TEDAS - Tail Event Driven ASset Allocation

Wolfgang Karl Härdle*
Sergey Nasekin*
David Lee Kuo Chuen**
Phoon Kok Fai**

BERLIN

ECONOMIC RISK

SFB 649

* Humboldt-Universität zu Berlin, Germany
** Singapore Management University, Singapore

TEDAS - Tail Event Driven ASset Allocation [*]

Wolfgang Karl Härdle, [†]   Sergey Nasekin, [‡]
David Lee Kuo Chuen, Phoon Kok Fai [§]

June 13, 2014

Abstract

Portfolio selection and risk management are very actively studied topics in quantitative finance and applied statistics. They are closely related to the dependency structure of portfolio assets or risk factors. The correlation structure across assets and opposite tail movements are essential to the asset allocation problem, since they determine the level of risk in a position. Correlation alone is not informative on the distributional details of the assets. By introducing TEDAS -Tail Event Driven ASset allocation, one studies the dependence between assets at different quantiles. In a hedging exercise, TEDAS uses adaptive Lasso based quantile regression in order to determine an active set of negative non-zero coefficients. Based on these active risk factors, an adjustment for intertemporal correlation is made. Finally, the asset allocation weights are determined via a Cornish-Fisher Value-at-Risk optimization. TEDAS is studied in simulation and a practical utility-based example using hedge fund indices.

Key words: portfolio optimization, asset allocation, adaptive lasso, quantile regression, value-at-risk

JEL Classification: C00, C14, C50, C58

[*]This research was supported by the Deutsche Forschungsgemeinschaft through IRTG 1792, SFB 649 and Lee Kong Chian School of Business, Sim Kee Boon Institute for Financial Economics, Singapore Management University
[†]C.A.S.E.- Center for Applied Statistics & Economics, Humboldt-Universität zu Berlin, Berlin, Germany. Lee Kong Chian School of Business, Singapore Management University, Singapore
[‡]C.A.S.E.- Center for Applied Statistics & Economics, Humboldt-Universität zu Berlin, Berlin, Germany
[§]Lee Kong Chian School of Business, Singapore Management University, Singapore

1   Introduction

Portfolio selection and risk management are important concepts in quantitative finance and applied statistics. Their applications deal with the estimation of correlation structure of portfolio assets or risk factors. The correlation structure, or, more generally, the dependence across assets is a main component of the portfolio allocation problem since it determines the level of risk in the investment position. However, the correlation is not informative on the distributional details of the portfolio. It does not specify the dependence between assets at different quantiles, but refers to relations with respect to their mean values, which may be weak, while relations or even dependence in tails or, more broadly, quantiles or expectiles, may be significant. Indeed, assets that have negative correlation when the markets are stable, may exhibit positive correlation during volatile periods. Modelling tail dependence is therefore a more informative and flexible approach to hedging and portfolio allocation.

The question of asset choice is crucial, since securities exhibiting low or even negative correlation, are more preferable, as they tend to decrease the overall portfolio risk. In earlier, such as Lintner (1983) and more recent literature, as Cvitanić et al. (2003), Favre and Galeano (2002), Giamouridis and Vrontos (2007), Lhabitant and Learned (2002), McFall Lamm (1999), McFall Lamm (2003) and in many other sources, the use of hedge funds as portfolio assets along with conventional securities such as stocks or bonds has been advocated because they provide superior risk-adjusted returns over the conventional assets ("beta benefit" and "alpha benefit") due to their dynamic nature, non-equity related strategies and other features.

Given the hedge fund alternative, the question arises how to effectively select from thousands of hedge funds. Statistically speaking, the number of covariates p is larger than the number of observations n, which may lead to estimation problems, such as, for instance, multicollinearity. The choice of the allocation procedure is also crucial, since it is necessary to account for the distributional properties of the portfolio and calculate the risk of the position accordingly. It has become an established fact that most hedge fund strategies exhibit asymmetric return patterns characterized by negative skew and excess kurtosis due to using leverage and financial derivatives. Therefore a successful portfolio allocation and risk measurement procedure should be able to match higher moments of the portfolio distribution such as skewness and kurtosis.

Another important point is the modelling of time-varying variance-covariance structure of the portfolio. It is a well-established fact that financial time series exhibit volatility clustering, when large changes tend to follow large changes and vice versa. The so-called "leverage effect" refers to the relationship between asset returns and both implied and realized volatility: volatility increases when the asset price falls. Correlations also tend to be unstable and change in time. Therefore a suitable model is necessary to address the issue of changing volatility and correlation structure while

being computationally feasible.

To deal with the issues of portfolio allocation and risk measurement indicated above, a framework is offered in this study which includes the estimation of quantile dependence between assets in the case of high dimensionality with $p > n$, an alternative distribution-based asset allocation procedure and models time-varying variance-covariance structure of assets applied to the universe of hedge funds. The Adaptive L_1 (LASSO - Least Absolute Shrinkage and Selection Operator) penalized quantile regression is used to simultaneously pursue variable selection and measure causal relations between variables at tail quantiles. Hedge funds are shown to exhibit superior performance as hedging assets for such tail events. For the asset universe with reduced dimension, a portfolio allocation procedure is implemented using the Value-at-Risk (VaR) as a portfolio risk measure which is minimized to obtain optimal asset weights. The VaR is adjusted via the Cornish-Fisher quantile expansion which allows to optimize over higher moments of the portfolio return distribution, so that the effect of "fat tails" is captured. Time-varying volatility and correlation structure of the portfolio is modeled with multivariate general autoregressive conditional heteroscedasticity (GARCH) such as the Dynamic Conditional Correlation (DCC) and the orthogonal GARCH models to obtain a further improvement of risk management process.

The present study is structured as follows. The first section motivates the use of hedge funds indices as a proxy for hedge funds and introduces the portfolio risk measure based on the adjusted Cornish-Fisher portfolio Value-at-Risk and time-varying variance-covariance structure. The second section outlines the non-positive Lasso selection and shrinkage method and quantile regression as well as their joint implementation - the Adaptive Lasso quantile regression estimator. The third section provides a Monte-Carlo simulation analysis and concludes with an empirical application designed as a "Tail Event Driven ASset allocation (TEDAS) strategy" with a discussion of its risk-return and utility characteristics in comparison to alternative strategies.

2   Portfolio Management and Hedge Funds

2.1   Asset Allocation Problem

Many portfolio managers rely on the Markowitz (mean-variance or risk-return) rule which combines assets into an "efficient" portfolio offering risk-adjusted target returns. Risk-return optimization is based on four inputs: the weights of total funds invested in each security w_i, $i = 1, \ldots, d$, the expected returns μ approximated as averages $\bar{r}$, volatilities (standard deviations) σ_i associated with each security and covariances σ_{ij}, $j = 1, \ldots, d; i \neq j$ between returns. Portfolio weights w_i are obtained from the

quadratic optimization problem, see Brandimarte (2006), p. 74

$$
\begin{aligned}
\operatorname*{minimize}_{w \in \mathbb{R}^d} \quad & \sigma_p^2(w) \overset{\mathrm{def}}{=} w^\top \Sigma w \\
\text{subject to} \quad & w^\top \mu = r_T, \\
& \sum_{i=1}^{d} w_i = 1, \\
& w_i \geq 0
\end{aligned}
\tag{1}
$$

where $\Sigma \in \mathbb{R}^{d \times d}$ is the covariance matrix for d portfolio asset returns, r_T is the "target" return for the portfolio assigned by the investor. The Markowitz rule simultaneously solves two problems: diversification and asset allocation. Diversification reduces specific risk; asset allocation allows to combine assets so that the portfolio risk can be lowered while the expected returns are not necessarily reduced. The exact shape of the curve of possible allocations depends on correlations between assets: the smaller the correlation, the smaller the risk of the portfolio. Therefore one prefers to find assets that offer an acceptable return while being less than perfectly correlated or even negatively correlated.

The potential benefits of including managed funds into asset portfolios were observed several decades ago. Lintner (1983) stated that "the improvements from holding an efficiently-selected portfolio of managed accounts or funds are so large that the return-risk tradeoffs provided by augmented portfolios clearly dominate the tradeoffs available from a portfolio of stocks alone or from portfolios of stocks and bonds". Recent results show that these findings are quite robust: Table 1 shows correlation results obtained for the conventional assets and hedge funds' indices in the period from 2000 to 2012 (monthly data). The source for the hedge funds indices returns' data is the Eurekahedge provider. The data on MSCI Country and Regional Indices are taken from Morgan Stanley Capital Index (MSCI).

The correlations between traditional equity markets are large and positive which diminishes the possible diversification benefits and confirms the world equity markets' trend towards greater global integration. The situation with hedge funds' indices is different. An important reason is that hedge funds are dynamically rebalanced portfolios unlike static assets such as stocks or bonds. Another reason is that most strategies trade essentially in non-equity-related spreads. Because of the strategies' diversity, hedge fund returns generally display moderate to low correlation with traditional equity and bond indices. In addition, hedge fund strategies have low correlations with each other which makes the idea of diversifying among loosely correlated funds natural. Consequently, many funds' strategies offer good opportunities for diversification. Table 1 shows that there might be good opportunities for diversification coming from CTA/Managed Futures, Global Macro, Asia Macro and Distressed Debt strategies. Fixed Income strategies show a high level of correlation with stock indices, which can be explained by a systemic rise in cross-asset correlation due to crisis events in the

global economy in recent years.

Table 1: Correlation statistics for MSCI and hedge funds' indices returns

Hedge Fund Indices	MSCI Indices								
	WRD	EUR	US	UK	FR	SW	GER	JAP	PAC
Asia CTA	−0.01	0.02	−0.02	−0.06	0.01	−0.09	0.04	−0.03	0.02
Asia Distressed Debt	0.30	0.30	0.24	0.31	0.31	0.26	0.27	0.26	0.34
Asia Macro	−0.01	−0.01	−0.04	0.01	−0.02	0.07	−0.03	0.06	0.06
Global CTA FoF	0.02	0.08	−0.08	0.09	0.10	0.09	0.07	0.06	0.10
Global Event Driven FoF	0.65	0.59	0.58	0.66	0.59	0.50	0.57	0.47	0.67
Global Macro FoF	0.19	0.22	0.07	0.24	0.22	0.18	0.20	0.23	0.31
CTA/Managed Futures	−0.04	0.02	−0.13	0.03	0.03	0.07	−0.01	0.04	0.05
Event Driven	0.82	0.75	0.75	0.78	0.75	0.64	0.75	0.62	0.83
Fixed Income	0.70	0.65	0.63	0.70	0.65	0.56	0.62	0.51	0.78
Long Short Equities	0.82	0.78	0.74	0.76	0.77	0.64	0.77	0.64	0.82
Asia inc Japan Distr. Debt	0.30	0.30	0.24	0.31	0.31	0.26	0.27	0.26	0.34
Asia inc Japan Macro	0.34	0.33	0.31	0.27	0.33	0.24	0.35	0.31	0.40

Calculations based on monthly data Jan. 2000 - Jul. 2012

WRD - World, EUR - Eurozone, FR - France, SW - Switzerland, PAC - Pacific ex. Japan

FoF means "fund of funds"

It has become an established fact that various hedge fund strategies exhibit asymmetric return patterns characterized by negative skew and excess kurtosis. For instance, as is noted by McFall Lamm (2003), the data for EAI, HFR and CSFB's Tremont hedge fund indices demonstrate a significant departure from normality. The results of Lhabitant and Learned (2002) show that skewness variations are not uniform across styles. Skewness and kurtosis phenomena for hedge funds portfolios may occur, inter alia, due to the following reasons, as noted by Lhabitant (2002):

- hedge fund strategies are often based on financial derivatives, and use other dynamic strategies whose returns are not normally distributed and exhibit skewness;

- hedge funds tend to use leverage to magnify returns, which results in frequent price jumps and leptokurtic returns' distributions.

Especially returns for distressed debt, fixed income, and merger arbitrage strategies have asymmetric distributions and exhibit significant non-Gaussian behaviour. The study of recent data on hedge funds and conventional assets' indices returns and risk statistics generally confirms these findings, as shown in Table 2. The estimation results also demonstrate that hedge funds on the whole yield risk-adjusted returns superior to those of stock indices.

Table 2: Returns and risk characteristics for hedge fund and MSCI indices

Hedge Fund/ MSCI Index	Return and Risk Measures					
	Aver. Ret.	Stand. Dev.	Skew	Kurt.	Sharpe Ratio	Value- -at-Risk
Asia CTA	0.12	0.24	3.37	11.97	0.37	0.25
Asia Distressed Debt	0.12	0.06	0.82	1.65	1.58	−0.02
Asia Macro	0.10	0.22	−2.66	8.63	0.34	0.22
Global CTA FoF	0.06	0.08	−0.81	0.22	0.42	0.07
Global Distr. Debt FoF	0.04	0.05	−2.30	6.30	0.28	0.04
Global Event Driven FoF	0.05	0.05	−1.78	4.83	0.41	0.03
Global Macro FoF	0.06	0.05	0.04	0.74	0.73	0.02
CTA/Managed Futures	0.11	0.08	−0.28	−0.27	1.13	0.01
Event Driven	0.10	0.07	−0.57	2.02	1.01	0.02
Fixed Income	0.08	0.04	1.21	4.34	1.38	−0.02
Long Short Equities	0.09	0.08	−0.67	1.64	0.75	0.04
Asia inc Japan Distr. Debt	0.12	0.06	0.82	1.65	1.58	−0.02
Asia inc Japan Macro	0.11	0.22	2.04	4.18	0.38	0.22
MSCI World	−0.01	0.17	−0.37	−0.96	−0.22	0.25
MSCI Eurozone	−0.04	0.24	−0.46	−0.41	−0.26	0.35
MSCI US	−0.01	0.16	−0.48	−0.90	−0.22	0.24
MSCI UK	−0.02	0.18	−0.09	−0.95	−0.24	0.27
MSCI France	−0.02	0.23	−0.34	−0.42	−0.22	0.33
MSCI Switzerland	0.03	0.17	−0.60	−0.07	0.00	0.23
MSCI Germany	−0.01	0.27	−0.68	0.09	−0.16	0.37
MSCI Japan	−0.04	0.18	0.00	−0.84	−0.40	0.29
MSCI Pacific ex. Japan	0.04	0.22	−0.36	−0.83	0.07	0.28

Data sources are Eurekahedge and MSCI; based on monthly data Jan. 2000 - Jul. 2012

all measures are annualized except for VaR (calculated monthly)

VaR calculated at 0.05% confidence level via log-normal approximation (see Dowd (2005))

2.1.1   Efficient Frontier Analysis

Given the potential diversification benefits of hedge funds, one must determine an allocation policy. Some investors adopt a pragmatic attitude and recommend allocating an arbitrarily pre-specified percentage of portfolios to hedge funds, for instance, 1%, 2% or 5%. Such an approach is quite common among practitioners. However, the findings of Lhabitant and Learned (2002) confirm that smart diversification outperforms naive one in terms of risk reduction.

In the study of Cvitanić et al. (2003), a utility-based model is proposed in which a non-myopic investor with incomplete information allocates wealth between a risk-free security, a passive portfolio (conventional asset) and an actively managed (hedge fund) portfolio based on the changes in the value of the expected alphas in the CAPM framework. The results imply that hedge funds have a low beta with respect to traditional stock and bond indexes while also having a so-called *alpha benefit*, that is, providing an abnormal return adjusted by risk in the CAPM framework. In McFall Lamm (1999) several reasons related to market inefficiency arguments are given to motivate superior risk-adjusted returns for hedge funds. First, there is a lack of transparency in hedge fund markets. Second, a high-return niche is created because a large pool of investable funds is effectively barred from moving into the industry. Finally, hedge funds as se-

curities are not as liquid as other financial products, using monthly or even quarterly redemptions. Because of this, as noted by Lhabitant (2002), hedge funds capture a long-term liquidity premium that increases their expected return.

The enhancement effect on the risk-return trade-off through allocating portfolio shares to hedge funds can be demonstrated with the so-called *efficient frontier*. The efficient frontier is constructed by solving the portfolio optimization problem (1) for different target return constraints r_T and then plotting them against the corresponding portfolio variance $w^\top \Sigma w$ values. As a result, a set of *optimal portfolios* which offer the highest possible expected return for a defined level of risk or the lowest risk for a given level of expected return is obtained. Obviously, portfolios that lie below the efficient frontier are *sub-optimal*: they do not provide enough return for the given level of risk.

As a demonstration that introducing hedge funds at the asset allocation level may give large potential benefits in a risk-return sense, in Lhabitant (2002) three efficient frontiers are compared, the first of which is constructed exclusively from portfolios of equity indices without hedge funds, the second with an imposed 5% cap on the hedge fund allocation and the third - without limits on hedge funds; it is found that the last frontier dominates the other two. Portfolios made solely of stocks are sub-optimal compared to those which include hedge funds.

The numerical results of generating efficient frontiers for portfolios including both stocks and hedge funds (the resulting allocations and risk-return profiles) are illustrated in Table 3. These results show that for the given risk and return range, between

Table 3: Efficient Frontier for the Portfolio of Hedge Funds and S&P 500 Stocks (%)

Asset Name	Portfolios on the Efficient Frontier												
	Pt.1	Pt.2	Pt.3	Pt.4	Pt.5	Pt.6	Pt.7	Pt.8	Pt.9	Pt.10	Pt.11	Pt.12	Pt.13
Portfolio Return	6.60	7.66	8.72	9.78	10.84	11.90	12.96	14.02	15.08	16.14	17.20	18.25	19.31
Portfolio Risk	2.61	2.64	2.76	3.00	3.43	4.27	6.08	8.93	12.15	15.53	18.99	22.50	26.02
	Portfolio Allocation												
Asia Arbitrage	2	2	0	0	0	0	0	0	0	0	0	0	0
Asia CTA	1	0	0	0	1	1	2	4	5	7	8	9	11
Asia Distressed Debt	7	11	17	23	31	44	48	50	50	45	40	35	28
Asia Macro	3	3	3	3	4	3	0	0	0	0	0	0	0
Global Macro FoF	16	2	0	0	0	0	0	0	0	0	0	0	0
Arbitrage	49	48	43	26	0	0	0	0	0	0	0	0	0
CTA/Managed Futures	1	6	9	14	19	22	16	8	1	0	0	0	0
Distressed Debt	0	0	0	0	0	2	4	0	0	0	0	0	0
Macro	5	14	14	7	0	0	0	0	0	0	0	0	0
Relative Value	0	0	3	17	36	10	0	0	0	0	0	0	0
Asia inc Japan Rel. Value	0	0	0	0	1	13	17	17	15	11	7	2	0
S&P Stocks (total weight)	*16*	*14*	*11*	*8*	*8*	*7*	*13*	*21*	*29*	*37*	*45*	*53*	*62*

38% and 93% of portfolio wealth is allocated to hedge funds. The fact that more weight is allocated to conventional stocks as the level of risk increases, confirms that hedge funds are in fact conservative securities. The findings about the hedge funds' strategies which are selected most frequently, such as Asia CTA, Asia Distressed Debt or Managed Futures, are consistent with the results obtained before in the correlation study in Table 1.

Thus hedge funds seem to be good candidates for diversification: they can substitute for bonds and cash as a defensive vehicle when equity prices decline. But assessing hedge funds' returns based on return and volatility criteria only may be misleading because of the potential underestimation of tail events due to skewness and kurtosis effects. Therefore one incorporates asymmetric return distributions when constructing hedge fund portfolios to minimize downside risk.

2.1.2 Alternative Portfolio Optimization Methods

Standard Markowitz rule is based on finding a tradeoff between risk and return minimizing portfolio variance as a risk measure. Another popular risk measure, Value-at-Risk (VaR), is based on tail properties of portfolio loss; it measures the maximum portfolio loss given confidence level α. VaR is a quantile of the probability distribution of future wealth, see Franke et al. (2011):

$$q_{\alpha,t} = F_{t+1}^{-1}(\alpha) \stackrel{\text{def}}{=} \inf\{x; F_{t+1}(x) \geq \alpha\}. \tag{2}$$

where F_{t+1} denotes the P&L distribution function.

A modification of VaR via the Cornish-Fisher (CF) expansion improves its precision adjusting estimated quantiles for non-normality. As discussed by Favre and Galeano (2002), VaR based only on volatility, underestimates portfolio risk. The CF expansion, see Abramowitz and Stegun (1965), approximates the quantile, e.g., the VaR (2), of an arbitrary random variable Y with mean μ, variance σ^2 and cdf F_Y via a standard normal variate $z_\alpha \stackrel{\text{def}}{=} \Phi^{-1}(\alpha)$ and higher moments. Let y_α be α-quantile $F_Y(y_\alpha) = \alpha$, then the CF approximation yields:

$$y_\alpha \simeq \mu - \sigma q_\alpha,$$

where

$$\begin{aligned}
q_\alpha = z_\alpha &+ \{\gamma_1 h_1(z_\alpha)\} \\
&+ \{\gamma_2 h_2(z_\alpha) + \gamma_1^2 h_{11}(z_\alpha)\} \\
&+ \{\gamma_3 h_3(z_\alpha) + \gamma_1 \gamma_2 h_{12}(z_\alpha) + \gamma_1^3 h_{111}(z_\alpha)\} + \dots
\end{aligned}$$

approximates q_α in (2); $\gamma_{r-2} = \kappa_r/\kappa_2^{r/2}$, $r = 3, 4, \dots$; κ_r are rth cumulants of the distribution of Y, $h_1(x) = \frac{1}{6} He_2(x)$, $h_2(x) = \frac{1}{24} He_3(x)$, $h_3(x) = \frac{1}{120} He_4(x)$, $h_{11}(x) = -\frac{1}{36}\{2He_3(x) + He_1(x)\}$, $h_{12}(x) = -\frac{1}{24}\{He_4(x) + He_2(x)\}$, $h_{111}(x) = \frac{1}{324}\{12He_4(x) + 19He_2(x)\}$ where $He_n(x)$ are Hermite polynomials of order n. For the CF-VaR expansion we consider cumulants up to the 4th order:

$$q_\alpha = z_\alpha + (z_\alpha^2 - 1)\frac{S}{6} + (z_\alpha^3 - 3z_\alpha)\frac{K}{24} - (2z_\alpha^3 - 5z_\alpha)\frac{S^2}{36},$$

where S is the skewness and K is kurtosis.

To incorporate asymmetry explicitly into the allocation procedure, one calculates the *portfolio* skewness and kurtosis making optimization over higher moments possible and thereby refining risk assessment. The portfolio skewness S_P and excess kurtosis K_P are given by moment expressions:

$$S_P(w) = \frac{1}{\sigma_P^3(w)}(m_3 - 3m_2m_1 + 2m_1^3)$$

$$K_P(w) = \frac{1}{\sigma_P^4(w)}(m_4 - 4m_3m_1 + 6m_2m_1^2 + 3m_1^4) - 3,$$

where all four moments are functions of $w \in \mathbb{R}^d$. Given portfolio weights w, the mean m_1 is $m_1 = \mu_P(w) \stackrel{\text{def}}{=} w^\top \bar{r}$ and the variance is $\sigma_P^2(w) = w^\top \hat{\Sigma} w$. The non-central second moment is $m_2 = \sigma_P^2 + m_1^2$ and the non-central third and fourth moments are calculated from the random variable $Y \stackrel{\text{def}}{=} w^\top r$, see Bhandari and Das (2009).

The CF-VaR expansion in the multivariate case takes the form (with $w = w_t$):

$$q_\alpha(w_t) = z_\alpha + (z_\alpha^2 - 1)\frac{S_P(w_t)}{6} + (z_\alpha^3 - 3z_\alpha)\frac{K_P(w_t)}{24} - (2z_\alpha^3 - 5z_\alpha)\frac{S_P(w_t)^2}{36},$$

and the modified risk-return optimization problem (1) is solved:

$$
\begin{aligned}
& \underset{w_t \in \mathbb{R}^d}{\text{minimize}} && W_t \cdot \{-q_\alpha(w_t) \cdot \sigma_P(w_t)\} \\
& \text{subject to} && w_t^\top \mu = r_T, \;\; w_t^\top \mathbf{1} = 1, \;\; w_{t,i} \geq 0,
\end{aligned}
\tag{3}
$$

where $W_t \stackrel{\text{def}}{=} W_0 \cdot \prod_{j=1}^{t-1} w_{t-j}^\top (1 + r_{t-j})$, W_0 is the initial portfolio wealth in dollars, with $\alpha < 0.5$ being the probability (confidence) level and W_t denoting portfolio value at time t. This is a nonlinear optimization problem with linear constraints which can be solved by standard methods.

If $S_P(w_t)$, $K_P(w_t)$ are zero, then this problem is equivalent to Markowitz allocation. Indeed, then $q_\alpha(w_t) = z_\alpha$ and so the CF-VaR optimization is analogous to the initial risk-return case, up to a multiplicative constant. As argued by McFall Lamm (2003), portfolio allocation based on the CF-VaR optimization produces lower kurtosis of the distribution of chosen portfolios' returns and a positive skew which is logical and reflects the fact that such portfolios tend to reduce downside risk making extreme losses unlikely.

2.1.3 Time-Varying Covariance Structure

Financial time series exhibit volatility clustering and leverage effects. Time-varying structure in volatility comes from autoregressive dynamics in squared returns. Esti-

mation results for the sample ACFs of the squared hedge funds' returns in Figure 1 imply persistence in the variance of returns' series. A further check on conditional heteroscedasticity via the ARCH test by Engle (1982) leads to the same conclusion. This test uses the alternative hypothesis that in:

$$e_t^2 = \alpha_0 + \sum_{k=1}^{p} \alpha_k e_{t-k}^2 + u_t, \quad u_t \sim N(0, \sigma^2), \text{i.i.d.} \tag{4}$$

at least one α_k with $k = 1, \ldots, p$ is different from zero. The results for the same hedge funds returns' series are reported in Table 4. We conclude that ARCH effects are present for the indicated hedge funds' indices.
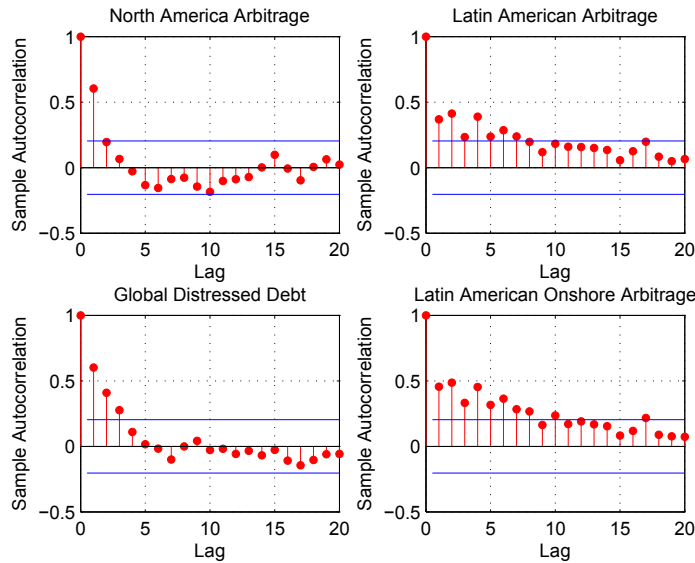
Figure 1: Sample ACFs of the squared returns for selected hedge fund indices

Table 4: Test for ARCH effects in selected hedge funds' returns residuals

Hedge Fund Name	Test Results	
	Test Stat.	P-value
North America Arbitrage	33.66	0.00
Latin American Arbitrage	6.73	0.03
Latin American Onshore Arbitrage	9.53	0.01
Global Distressed Debt FoF	60.06	0.00

The significance level for the test is 0.05; H0: no ARCH effect

$p = 2$ lags assumed

One needs to account for the time-varying structure of volatility as well as for correlation shifts in returns' covariance $\Sigma = \Sigma_t$. There have been many models proposed to deal with the issue of multivariate volatility. The simplest one is the Exponentially Weighted Moving Average estimator which emphasizes that recent mean-zero returns $r_1, \ldots, r_{t-1}$ are more relevant than earlier ones, see Tsay (2005).

Another simple volatility model is the "orthogonal GARCH" or principal component GARCH method, described, for instance, by Alexander (2001). This method imposes a univariate GARCH structure on the first principal components of a system of risk sources (returns). The orthogonal GARCH is built on a "factor" model for returns where it is assumed that the zero-mean returns matrix $Y_t \in \mathbb{R}^{n \times p}$ is generated by k factors $F_t \in \mathbb{R}^{n \times k}$ with $k < p$.

A class of models generalizes univariate volatility models to the multivariate case. The VEC model parameterizes the vector of all covariances and variances generalizing the univariate GARCH model for Σ_t; without further restrictions, this model will not guarantee positive definiteness of the predicted covariance matrix. A further extension, the BEKK model, Engle and Kroner (1995), imposes additional restrictions such that Σ_t is almost certainly positive definite, see Tsay (2005).

A more feasible model was proposed by Engle (2002) which separately estimates a series of univariate GARCH models and the correlation estimate, the so-called dynamic correlation model (DCC). This model has computational advantages over multivariate GARCH models in that the number of parameters to be estimated is independent of the number of correlated series. Some of these models are later applied to model the portfolio volatility structure in the TEDAS strategy.

2.2   High Dimensionality and TEDAS Strategy

According to the recent estimates by Preqin research and consultancy firm for 2013, there are currently over 5,200 fund management groups worldwide as of 2013 managing investment products in the hedge fund sector. They manage a combined 2.30 trillion USD globally. Therefore investors who intend to include hedge funds as assets into their portfolios, face a problem of high dimensionality of the universe of possible candidates. Therefore a technique to estimate parameters in a framework with a large number of inputs is needed. The problem of multicollinearity in a linear regression model causes coefficients to be poorly determined, imprecise and to exhibit high variance. Imposing size constraints on the coefficients alleviates the problem.

Penalized regression techniques exclude irrelevant covariates, making the model parsimonious and reducing its prediction error. The Lasso (LASSO - Least Absolute Shrinkage and Selection Operator) or L_1 penalized regression, discussed below in more detail, selects one variable from a group of highly correlated variables and ignores the others. Therefore it enables to estimate coefficients of a high-dimensional design matrix, where the number of covariates p may be much larger than the number of observations n.

On the other hand, the idea to hedge tail events motivates the use of some bench-

mark asset's conditional quantiles given the matrix of covariates. As we are primarily interested in the opposite quantile dependence in the high-dimensional setup, one uses the negative-signed Lasso-penalized quantile regression estimates to assess the relationship between the benchmark (core) asset (stock index) returns and satellite (hedging) securities (hedge funds' returns) across the conditional quantiles. This allows to do asset allocation more precisely and hedge core asset tail events, when the downside risk is especially high. Taking S&P 500 as the core asset and hedge funds as hedging satellites leads to the TEDAS - Tail Event Driven ASset allocation strategy, which is a flexible tool adjusting the traditional allocation approach for non-normality and utilizing the conditional quantile estimator for high-dimensional data.

2.3   Lasso Quantile Regression Estimator

2.3.1   Basic Setup

The Lasso estimator, see Tibshirani (1996), was first proposed for a linear model in the least-squares framework. It is used to avoid model overfitting by imposing the L_1-penalty on the coefficients and shrinking them to zero. Quantile regression estimation provides conditional quantile functions which describe the relation between response and regressors for some quantile level $\tau \in (0,1)$: consider a random sample from some distribution $\{(X_i, Y_i); i = 1, ..., n\}$, $X_i \in \mathbb{R}^p$, $Y_i \in \mathbb{R}$. Given the piecewise linear loss function

$$\rho_\tau(u) = |u|\{\tau - \mathbf{I}(u < 0)\}, \tag{5}$$

the quantile regression estimator is the solution to the convex optimization problem

$$\hat{\beta}_\tau = \arg \min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n \rho_\tau(Y_i - X_i^\top \beta); \tag{6}$$

the conditional quantile function $q_\tau(x)$ is given by:

$$q_\tau(x) \stackrel{\text{def}}{=} F_{y|x}^{-1}(\tau) = x^\top \beta(\tau) = \arg \min_{\beta \in \mathbb{R}^p} \mathsf{E}_{Y|X=x} \rho_\tau\{Y - X\beta\}, \tag{7}$$

The L_1-penalized quantile regression (LQR) estimator is then constructed as follows:

$$\hat{\beta}_{\tau,\lambda} = \quad \arg \min_{\beta \in \mathbb{R}^p} f(\beta)$$
$$\text{subject to} \quad g(\beta) \geq 0 \tag{8}$$

where

$$f(\beta) = \sum_{i=1}^{n} \rho_\tau(Y_i - X_i^\top \beta)$$

$$g(\beta) = t - \|\beta\|_1,$$

and t is the size constraint on $\|\beta\|_1$. In unrestricted form, LQR (8) is equivalent to:

$$\hat{\beta}_{\tau,\lambda} = \arg \min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau(Y_i - X_i^\top \beta) + \lambda \|\beta\|_1. \tag{9}$$

As first noted by Barrodale and Roberts (1974) and later by Koenker and Bassett (1978), (6) is equivalent to a linear program which is also the case for the LQR. There is a correspondence between λ and t which depends on the data X, Y and can be illustrated by the duality of (8), see Osborne et al. (2000). The choice of the regularization parameter λ or, equivalently, the coefficients' constraint t, is crucial for the Lasso estimator. It controls the level of penalization and the resulting shrinkage. The methods to select λ optimally, such as cross-validation, generalized cross-validation or information criteria are discussed, for instance, in Tibshirani (1996) or Efron and Tibshirani (1993).

In the context of the LQR, it is especially relevant to consider the case of high-dimensional sparse models where the overall number of regressors p is very large, possibly much larger than the sample size n, but the number of significant regressors for each conditional quantile of interest is at most q, which is smaller than the sample size, that is, $q = \mathcal{o}(n)$. A number of general regularity conditions needed for the derivation of the Lasso-penalized quantile regression estimator are usually introduced, as in Belloni and Chernozhukov (2011); similar ones are used in the setup for the Adaptive Lasso quantile regression further below.

2.3.2   Adaptive Lasso Quantile Regression

Adaptive L_1 (Lasso)-penalized quantile regression (ALQR) can simultaneously select the true model and provide a robust estimator possessing oracle properties, see appendix. The model uses the re-weighted penalty, where the weights $\hat{w}$ can be obtained from any root-n-consistent LQR estimator, as, for instance, from Belloni and Chernozhukov (2011). In general, the ALQR estimator is obtained as

$$\hat{\beta}_{\lambda_n}^{\text{adapt}} = \arg \min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau(Y_i - X_i^\top \beta) + \lambda_n \|\hat{w}^\top \beta\|_1 \tag{10}$$

It is essential to determine the regularization parameter λ_n so that the resulting estimator retains "good" properties such as asymptotic normality or variable selection

consistency, see appendix. Traditional procedures to determine λ_n, such as K-fold cross-validation, generalized cross-validation and BIC have several drawbacks. As p is increasing with the growth of the sample size, the number of potential models goes to infinity very quickly. There is no guarantee that, for instance, K-fold cross-validation provides a choice of λ_n with a proper rate. Zheng et al. (2013) suggest using a data-driven procedure to select the penalty level. The standard regularity conditions, as in Belloni and Chernozhukov (2011), are assumed to hold in this setup. For details, we refer to the appendix. A suitable choice for λ_n is

$$\lambda_n = \mathcal{O}\left\{\sqrt{q}\log(n \vee p)(\log n)^{\nu/2}\right\} \tag{11}$$

for some $\nu > 0$. The cardinality q of the non-zero coefficient set can be approximated by $\|\hat{\beta}^{\text{init}}\|_0$, where $\hat{\beta}^{\text{init}}$ is given by a consistent LQR estimator.

According to Section 1 of this study, it is necessary to estimate only those ALQR coefficients, which are non-zero and negative, which corresponds to the assets (hedge funds) oppositely related to the core asset in the sign (S&P 500) at different quantiles, when the dimensionality of the hedge funds' returns matrix X is high with $p > n$. This requirement amounts to adding one more constraint in the linear program formulation, see appendix.

3   Simulation Study and Data Analysis

3.1   Monte-Carlo Simulation

In the following simulation study, the LQR and the ALQR estimates are numerically compared. Model selection is performed with the Bayesian (Schwarz) Information Criterion (BIC), see Li and Zhu (2008). The BIC criterion under the piecewise linear loss function ρ_τ, as in (5), can be formulated as

$$\text{BIC}_{\lambda_n,\tau} \overset{\text{def}}{=} \log\left\{n^{-1} \cdot \sum_{i=1}^{n} \rho_\tau(Y_i - X_i^\top \hat{\beta}_\tau)\right\} + \frac{\log(n)}{2n} \cdot \widehat{\text{df}}(\lambda_n) \tag{12}$$

where $\widehat{\text{df}}(\lambda_n) \overset{\text{def}}{=} \|\hat{\beta}\|_0 = \hat{q}$. The linear model design is generated with $\beta = (-5, -5, -5, -5, -5, -5, 0, ..., 0)$, $X_i \sim \text{N}(0, \Sigma)$, $i = 1, \ldots, n$. The regressors are correlated with $\Sigma_{jk} = 0.5^{|j-k|}$, $j, k = 1, \ldots, p$, $n = 50$, $p = 300$; ε_i are independent, $\varepsilon_i \sim \text{N}(0, \sigma^2)$. Three levels of noise are considered: $\sigma = 0.1, 0.5, 1$. In the ALQR case, for $\hat{\beta}^{\text{init}}_{\tau,\lambda_n}$ the LQR estimator is used, where λ_n is chosen according to the BIC criterion given above. The weights for the adaptive setup are constructed according to the rule $w_j = \min(1/|\hat{\beta}^{\text{init}}_j|, \sqrt{n})$, which allows to select significant covariates in a more "adaptive" way

(so 0 is not an absorbing status anymore), see Zheng et al. (2013). The regularization parameter λ_n is selected consistently with (11):

$$\lambda_n = 0.25\sqrt{\|\hat{\beta}^{\text{init}}\|_0}\log(n \vee p)(\log n)^{0.1/2}.$$

The number of simulation replications is set to 100.

The accuracy of the model selection is assessed according to a number of criteria:

1. Standardized L_2-norm
$$\text{Dev} \overset{\text{def}}{=} \frac{\|\beta - \hat{\beta}\|_2}{\|\beta\|_2}$$

2. Sign consistency
$$\text{Acc} \overset{\text{def}}{=} \sum_{j=1}^{p} |\operatorname{sign}(\beta_j) - \operatorname{sign}(\hat{\beta}_j)|$$

3. Least angle
$$\text{Angle} \overset{\text{def}}{=} \frac{<\beta, \hat{\beta}>}{\|\beta\|_2 \cdot \|\hat{\beta}\|_2}$$

4. Estimate of true model dimension:
$$\text{Est} \overset{\text{def}}{=} \hat{q}$$

5. Empirical risk
$$\text{Risk} \overset{\text{def}}{=} \sqrt{n^{-1}\sum_{i=1}^{n}\left[X_i^\top(\beta - \hat{\beta})\right]^2}$$

The results of the simulation analysis under three levels of noise and three quantile indices $\tau = 0.1$, $\tau = 0.5$ and $\tau = 0.9$ are shown in Table 5. The ALQR method almost never over-estimates, while the LQR always does. Moreover, all the remaining accuracy criteria results confirm that the adaptive technique significantly improves the performance of quantile regression in model selection and estimation, compared with the LQR.

3.2 Data Description for Empirical Analysis

The empirical analysis aims at building an asset allocation strategy which utilizes the ALQR technique in hedging tail events. The selected coefficients capture opposite causal relations between the benchmark (the stock index S&P 500) and the hedging assets (hedge funds) at different quantiles. The input data for the empirical analysis have the following characteristics:

Table 5: Criteria Results under Different Models and Quantiles

Accuracy Crit. and Model		Noise Levels and Quantile Indices					
		$\sigma=0.1$		$\sigma=0.5$		$\sigma=1$	
		$\tau=0.1$	$\tau=0.9$	$\tau=0.1$	$\tau=0.9$	$\tau=0.1$	$\tau=0.9$
Dev	ALQR	*0.55(0.29)*	*0.60(0.27)*	*0.65(0.25)*	*0.55(0.28)*	*0.60(0.28)*	*0.57(0.27)*
	LQR	0.59(0.27)	0.63(0.25)	0.68(0.20)	0.60(0.22)	0.66(0.20)	0.63(0.21)
Acc	ALQR	*3.47(2.41)*	*3.78(2.47)*	*4.10(2.57)*	*3.15(2.25)*	*3.76(2.63)*	*3.38(2.49)*
	LQR	9.29(2.08)	10.03(2.47)	9.86(2.65)	9.47(2.68)	9.68(2.77)	9.65(2.62)
Angle	ALQR	*0.55(0.47)*	*0.63(0.47)*	*0.75(0.62)*	*0.53(0.46)*	*0.67(0.61)*	*0.57(0.48)*
	LQR	0.78(0.60)	0.89(0.61)	0.95(0.70)	0.73(0.51)	0.92(0.76)	0.80(0.56)
Est	ALQR	*5.85(1.10)*	*5.92(1.36)*	*5.82(1.29)*	*5.83(1.01)*	*5.80(1.30)*	*5.88(1.12)*
	LQR	12.33(1.83)	12.77(2.10)	12.48(1.88)	12.87(2.17)	12.44(1.93)	12.79(1.98)
Risk	ALQR	*5.64(3.40)*	*6.35(3.40)*	*6.89(3.46)*	*5.56(3.22)*	*6.45(3.68)*	*5.96(3.43)*
	LQR	6.90(3.49)	7.62(3.39)	7.91(2.99)	7.02(2.91)	7.88(3.10)	7.47(2.99)

Model notation: ALQR - Adaptive Lasso quantile regression; LQR - simple Lasso quantile regression

Standard deviations are given in brackets

Number of replications is 100

- 166 observations on monthly log-returns' series of 164 Eurekahedge hedge funds' indices in the period of 31.01.2000 - 31.10.2013 (source: *Bloomberg*);

- 166 observations on monthly log-returns of S&P 500 in the period of 31.01.2000 - 31.10.2013.

The data may show multicollinearity of the predictor matrix, but also in this case the Lasso results can be interpreted. The coefficient estimates are still unbiased, see Zou (2006).

3.3   TEDAS - Tail Event Driven ASset Allocation Strategy

The TEDAS strategy demonstrates how to select hedge funds which are oppositely related to S&P 500 in the lower tail and at higher quantiles, at which the index still has negative returns. Assuming for convenience that the median return of the hedged asset is positive, it is necessary to deal with the quantiles lower than the median. Here the ALQR is useful as it simultaneously addresses the problem of high dimensionality (excluding highly correlated covariates), provides consistent estimates of coefficients and measures causal tail relations between covariates (X, hedge funds' log-returns) and the response (Y, S&P 500 log-returns).

Given that several hedging assets have been chosen, the question of optimal portfolio composition arises. As was demonstrated, hedge fund strategies exhibit asymmetric return patterns characterized by negative skew and excess kurtosis. The traditional Markowitz risk-return analysis does not address these facts and the idea of tail risk

minimization is a natural alternative here. Together with the CF adjustment it makes the allocation more tractable in the case of non-normality.

The TEDAS strategy works as follows. Let l be the width of a moving window, $l = 80$, the set of quantile indices $\tau_{1,2,3,4,5} = (0.05, 0.15, 0.25, 0.35, 0.50)$. $F_n(x)$, x is in the domain of X_i, is defined to be the empirical distribution function of S&P 500 log-returns. $\hat{q}_\tau \overset{\text{def}}{=} F_n^{-1}(\tau)$ is the empirical quantile function of S&P 500 log-returns. The estimated negative coefficients of the ALQR are denoted as $\hat{\beta}_{\tau, \lambda_n}$. Assuming a wealth $W_0 = 1$ in the beginning of the investment period, at each time moment t; $t = l, \ldots, n$

1. determine the S&P 500 return r_t (assumed to be known, e.g., by means of a forecast)

2. choose $\tau_{j,t}, j = 1, \ldots, 5$ corresponding to the right-hand side $\hat{q}_{\tau_j,t}$ in one of the conditions which holds simultaneously: $r_t \le \hat{q}_{\tau_1,t}$, $\hat{q}_{\tau_1,t} < r_t \le \hat{q}_{\tau_2,t}$, $\hat{q}_{\tau_2,t} < r_t \le \hat{q}_{\tau_3,t}$, $\hat{q}_{\tau_3,t} < r_t \le \hat{q}_{\tau_4,t}$, $\hat{q}_{\tau_4,t} < r_t \le \hat{q}_{\tau_5,t}$

3. solve the ALQR problem for $\hat{\beta}_{\tau_j,t, \lambda_n}$ on the moving window using the observations $X \in \mathbb{R}^{t-l+1,\ldots,t \times p}$, $Y \in \mathbb{R}^{t-l+1,\ldots,t}$, buy the hedge funds with $\hat{\beta}_{\tau_j,t, \lambda_n} \ne 0$ taken with optimal weights (liquidating previous portfolio)

4. if none of the inequalities from Step 2 holds, invest into the benchmark asset (S&P 500) at r_t (liquidating previous portfolio).

This strategy is referred to as TEDAS Strategy 1. It is assumed that the optimal weights in the Step 3 are chosen as the solution of the CF-VaR optimization problem (3). The VaR confidence level α is set to 1% and the target portfolio return r_T is set to be the 70%-quantile of the mean return vector at each step. The covariance matrix of the asset returns is estimated on the recursive window from the dynamic conditional correlation model. The GARCH process for individual assets is for simplicity assumed to be GARCH (1,1) model with mean equation specified as ARMA (1,1). Higher-order moments used for the estimation of co-skewness and co-kurtosis tensors, are assumed to be constant in time.

Three alternative strategies are considered for comparison: Strategy 2 is the base case "buy-and-hold" strategy for the S&P 500 when the portfolio solely consists of the S&P 500 index and is held without rebalancing until the end of the investment period. Strategy 3 is based on the ALQR as before, but "naive" diversification is applied to the assets (the same hedge funds with non-zero coefficients): every asset receives an equal portfolio weight. Strategy 4 assumes investing into S&P 500 whenever $r_t > \hat{q}_{\tau_5,t}$ and doing asset allocation on the same set of hedge funds with a simple unadjusted variance-covariance VaR as the objective function and time-varying covariance structure modeled by the Orthogonal GARCH model, as outlined above. The threshold

level of the proportion of total variation explained by the first k principal components used to select the number of the first k most important factors is set to 95%. The comparison of the three strategies' cumulative returns is given Figure 2.
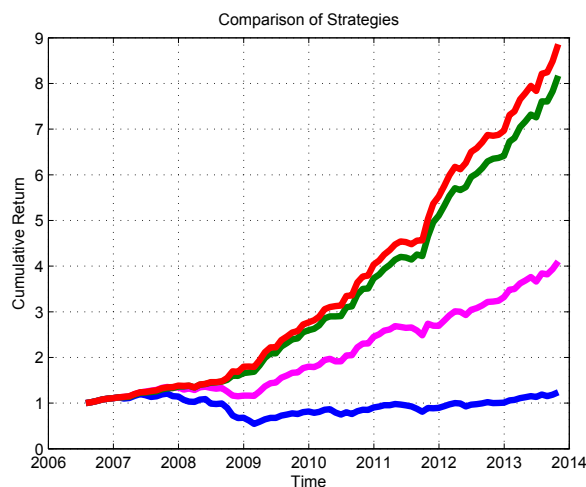

Figure 2: Strategies' cumulative returns' comparison: Strategy 1 (in red), Strategy 2 (in blue), Strategy 3 (in green), Strategy 4 (in magenta)

As is seen, Strategy 1 performs best in terms of cumulative return after 87 periods of moving-window estimation (cumulative return 785.7%): the "naive" Strategy 3, still based on the ALQR approach, yields 717.0%. Both of these two strategies outperform Strategies 2 (23.6%) and 4 (310.4%) based on portfolio allocation without ALQR. Even if one incorporates the adjustment for the three biases traditionally occurring if one uses hedge fund indices instead of hedge funds ("survivorship", "selection", and "instant history" biases, see Fung and Hsieh (2002)), the performance of strategies 1 and 3 is still better than the base case.

The investors are interested in the "risk-adjusted" return and want to know whether the return was achieved through better allocation and not by simply trading higher expected returns for higher uncertainty. It is therefore necessary to compare the presented strategies in terms of risk to assess the risk-adjusted returns. Risk in this case is measured as the value of the objective function (CF-VaR) at the optimal solution w (for Strategies 1 and 4), as CF-VaR of the portfolio with equal weights for the "naive" Strategy 3 and of the position in S&P 500 only for the "buy-and-hold" Strategy 2. It should be noted that the simple variance-covariance VaR tends to underestimate portfolio risk and the real risk will, most likely, be larger. The comparison of risks for the 4 strategies is shown in Figure 3.

The upper picture gives the VaR in dollar terms, while the lower one - in relative terms (% of portfolio value). It is not possible to decide which strategy is "better" without an expected utility analysis. Consider an investor with the logarithmic Bernoulli

18
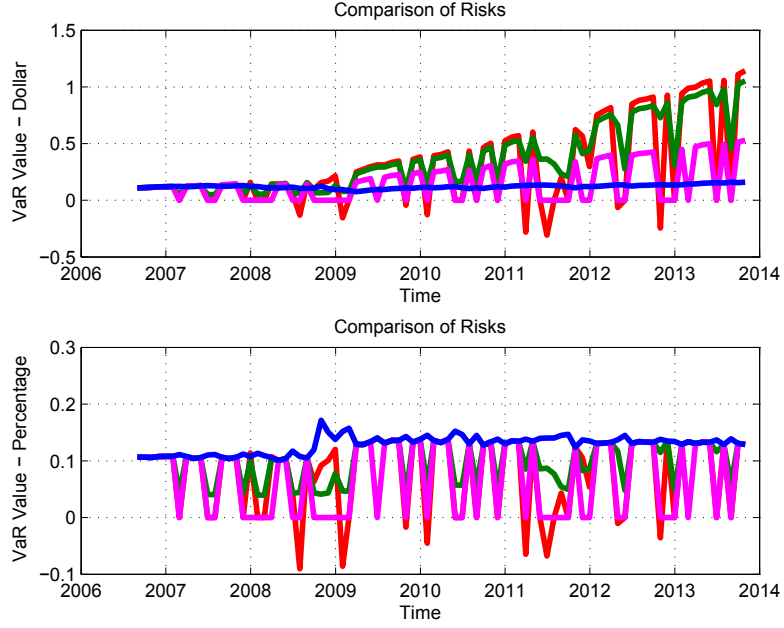
Figure 3: Absolute (upper) and relative (lower) VaR for Strategies: 1 (in red), 2 (in blue), 3 (in green), 4(in magenta)

utility function $u(x) = \log(x)$. Expanding the utility function in a Taylor series around the end of period expected wealth $\overline{W} \stackrel{\text{def}}{=} W_0 \cdot w^{\top}(1+\bar{r})$ up to the 4th order and taking expectations yields von Neumann-Morgenstern expected utility:

$$
\begin{aligned}
\mathsf{E}\{u(W_t)\} = \mathsf{E}\{u(Y)\} &\simeq u(m_1) + u'(m_1)\,\mathsf{E}(Y - m_1) + \frac{1}{2}u''(m_1)\,\mathsf{E}\{(Y - m_1)^2\} \\
&\quad + \frac{1}{3!}u^{(3)}(m_1)\,\mathsf{E}\{(Y - m_1)^3\} + \frac{1}{4!}u^{(4)}(m_1)\,\mathsf{E}\{(Y - m_1)^4\} \\
&= u(\overline{W}) + u'(\overline{W})\,\mathsf{E}(W_t - \overline{W}) + \frac{1}{2}u''(\overline{W})\,\mathsf{E}\{(W_t - \overline{W})^2\} \\
&\quad + \frac{1}{3!}u^{(3)}(\overline{W})\,\mathsf{E}\{(W_t - \overline{W}^3\} + \frac{1}{4!}u^{(4)}(\overline{W})\,\mathsf{E}\{(W_t - \overline{W})^4\} \\
&\simeq \log(\overline{W}) - \frac{1}{2\overline{W}^2}\sigma_P^2 + \frac{1}{3!\overline{W}^3}S_P - \frac{1}{4!\overline{W}^4}K_P
\end{aligned}
$$

This representation extends the Markowitz quadratic utility assumption and describes the preferences of a risk-averse investor influenced by the first four moments. Figure 4 demonstrates that the TEDAS Strategy 1 is indeed the best in terms of expected utility.

The algorithm in Strategy 1 rebalances the portfolio to hedge the benchmark asset 36 times out of 87 moving-window estimation periods. The histograms of the selected
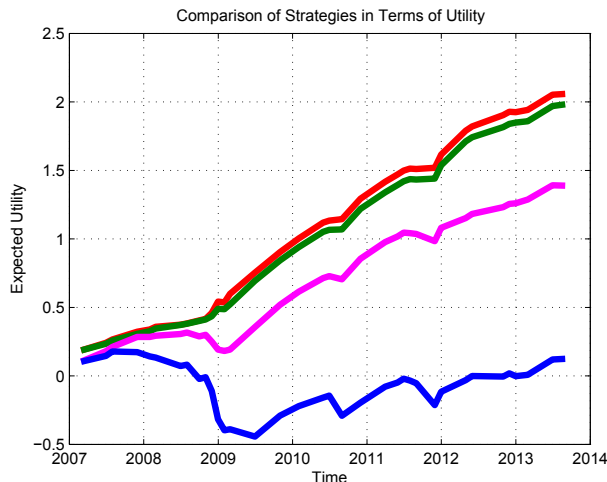
Figure 4: Comparison of expected utilities for 4 strategies: Strategy 1 (in red), Strategy 2 (in blue), Strategy 3 (in green), Strategy 4 (in magenta)

Table 6: The selected hedge funds for $\tau = 0.05$

Top 9 influential hedge funds	Frequency
Emerging Markets Arbitrage Hedge Fund Index	36
Emerging Markets CTA/Managed Futures Hedge Fund Index	36
CTA/Managed Futures Hedge Fund Index	34
Asia inc. Japan Macro Hedge Fund Index	26
Europe CTA Managed Futures Hedge Fund Index	20
Asia CTA Hedge Fund Index	13
Asia Convertible Arbitrage Hedge Fund Index	7
Europe Relative Value Hedge Fund Index	7
Europe Arbitrage Hedge Fund Index	5

variables' number $\hat{q}$ are shown in Figure 5. Frequencies for the hedge funds with non-zero coefficients selected once or more times are given in Figure 6. The names of hedge funds with significant (non-zero) negative coefficients, chosen most frequently in the 36 rebalancing attempts are shown in Table 6 for $\tau = 0.05$. Arbitrage and directional (macro and managed futures) seem to be the best candidates for tail hedging. This is rather intuitive, because these strategies trade essentially in non-equity related spreads. The phenomenon that the selected hedge funds are primarily oriented on the world region of Americas, may be induced by the fact that the benchmark index in this case is the S&P 500 and the directional strategies usually especially closely monitor the dynamics of market directions which is inevitably reflected in the behaviour S&P 500 as the most important indicator of the American market. The choice of emerging markets and arbitrage strategies may be motivated by the similar directionality of their bets and their global view on market dynamics as well as inter-market relationships and possible arbitrage opportunities.

The empirical analysis demonstrates that the TEDAS Strategy 1 shows superior
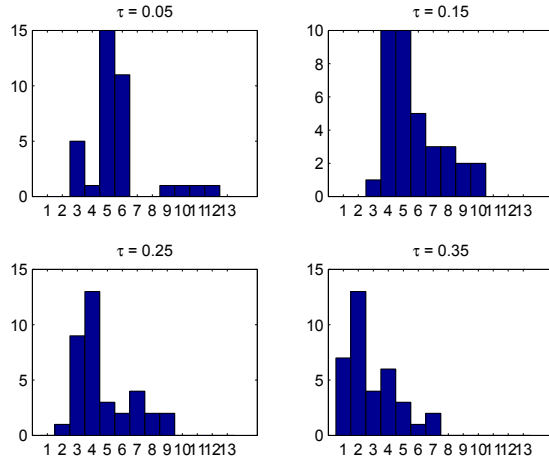
Figure 5: Frequency of the number of selected variables for 4 different τ
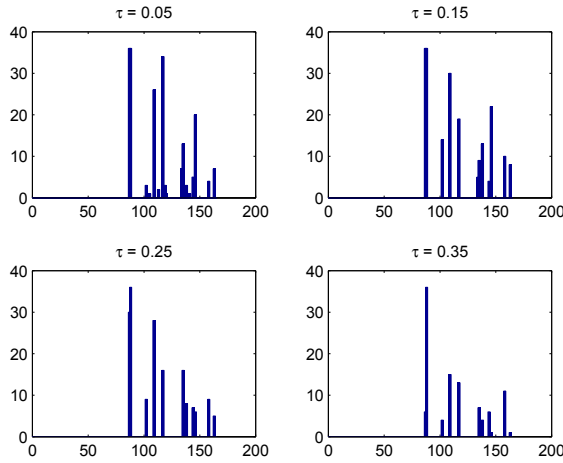
Figure 6: Frequency of selected hedge funds for 4 different τ

performance among other strategies, since it allows to address the problem of high dimensionality of the assets' returns matrices in portfolio allocation problems and makes it possible to hedge portfolio returns at different quantile levels. It also accounts for the non-normality of the portfolio returns' distribution as well as for time-varying structure of volatility and gives better performance in terms of risk-adjusted return, providing higher expected utility.

4   Conclusion

This study represents an innovative attempt to analyze the performance of an asset allocation strategy based on the TEDAS approach applied to the universe of hedge

funds. TEDAS incorporates "least absolute shrinkage and selection" which does variable selection and shrinks noise coefficients to zero and simultaneously discloses causal relationships between tail events of the benchmark asset and the covariate hedge funds returns in the high-dimensional framework. TEDAS selects portfolio weights through the minimization of the portfolio objective risk function modeled as the variance-covariance Value-at-Risk adjusted for higher portfolio return distribution moments such as skewness and kurtosis via the CF expansion; the variance-covariance return structure is allowed to be time-varying and is modeled with the dynamic conditional correlation approach to capture the effect of possible asymmetric volatility clustering.

The results prove the superior risk-return and expected utility profile of TEDAS compared to the base case S&P 500 "buy-and-hold", "naive" diversification strategies and a strategy using the so-called "orthogonal GARCH" and the simple Value-at-Risk objective portfolio risk function applied to the original universe of hedge funds without prior dimension reduction through Lasso and quantile regression estimation. These findings indicate the potential advantage of quantitative finance methods based on the ALQR. They also imply that hedge funds can be viable alternatives to conventional assets as a defensive vehicle as equity prices decline and an efficient tool for portfolio diversification and tail events hedging. It is revealed that arbitrage and directional strategies provide the best opportunities.

The ALQR technique allows estimation in the case of growing dimensions and high dimensionality, when, for the $n \times p$ returns' matrix X it is the case that $p > n$. This is convenient for the choice from a large universe of assets. The true dimensionality of the model occurs to be much smaller than the original model and the possibility to estimate this dimensionality is an obvious advantage of the adaptive method which uses an adaptively weighted penalty. The superior performance of the ALQR algorithm is confirmed by several accuracy measures in a Monte-Carlo experiment.

In general, the quantile regression framework with Lasso shrinkage applied to an asset selection problem such as hedging/trading along with a suitable choice of the objective portfolio risk function provides a convenient tool for risk measurement and modelling.

5   Appendix

5.1   Regularity conditions for Lasso quantile regression

A "restricted set" needs to be defined for one of the following conditions: define $\overline{T}_\tau(\delta, m) \subset \{1, \ldots, p\} \backslash T_\tau$ as the support of the m largest in absolute value components of the vector δ outside of $T_\tau \stackrel{\text{def}}{=} \text{support}(\beta_\tau)$ and $\overline{T}_\tau(\delta, m)$ is the empty set if

$m = 0$. Then, for some $c_0 \geq 0$ and each $\tau \in \mathcal{T}$ the *restricted set* is defined as

$$A_\tau \overset{\text{def}}{=} \{\delta \in \mathbb{R}^p : \|\delta_{T_\tau^C}\|_1 \leq c_0 \|\delta_{T_\tau}\|_1, \|\delta_{T_\tau^C}\|_0 \leq n\}$$

. The regularity conditions are then stated as follows:

D.1. *Sampling and Smoothness.* The data $(Y_i, X_i^\top)^\top \in \mathbb{R}^{1+p}$, $i = 1, \ldots, n$ are an i.i.d. sequence of vectors, $X_{i1} = 1$, $n \wedge p \geq 3$. For each x in the support of X_i and $\forall y \in \mathbb{R}$, the conditional density $f_{Y_i|X_i}(y|x)$ is continuously differentiable in y; $f_{Y_i|X_i}(y|x)$ and $\frac{\partial}{\partial y} f_{Y_i|X_i}(y|x)$ are bounded in absolute value by constants $\overline{f}$ and $\overline{f'}$, uniformly in y and x. Also, the conditional density of Y_i evaluated at the conditional quantile $X_i^\top \beta_\tau$ is bounded away from zero uniformly for any x in the support of X_i, that is, $f_{Y_i|X_i}(x^\top \beta_\tau | x) > \underline{f} > 0$.

D.2. *Sparsity and Smoothness of the τth Conditional Quantile Function $F_{y|x}^{-1}(\tau|x)$:* $\tau \to \beta_\tau$. Let $\mathcal{T}$ be a compact subset of (0,1). The coefficients β_τ are sparse (have multiple zero elements) and smooth with respect to $\tau \in \mathcal{T}$:

$$\sup_{\tau \in \mathcal{T}} \|\beta_\tau\|_0 \leq q \quad and \quad \|\beta_\tau - \beta_{\tau'}\| \leq L|\tau - \tau'|, \quad \forall \tau, \tau' \in \mathcal{T}$$

where $q \geq 1$ and $\log L \leq C_L \log(p \vee n)$ for some constant C_L.

D.3. *Well-Behaved Covariates.* Covariates are normalized such that $\sigma_j^2 = \mathsf{E}(X_{ij}^2) = 1$ for all $j = 1, \ldots, p$ and $\hat{\sigma}_j^2 = (1/n) \sum_{i=1}^n X_{ij}^2$ obeys

$$\mathsf{P}(\max_{1 \leq j \leq p} |\hat{\sigma}_j - 1| \leq 0.5) \geq 1 - \gamma \to 1 \text{ as } n \to \infty,$$

for some probability level $1 - \gamma$.

D.4. *Restricted Identifiability and Nonlinearity.* For some constants $m \geq 0$ and $c_0 \geq 9$, the matrix $\mathsf{E}(X_i X_i^\top)$ satisfies the "restricted eigenvalue" condition

$$\kappa_m^2 \overset{\text{def}}{=} \inf_{\tau \in \mathcal{T}} \inf_{\delta \in A_\tau, \delta \neq 0} \frac{\delta^\top \mathsf{E}(X_i X_i^\top)\delta}{\|\delta_{T_\tau \cup \overline{T}_\tau(\delta,m)}\|^2} > 0 \tag{13}$$

and $\log(\underline{f}\kappa_0^2) \leq C_f \log(n \vee p)$ for some constant C_f. Moreover, the "restricted non-linear impact coefficient η" is defined to be as

$$\eta \overset{\text{def}}{=} \frac{3\underline{f}^{3/2}}{8\overline{f'}} \inf_{\tau \in \mathcal{T}} \inf_{\delta \in A_\tau, \delta \neq 0} \frac{\mathsf{E}[|X_i^\top \delta|^2]^{3/2}}{\mathsf{E}[|X_i^\top \delta|^3]} > 0 \tag{14}$$

Condition D.1 is a smoothness assumption on the conditional density of Y given X; it does not impose normality or homoscedasticity assumptions. Also the assumption

that the conditional density is bounded from below at the conditional quantile is standard. Condition D.2 imposes sparsity and smoothness on the behavior of the quantile regression coefficients β_τ as the quantile index τ is varied. The third condition, D.3, requires that $\hat{\sigma}_j$ does not deviate strongly from σ_j and normalizes $\sigma_j^2 = 1$. Condition D.4 requires that there exists a constant C_f, such that $\kappa_0^2 \leq C_f$, which together with the fact that κ_m^2 is non-increasing in m, entails that the smallest eigenvalue of the covariance matrix $\Sigma_q \overset{\text{def}}{=} \mathsf{E}(X_i^1 X_i^{1\top})$ is finite and bounded away from 0. Here X^1 is defined from the following representation:

$$Y = X\beta + \varepsilon = X^1\beta^1 + X^2\beta^2 + \varepsilon \tag{15}$$

where $X = (X^1, X^2)$, $X^1 \in \mathbb{R}^{n \times q}$, $X^2 \in \mathbb{R}^{n \times (p-q)}$; $\beta^1 \in \mathbb{R}^q$ are true nonzero coefficients, $\beta^2 \in \mathbb{R}^{p-q} = 0$ are noise coefficients and $q = \|\beta\|_0$ is the cardinality of the true β, or the number of non-zero coefficients. The "restricted non-linear impact coefficient" η from D.4, as noted in Belloni and Chernozhukov (2011), p. 7, controls the quality of minoration of the quantile regression objective function by a quadratic function over the restricted set.

Under the conditions D.1-D.4 asymptotic results are obtained in Belloni and Chernozhukov (2011) for the linear setup, which is as before, in particular, $X = (1, Z)^\top$, $Z_i \sim \mathrm{N}(0, \Sigma)$, where $Z = (X_2, \ldots, X_n)^\top$ (the covariates' matrix without constant). The normality of errors can be easily relaxed by allowing for the disturbance ε to have a smooth density that obeys the conditions stated in D.1. The sparsity and variable selection consistency conditions are satisfied and the L_1-penalized quantile regression estimator is consistent at the near-oracle rate. It can be seen, however, that the penalty for each variable in (10) is of the same order, λ_n/n, and hence not quite adaptive in the sense that it surely does not overpenalize significant variables. A further attempt to improve the estimator is the ALQR method which requires additional regularity assumptions:

D.5. *Growth Rate of Covariates.* The growth rate of significant variables and all variables allowed is assumed to satisfy

$$\frac{q^3\{\log(n \vee p)\}^{2+\nu}}{n} \to 0, \quad \text{for some } \nu > 0.$$

D.6. *Moments of Covariates.* Covariates satisfy the Cramér condition

$$\mathsf{E}[|X_{ij}|^k] \leq 0.5 C_m M^{k-2} k!$$

for some constants C_m, M, $\forall k \geq 2$, $j = 1, \ldots, p$.

D.7. *Well-Separated Regression Coefficients.* It is assumed that $\exists b_0 > 0$, such that $\forall j \leq q$, $|\hat{\beta}_j| > b_0$.

Condition D.5 limits the size of significant variables to be less than $n^{1/3}$ which is sufficient to obtain the consistency and asymptotic normality under the full model and discontinuous score function of $\rho_\tau(\cdot)$. Condition D.6 is necessary to establish the sparsity property of the Adaptive L_1 quantile estimator through Bernstein's inequality. Condition D.7 is essential for establishing the oracle consistency property and also assumes that the parameter values of the true model are uniformly bounded away from 0.

Assuming that the regularity conditions D.1, D.4-D.7 are satisfied and that λ_n satisfies

$$\lambda_n q/\sqrt{n} \to 0 \text{ and } \lambda_n/\{\sqrt{q}\log(n \vee p)\} \to \infty, \tag{16}$$

the ALQR estimator satisfies the following oracle properties, as in Zheng et al. (2013):

1. Variable selection consistency:

$$\mathrm{P}(\beta^2 = 0) \geq 1 - 6\exp\left\{-\frac{\log(n \vee p)}{4}\right\}.$$

2. Estimation consistency:

$$\|\beta - \hat{\beta}\| = \mathcal{O}_p\left(\sqrt{\frac{q}{n}}\right).$$

3. Asymptotic normality: let $u_q^2 \overset{\text{def}}{=} \alpha^\top \Sigma_q \alpha$, $\forall \alpha \in \mathbb{R}^q$, such that $\|\alpha\| < \infty$, then

$$n^{1/2}u_q^{-1}\alpha^\top(\beta^1 - \hat{\beta}^1) \overset{\mathcal{L}}{\to} \mathrm{N}\left\{0, \frac{(1-\tau)\tau}{f^2(\gamma^*)}\right\}$$

where γ^* is the τth quantile and f is the pdf of ε.

5.2   Non-positive Lasso quantile regression optimization problem

The Lasso-penalized QR problem with an additional non-positivity constraint takes the following form:

$$
\begin{aligned}
\underset{(\xi,\zeta,\eta,\tilde{\beta}) \in \mathbb{R}_+^{2n} \times \mathbb{R}^{2p}}{\text{minimize}} \quad & \tau 1_n^\top \xi + (1-\tau)1_n^\top \zeta + \lambda 1_n^\top \eta \\
\text{subject to} \quad & \xi - \zeta = Y + X\tilde{\beta}, \\
& \xi \geq 0, \\
& \zeta \geq 0, \\
& \eta \geq \tilde{\beta}, \\
& \eta \geq -\tilde{\beta}, \\
& \tilde{\beta} \geq 0, \quad \tilde{\beta} \overset{\text{def}}{=} -\beta
\end{aligned}
\tag{17}
$$

with "slack" variables ξ, ζ and η.

Transformed into matrix form, this problem can be equivalently re-written as

$$
\begin{aligned}
\text{minimize} \quad & c^\top x \\
\text{subject to} \quad & Ax = b, \ \ Bx \leq 0, \\
& l_i \leq x_i \leq u_i, \ \ i \in \mathcal{I} \\
& 0 \leq x_i, \ \ i \in \mathcal{J}
\end{aligned}
$$

where $\mathcal{I}$, $\mathcal{J}$ are disjoint index sets such that $\mathcal{I} \cup \mathcal{J} = \{1, 2, \ldots, n\}$, $\mathcal{I} \cap \mathcal{J} = \emptyset$, $A = \begin{pmatrix} I_n & -I_n & 0 & X \end{pmatrix}$, $b = Y$, $x = \begin{pmatrix} \xi & \zeta & \eta & \beta \end{pmatrix}^\top$, and

$$
c = \begin{pmatrix} \tau 1_n \\ (1 - \tau)1_n \\ \lambda 1_p \\ 01_p \end{pmatrix}, \quad B = \begin{pmatrix} -E_{p \times n} & 0 & 0 & 0 \\ 0 & -E_{p \times n} & 0 & 0 \\ 0 & 0 & -I_p & I_p \\ 0 & 0 & -I_p & -I_p \\ 0 & 0 & 0 & I_p \end{pmatrix}
$$

where I_p is $p \times p$ identity matrix; $E_{p \times n} = \begin{pmatrix} I_n \\ 0 \end{pmatrix}$. Denote the dimension of matrix A as $m \times \tilde{n}$. Without loss of generality, as noted by Zhang (1998), it is also assumed that for some positive integer $n_u \leq \tilde{n}$

$$
\mathcal{I} = \{1, 2, \ldots, n_u\}, \quad \mathcal{J} = \{n_u + 1, n_u + 2, \ldots, \tilde{n}\}.
$$

A suitable algorithm to compute the solution is an efficient large-scale primal-dual infeasible-interior-point algorithm using the Newton method as solver, outlined in Zhang (1998).

6   References

Abramowitz, M. and Stegun, I.A. (1965): Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables. *Dover Publications.*

Alexander, C.O. (2001): Orthogonal GARCH, in "Mastering Risk", Volume II. *FT-Prentice Hall*, 21–38.

Barrodale, I. and Roberts, F. (1974): "Solution of an Overdetermined System of Equations in the L_1 Norm", *Communications of the ACM*, **17**, 319-320.

Belloni, A., Chernozhukov, V. (2011): "L_1-penalized Quantile Regression in High-Dimensional Sparse Models", *Annals of Statistics*, **39(1)**, 82–130.

Bhandari, R. and Das, S.R. (2009): "Options on Portfolios with Higher-Order Moments", *Finance Research Letters*, **6**, 122-129.

Brandimarte, P. (2006): Numerical Methods in Finance and Economics. A Matlab-Based Introduction. *Wiley.*

Byström, H.N.E. (2004): "Orthogonal GARCH and Covariance Matrix Forecasting: The Nordic Stock Markets During the Asian Financial Crisis 1997-1998" *The European Journal of Finance*, **10(1)**, 44–67

Cvitanić, J., Lazrak, A., Martellini, L., and Zapatero, F. (2003): "Optimal Allocation to Hedge Funds: an Empirical Analysis" *Quantitative Finance*, **3(1)**, 28–39

Dowd, K. (2005): Measuring Market Risk. 2nd Edition. *Wiley.*

Efron, B. and Tibshirani, R.J. (1993): An Introduction to the Bootstrap. *Chapman & Hall/CRC.*

Engle, R. (1982): "Autoregressive Conditional Heteroscedasticity with Estimates of Variance of United Kingdom Inflation", *Econometrica*, **50(4)**, 987–1007.

Engle, R. and Kroner, K. (1995): "Multivariate Simultaneous GARCH", *Econometric Theory*, **11**, 122–150.

Engle, R. (2002): "Dynamic Conditional Correlation: a Simple Class of Multivariate GARCH Models", *Journal of Business and Economic Statistics*, **20(3)**, 339–350.

Favre, L. and Galeano, J.-A. (2002): "Mean-Modified Value-at-Risk Optimization with Hedge Funds" *The Journal of Alternative Investments*, **5(2)**, 21–25

Franke, J. and Härdle, W.K., Hafner, C.M. (2011): Statistics of Financial Markets. An Introduction. 3rd Edition. *Springer.*

Fung, W. and Hsieh, D.A. (2002): "Hedge-Fund Benchmarks: Information Content and Biases" *Financial Analysts Journal*, **58(1)**, pp.22-34

Giamouridis, D. and Vrontos, I.D. (2007): "Hedge Fund Portfolio Construction: a Comparison of Static and Dynamic Approaches" *Journal of Banking and Finance*, **31**, 199–217

Koenker, R. and Bassett, G. (1978): "Regression Quantiles", *Econometrica*, **46**, 33-50

Lhabitant, F.-S. (2002): Hedge Funds. Myths and Limits. *Wiley.*

Lhabitant, F.-S. and Learned, M. (2002): "Hedge Fund Diversification: How Much Is Enough". *The Journal of Alternative Investments*, **5(3)**, 23–49

Li, Y. and Zhu, J. (2008): "L_1-Norm Quantile Regression". *Journal of Computational and Graphical Statistics,*, **17(1)**, 1–23

Lintner, J. (1983): The Potential Role of Managed Commodity-Financial Futures Accounts (and/or Funds) in Portfolios of Stocks and Bonds. *Presentation to the Financial Analysts Federation.*

McFall Lamm, R. (1999): Portfolios of Alternative Assets: Why Not 100% Hedge Funds? *Journal of Investing*, **8(4)**, 87–97

McFall Lamm, R. (2003): "Asymmetric Returns and Optimal Hedge Fund Portfolios". *The Journal of Alternative Investments*, **6(2)**, 9–21

Osborne, M.R., Presnell, B. and Turlach, B.A. (2000): "On the LASSO and Its Dual". *Journal of Computational and Graphical Statistics*, **9(2)**, 319–337

Tibshirani, R. (1996): "Regression Shrinkage and Selection via the Lasso". *Journal of Royal Statistical Society*, **58(1)**, 267–288

Tsay, R. (2005): Analysis of Financial Time Series. 2nd Edition. *Wiley*

Zhang, Y. (1998): "Solving Large-Scale Linear Programs by Interior-Point Methods under the Matlab Environment", *Optimization Methods and Software*, **10**, 1–31

Zheng, Qi, Gallagher, C. and Kulasekera, K.B. (2013): "Adaptive Penalized Quantile Regression for High-Dimensional Data". *Journal of Computational and Graphical Statistics*, **143**, 1029–1038

Zou, H. (2006): "The Adaptive Lasso and its Oracle Properties". *Journal of Statistical Planning and Inference*, **101(476)**, 1418–1429

SFB 649 Discussion Paper Series 2014

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Principal Component Analysis in an Asymmetric Norm" by Ngoc Mai Tran, Maria Osipenko and Wolfgang Karl Härdle, January 2014.
002 "A Simultaneous Confidence Corridor for Varying Coefficient Regression with Sparse Functional Data" by Lijie Gu, Li Wang, Wolfgang Karl Härdle and Lijian Yang, January 2014.
003 "An Extended Single Index Model with Missing Response at Random" by Qihua Wang, Tao Zhang, Wolfgang Karl Härdle, January 2014.
004 "Structural Vector Autoregressive Analysis in a Data Rich Environment: A Survey" by Helmut Lütkepohl, January 2014.
005 "Functional stable limit theorems for efficient spectral covolatility estimators" by Randolf Altmeyer and Markus Bibinger, January 2014.
006 "A consistent two-factor model for pricing temperature derivatives" by Andreas Groll, Brenda López-Cabrera and Thilo Meyer-Brandis, January 2014.
007 "Confidence Bands for Impulse Responses: Bonferroni versus Wald" by Helmut Lütkepohl, Anna Staszewska-Bystrova and Peter Winker, January 2014.
008 "Simultaneous Confidence Corridors and Variable Selection for Generalized Additive Models" by Shuzhuan Zheng, Rong Liu, Lijian Yang and Wolfgang Karl Härdle, January 2014.
009 "Structural Vector Autoregressions: Checking Identifying Long-run Restrictions via Heteroskedasticity" by Helmut Lütkepohl and Anton Velinov, January 2014.
010 "Efficient Iterative Maximum Likelihood Estimation of High-Parameterized Time Series Models" by Nikolaus Hautsch, Ostap Okhrin and Alexander Ristig, January 2014.
011 "Fiscal Devaluation in a Monetary Union" by Philipp Engler, Giovanni Ganelli, Juha Tervala and Simon Voigts, January 2014.
012 "Nonparametric Estimates for Conditional Quantiles of Time Series" by Jürgen Franke, Peter Mwita and Weining Wang, January 2014.
013 "Product Market Deregulation and Employment Outcomes: Evidence from the German Retail Sector" by Charlotte Senftleben-König, January 2014.
014 "Estimation procedures for exchangeable Marshall copulas with hydrological application" by Fabrizio Durante and Ostap Okhrin, January 2014.
015 "Ladislaus von Bortkiewicz - statistician, economist, and a European intellectual" by Wolfgang Karl Härdle and Annette B. Vogt, February 2014.
016 "An Application of Principal Component Analysis on Multivariate Time-Stationary Spatio-Temporal Data" by Stephan Stahlschmidt, Wolfgang Karl Härdle and Helmut Thome, February 2014.
017 "The composition of government spending and the multiplier at the Zero Lower Bound" by Julien Albertini, Arthur Poirier and Jordan Roulleau-Pasdeloup, February 2014.
018 "Interacting Product and Labor Market Regulation and the Impact of Immigration on Native Wages" by Susanne Prantl and Alexandra Spitz-Oener, February 2014.

SFB 649 Discussion Paper Series 2014

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.