.69	0.81	0.73	0.65	0.82	0.79	0.57	0.58	0.80	0.68	0.72	
corr(cov,cosk)	0.60	0.65	0.52	0.73	0.72	0.70	0.70	0.41	0.58	0.69	0.66	0.64	
corr(cov,kokurt)	0.72	0.70	0.61	0.76	0.68	0.70	0.73	0.61	0.61	0.74	0.69	0.68	
	qnt3	qnt4	hi20	low	p2	p3	p4	p5	p6	p7	p8	p9	high
corr(cov,skew)	-0.38	-0.48	-0.51	0.32	0.04	0.10	-0.19	-0.38	-0.38	-0.49	-0.40	-0.52	-0.51
corr(cov,kurt)	0.63	0.77	0.83	0.62	0.73	0.66	0.87	0.85	0.62	0.74	0.77	0.73	0.79

corr(cov,cosk)	0.68	0.70	0.82	0.46	0.66	0.63	0.71	0.64	0.62	0.82	0.81	0.78	0.85
corr(cov,cokurt)	0.65	0.69	0.79	0.81	0.78	0.81	0.77	0.78	0.67	0.81	0.76	0.76	0.81

Table 8. Factor Pricing in a Time-Series Regression

The table presents coefficients estimated by regressing excess portfolio returns on the three studied conditional moments (covariance, skewness and kurtosis). For comparison, the last two columns show the regression coefficients with conditional covariance as the only explanatory variable. Insignificant estimates under 10% level are highlighted.

	cov	skew	intercept	kurt	Cov	Intercept
Rm	0.0593 (0.013)	0.2090 (0.071)	-0.8887 (0.125)		0.0340 (0.076)	-0.3301 (0.501)
Durables	0.0761 (0.009)	0.0914 (0.099)	-1.4705 (0.062)		0.0825 (0.005)	-1.6117 (0.040)
Energy	-0.0077 (0.799)	-0.1697 (0.215)	0.8589 (0.145)	-0.1163 (0.010)	0.0020 (0.945)	0.6856 (0.231)
HiTech	0.0434 (0.123)	0.4391 (0.122)	-0.8529 (0.331)		0.0203 (0.395)	-0.1332 (0.858)
Health	0.0367 (0.136)	-0.0931 (0.313)	-0.1536 (0.764)		0.0325 (0.181)	-0.0790 (0.876)
Manufacturing	0.0755 (0.006)	0.2046 (0.089)	-1.0727 (0.095)		0.0494 (0.030)	-0.6342 (0.281)
Non-durables	0.0478 (0.018)	0.1682 (0.180)	-0.2106 (0.615)		0.0378 (0.043)	-0.0438 (0.913)
Shops	0.0802 (0.001)	0.2302 (0.048)	-1.2864 (0.033)		0.0616 (0.005)	-0.8724 (0.123)
Telecommunications	0.0196 (0.399)	0.1441 (0.079)	0.1381 (0.764)		0.0160 (0.491)	0.2056 (0.655)
Utilities	0.0094 (0.756)	0.0612 (0.393)	0.3702 (0.359)		0.0099 (0.742)	0.3428 (0.394)
Other	0.0528 (0.010)	0.2248 (0.027)	-0.8733 (0.126)		0.0353 (0.062)	-0.4384 (0.415)
Lo20	0.0802 (0.003)	0.0980 (0.375)	-1.5469 (0.046)		0.0724 (0.004)	-1.3397 (0.069)

Qnt2	0.0950	0.1694	-1.9096		0.0786	-1.4646
	(0.001)	(0.190)	(0.020)		(0.002)	(0.049)
Qnt3	0.0935	0.1291	-1.6496		0.0767	-1.2957
	(0.001)	(0.126)	(0.030)		(0.004)	(0.074)
Qnt4	0.0692	0.1920	-1.1193		0.0517	-0.7326
	(0.002)	(0.096)	(0.063)		(0.007)	(0.187)
Hi20	0.0501	0.1054	-0.5502		0.0318	-0.2282
	(0.108)	(0.252)	(0.396)		(0.234)	(0.696)
Low	0.0418	0.2506	-1.6724	0.0081	0.0633	-2.2110
	(0.062)	(0.003)	(0.029)	(0.739)	(0.003)	(0.003)
P2	0.0416	0.2677	-0.8402	0.0083	0.0425	-0.9012
	(0.005)	(0.115)	(0.091)	(0.777)	(0.004)	(0.069)
P3	0.0419	0.2533	-0.6539	-0.0002	0.0439	-0.6394
	(0.004)	(0.152)	(0.133)	(0.990)	(0.003)	(0.142)
P4	0.0613	0.1662	-0.8642	-0.0026	0.0562	-0.7673
	(0.001)	(0.135)	(0.055)	(0.920)	(0.001)	(0.086)
P5	0.0563	0.1800	-0.8396	0.0087	0.0458	-0.6145
	(0.003)	(0.137)	(0.064)	(0.726)	(0.008)	(0.150)
P6	0.0694	0.1377	-1.0294		0.0503	-0.6568
	(0.001)	(0.018)	(0.043)		(0.011)	(0.177)
P7	0.0409	0.0167	-0.3170		0.0367	-0.2408
	(0.120)	(0.739)	(0.576)		(0.111)	(0.643)
P8	0.0224	-0.0425	0.1834		0.0275	0.0952
	(0.367)	(0.607)	(0.739)		(0.225)	(0.856)
P9	0.0150	0.0145	0.3400		0.0129	0.3812
	(0.510)	(0.864)	(0.538)		(0.504)	(0.442)
High	0.0019	-0.0200	0.9431		0.0037	0.9035
	(0.937)	(0.881)	(0.165)		(0.856)	(0.149)

Table 9. Cross-Sectional Asset Pricing Tests

The table shows estimation results from cross-sectional tests of the conditional four-moment CAPM (initial columns) and conditional two-moment CAPM (the middle columns). The last three columns illustrate correlations between beta, gamma and delta risk factors. Insignificant estimates at 10% are highlighted.

	β	γ	δ	intercept	R-squared	β	intercept	R-squared	$corr(\beta,\gamma)$	$corr(\beta,\delta)$	$corr(\gamma,\delta)$
2001	0.5571 (0.264)	1.7989 (0.649)	-36.662 (0.071)	-0.4670 (0.268)	19%	0.0810 (0.861)	-0.5837 (0.183)	1%	-0.28	0.43	-0.38
2002	-2.353 (0)	0.1812 (0.774)	-17.861 (0.025)	1.0538 (0.170)	58%	-1.9375 (0)	0.1385 (0.755)	46%	0.50	-0.32	0.06
2003	3.3135 (0)	0.0358 (0.877)	-34.850 (0.020)	-0.3968 (0.574)	59%	2.5199 (0)	0.0558 (0.928)	45%	-0.25	0.49	-0.39
2007	-1.021 (0.182)	-7.3834 (0.029)	-35.544 (0.910)	1.6932 (0.028)	38%	-1.4839 (0.073)	1.5237 (0.071)	13%	0.22	0.32	0.57
2008	-5.412 (0)	2.3582 (0.217)	-0.5906 (0.846)	2.7005 (0.015)	76%	-5.1892 (0)	1.2926 (0.068)	72%	0.21	0.05	-0.62
2009	3.0429 (0.004)	-	-21.260 (0.022)	0.8765 (0.491)	66%	4.4729 (0)	-1.5164 (0.094)	57%	-0.81	-0.62	0.53
2010	2.6141 (0)	0.1484 (0.638)	-	-1.0559 (0.072)	56%	2.6267 (0)	-0.9651 (0.074)	56%	0.05	-0.08	0.84

EViews Codes

```
' GARCH 1  bi-variate BEKK of Engle and Kroner (1995):
' y = mu + res
' res ~ N(0,H)
' H = omega*omega' + beta H(-1) beta' + alpha res(-1) res(-1)' alpha'
' where    y = 2 x 1
'             mu = 2 x 1
'             lambda = 2 x 1
'      H = 2 x 2 (symmetric)
'         H(1,1) = variance of y1   (saved as var_y1)
'         H(1,2) = cov of y1 and y2 (saved as cov_y1y2)
'         H(2,2) = variance of y2   (saved as  var_y2)
'  omega = 2 x 2 low triangular
'  beta = 2 x 2 diagonal
'  alpha = 2 x 2 diagonal

'change path to program path
%path = @runpath
cd %path
' load workfile
load excess.wf1
' input data (dependent variables of both series must be continuous)
smpl @all
series y1 = durbl
series y2 = rm

' set sample for GARCH estimation
sample s0 1970M01 2010M12
```

```
sample s1 1970M02 2010M12

' load data

smpl s0


'get starting values for parameters from univariate GARCH (1,1)

equation eq1.arch(m=100,c=1e-5) y1 c

equation eq2.arch(m=100,c=1e-5) y2 c


' declare coef vectors to use in bi-variate GARCH model (please see introduction for details)

coef(2) mu

 mu(1) = eq1.c(1)

 mu(2)= eq2.c(1)


coef(3) omega

 omega(1)=(eq1.c(2))^.5

 omega(2)=0

 omega(3)=eq2.c(2)^.5


coef(2) alpha

 alpha(1) = (eq1.c(3))^.5

 alpha(2) = (eq2.c(3))^.5


coef(2) beta

 beta(1)= (eq1.c(4))^.5

 beta(2)= (eq2.c(4))^.5


' constant adjustment for log likelihood (i.e. we define 2log(2pi))

!mlog2pi = 2*log(2*@acos(-1))
```

```
'old values
' use var-cov of sample in "s1" as starting value of variance-covariance matrix
series cov_y1y2 = @cov(y1-mu(1), y2-mu(2))
series var_y1 = @var(y1)
series var_y2 = @var(y2)

series res2 = y2-mu(2)
series sqres1 = (y1-mu(1))^2
series sqres2 = (y2-mu(2))^2
series res1res2 = (y1-mu(1))*(y2-mu(2))

' LOG LIKELIHOOD - set up the likelihood
' 1) open a new blank likelihood object (L.O.) name bvgarch
' 2) specify the log likelihood model by append

    ' squared errors and cross errors
 logl bvgarch
bvgarch.append @logl logl
bvgarch.append sqres1 = (y1-mu(1))^2
bvgarch.append sqres2 = (y2-mu(2))^2
bvgarch.append res1res2 = (y1-mu(1))*(y2-mu(2))
bvgarch.append res2 = y2-mu(2)

' calculate the variance and covariance series
bvgarch.append var_y1  =  omega(1)^2 + beta(1)^2*var_y1(-1) + alpha(1)^2*sqres1(-1)
bvgarch.append cov_y1y2 = omega(1)*omega(2) + beta(2)*beta(1)*cov_y1y2(-1) +
alpha(2)*alpha(1)*res1res2(-1)
bvgarch.append var_y2  = omega(3)^2 + omega(2)^2 + beta(2)^2*var_y2(-1) + alpha(2)^2*sqres2(-1)
```

```
' determinant of the variance-covariance matrix
bvgarch.append deth = var_y1*var_y2 - cov_y1y2^2

' inverse elements of the variance-covariance matrix
bvgarch.append invh1 = var_y2/deth
bvgarch.append invh2 = -cov_y1y2/deth
bvgarch.append invh3 = var_y1/deth

' log-likelihood series
bvgarch.append logl =-0.5*(!mlog2pi + (invh1*sqres1+2*invh2*res1res2+invh3*sqres2) + log(deth))

' remove some of the intermediary series
bvgarch.append @temp invh1 invh2 invh3 sqres1 sqres2 deth

' estimate the model
smpl s1
bvgarch.ml(showopts, m=500, c=1e-5)

' change below to display different output
show bvgarch.output
graph varcov.line var_y1 var_y2 cov_y1y2
show varcov
```

GARCH 3:

```
'  y = mu + res -> y = H*lambda + mu + res
'  res ~ N(0,H)
'  H = omega*omega' + beta H(-1) beta' + alpha res(-1) res(-1)' alpha'
' where       y = 2 x 1
'             mu = 2 x 1
```

```
'                    lambda = 2 x 1
'       H = 2 x 2 (symmetric)
'           H(1,1) = variance of y1   (saved as var_y1)
'           H(1,2) = cov of y1 and y2 (saved as cov_y1y2)
'           H(2,2) = variance of y2   (saved as  var_y2)
'  omega = 2 x 2 low triangular
'  beta = 2 x 2 diagonal
'  alpha = 2 x 2 diagonal

'change path to program path
%path = @runpath
cd %path
' load workfile
load excess.wf1
' input data (dependent variables of both series must be continuous)
smpl @all
series y1 = qnt3
series y2 = rm
' set sample for GARCH estimation (not the whole series > leave some observations for forecasting)
sample s0 1970M01 2010M12
sample s1 1970M02 2010M12
' load data
smpl s0

'get starting values for parameters from univariate GARCH-M (1,1); archm=var shows the inclusion of var in the mean eq
equation eq1.arch(archm=var,m=100,c=1e-5) y1 c
equation eq2.arch(archm=var,m=100,c=1e-5) y2 c
'save the conditional variances
```

```
eq1.makegarch garch1
eq2.makegarch garch2
' declare coef vectors to use in bi-variate GARCH model (please see introduction for details)
    coef(2) lambda
    lambda(1) = eq1.c(1)
    lambda(2) = eq2.c(1)

    coef(2) mu
    mu(1) = eq1.c(2)
    mu(2)= eq2.c(2)

    coef(3) omega
    omega(1)=(eq1.c(3))^.5
    omega(2)=0     ' because we don't have it in the univariate GARCH-M
    omega(3)=(eq2.c(3))^.5

    coef(2) alpha
    alpha(1) = (eq1.c(4))^.5
    alpha(2) = (eq2.c(4))^.5

    coef(2) beta
    beta(1)= (eq1.c(5))^.5
    beta(2)= (eq2.c(5))^.5

' constant adjustment for log likelihood (i.e. we define 2log(2pi))
!mlog2pi = 2*log(2*@acos(-1))
'old values
' use var-cov of sample in "s1" as starting value of variance-covariance matrix
'series cov_y1y2 = @cov(y1-mu(1), y2-mu(2))
```

```
'series var_y1 = @var(y1)
'series var_y2 = @var(y2)
'series sqres1 = (y1-mu(1))^2
'series sqres2 = (y2-mu(2))^2
'series res1res2 = (y1-mu(1))*(y2-mu(2))

    series cov_y1y2 = @cov(y1-mu(1)-lambda(1)*garch1, y2-mu(2)-lambda(2)*garch2)
    series var_y1 = @var(y1-lambda(1)*garch1)
    series var_y2 = @var(y2-lambda(2)*garch2)
    series sqres1 = (y1-mu(1)-lambda(1)*garch1)^2
    series sqres2 = (y2-mu(2)-lambda(2)*garch2)^2
    series res1res2 = (y1-mu(1)-lambda(1)*garch1)*(y2-mu(2)-lambda(2)*garch2)

' LOG LIKELIHOOD
logl bvgarch
'old values
'bvgarch.append @logl logl
'bvgarch.append sqres1 = (y1-mu(1))^2
'bvgarch.append sqres2 = (y2-mu(2))^2
'bvgarch.append res1res2 = (y1-mu(1))*(y2-mu(2))

    ' squared errors and cross errors
    bvgarch.append @logl logl
    bvgarch.append sqres1 = (y1-mu(1)-lambda(1)*garch1)^2
    bvgarch.append sqres2 = (y2-mu(2)-lambda(2)*garch2)^2
    bvgarch.append res1res2 = (y1-mu(1)-lambda(1)*garch1)*(y2-mu(2)-lambda(2)*garch2)

' calculate the variance and covariance series
bvgarch.append var_y1  =  omega(1)^2 + beta(1)^2*var_y1(-1) + alpha(1)^2*sqres1(-1)
```

```
bvgarch.append var_y2  = omega(3)^2 + omega(2)^2 + beta(2)^2*var_y2(-1) + alpha(2)^2*sqres2(-1)

bvgarch.append cov_y1y2 = omega(1)*omega(2) + beta(2)*beta(1)*cov_y1y2(-1) +
alpha(2)*alpha(1)*res1res2(-1)


' determinant of the variance-covariance matrix

bvgarch.append deth = var_y1*var_y2 - cov_y1y2^2


' inverse elements of the variance-covariance matrix

bvgarch.append invh1 = var_y2/deth

bvgarch.append invh3 = var_y1/deth

bvgarch.append invh2 = -cov_y1y2/deth


' log-likelihood series

bvgarch.append logl =-0.5*(!mlog2pi + (invh1*sqres1+2*invh2*res1res2+invh3*sqres2) + log(deth))


' remove some of the intermediary series

bvgarch.append @temp invh1 invh2 invh3 sqres1 sqres2 res1res2 deth


' estimate the model

smpl s1

bvgarch.ml(showopts, m=100, c=1e-5)
```

León, Rubio and Serna (2003) model

```
'  r = m + res

'  res ~ N(0,h)

'  h = a1 + b1*h(-1) + last1*res(-1)*res1(-1)'    (variance)

'  s = a2 + b2*s(-1) + last2*e3(-1)                (skew)

'  k = a3 + b3*k(-1) + last3*e4(-1)                (kurt)

' where        r = 1 x 1            m = 1 x 1

'              a = 3 x 1            b = 3 x 1           last = 3 x 1
```

```
' input data (dependent variables of both series must be continuous)
smpl s0
series r = p7

' declare coef vectors to use in bi-variate GARCH model (please see introduction for details)
coef(1) m
 m(1) = mu(1)

coef(3) a
 a(1)= 0
 a(2)= 0
 a(3)= 0

coef(3) b
 b(1) = beta(1)^2
 b(2) = -log(1/0.6 -1)
 b(3) = -log(1/0.6 -1)

coef(3) last
 last(1) = alpha(1)^2
 last(2) = -log(1/0.4 -1)
 last(3) = -log(1/0.4 -1)

'set initial skew and kurt to their unconditional values
series h = @var(r)
series e=(res1-@mean(res1))/@stdev(res1)
series s = -0.51
series k = 2.59
```

```
' LOG LIKELIHOOD
' squared errors and cross errors
 logl lastgarch
lastgarch.append @logl logl
'lastgarch.append res = r-m(1)
lastgarch.append e=(res1-@mean(res1))/@stdev(res1)

' calculate the variance and covariance series
‘lastgarch.append h  = a(1) + b(1)*h(-1) + last(1)*(e(-1)^2)*h(-1)
lastgarch.append s  = a(2) + @logit(b(2))*s(-1) + @logit(last(2))*(e(-1)^3)
lastgarch.append k  = a(3) + @logit(b(3))*k(-1) + @logit(last(3))*(e(-1)^4)
lastgarch.append x = 1+s/6*(e^3-3*e)+(k-3)/24*(e^4-6*(e^2)+3)
lastgarch.append y = x^2
lastgarch.append gama = 1+ s^2/6 +(k-3)^2/24

' log-likelihood series
lastgarch.append logl =-0.5*log(var_y1) - 0.5*(e^2) - log(gama) + log(y)

' remove some of the intermediary series
lastgarch.append @temp res e x y gama

' estimate the model
smpl s1
lastgarch.ml(showopts, m=500, c=1e-7)
genr b2=@logit(b(2))
genr b3=@logit(b(3))
genr last2=@logit(last(2))
genr last3=@logit(last(3))
```