TEDAS - Tail Event Driven ASset Allocation: equity and mutual funds' markets

Wolfgang Karl Härdle

Sergey Nasekin

Alla Petukhina

Ladislaus von Bortkiewicz Chair of Statistics
C.A.S.E. – Center for Applied Statistics
and Economics
Humboldt–Universität zu Berlin
http://lvb.wiwi.hu-berlin.de
http://case.hu-berlin.de

Strategies comparison: hedge funds' indices

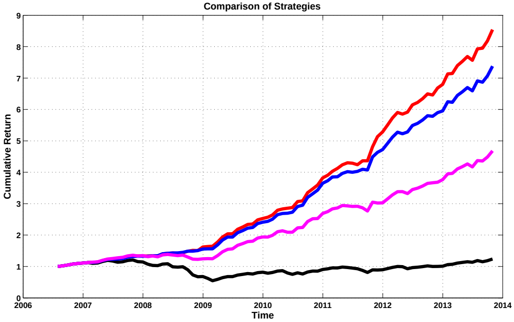


Figure 1: Strategies' cumulative returns' comparison: TEDAS Basic, S&P500, TEDAS Naïve, RR

- Härdle et al. (2014)
  - TEDAS applied to hedge funds' indices performs better than benchmark models
- Limitation of using hedge indices as portfolio assets
- Application of TEDAS approach to global mutual funds' data and German stock market

Core & Satellites

Mutual funds, SDAX, MDAX and TecDAX constituents

- diversification - reduction of the portfolio risk
- construction - a more diverse universe of assets
- allocation - a higher risk-adjusted return.

- Comparison of the TEDAS with more benchmark strategies:
  - 60/40 portfolio
  - Risk Parity (equal risk portfolio contribution)
  - Mean-Variance strategy
- TEDAS parameters optimisation

Tail Risk

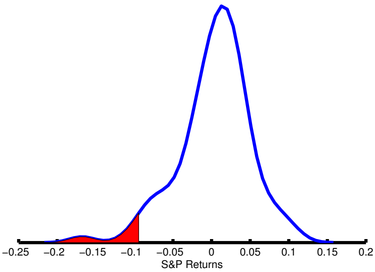


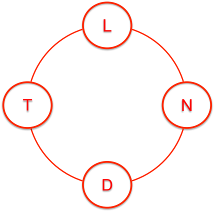
Figure 2: Estimated density of S&P 500 returns

The TLND challenge

- **T**ail dependence
- **L**arge universe: $p > n$
- **N**on normality
- **D**ynamics

TEDAS Objectives

- Hedge tail events
  - Quantile regression
  - Variable selection in high dimensions
- Improve Asset Allocation
  - Higher-order moments' optimization
  - Modelling of moments' dynamics



Outline

1. Motivation ✓
2. TEDAS framework
3. Data
4. Empirical Results
5. Discussion: choice of τ-spine
6. Conclusions

Tail Events

- $Y \in \mathbb{R}^n$ core log-returns; $X \in \mathbb{R}^{n \times p}$ satellites' log-returns, $p > n$
-

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

$$\rho_\tau(u) = u\{\tau - \mathbf{I}(u < 0)\}$$

- L_1 penalty $\lambda_n \|\hat{\omega}^\top \beta\|_1$ to nullify "excessive" coefficients; λ_n and $\hat{\omega}$ controlling penalization; constraining $\beta \leq 0$ yields ALQR ▸ Details

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

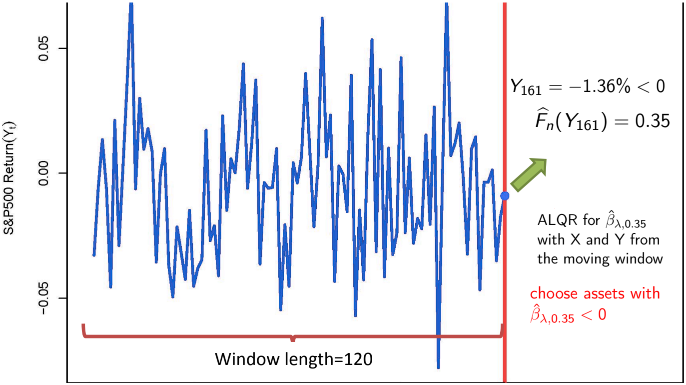
TEDAS Step 1

Initial wealth $W_0 = \$1$, $t = l, \ldots, n$; $l = 120$ length of the moving window

- Portfolio constituents' selection

  1. determine core asset return Y_t, set $\tau_t = \widehat{F}_n(Y_t)$
     $\tau_{j=1,\ldots,5} = (0.05, 0.15, 0.25, 0.35, 0.50)$
  2. ALQR for $\hat{\beta}_{\tau_t,\lambda_n}$ using the observations $X \in \mathbb{R}^{t-l+1,\ldots,t \times p}$, $Y \in \mathbb{R}^{t-l+1,\ldots,t}$
  3. Select $\tau_{j,t}$ according to the right-side $\hat{q}_{\tau_{j,t}}$ in: $Y_t \leq \hat{q}_{\tau_{\mathbf{1},t}}$ or $\hat{q}_{\tau_{\mathbf{1},t}} < Y_t \leq \hat{q}_{\tau_{j,t}}$

TEDAS Step 1


S&P500 Return(Y_t)
0.05
0.00
−0.05
$Y_{161} = -1.36\% < 0$
$\widehat{F}_n(Y_{161}) = 0.35$
ALQR for $\hat{\beta}_{\lambda,0.35}$
with X and Y from
the moving window
choose assets with
$\hat{\beta}_{\lambda,0.35} < 0$
Window length=120

TEDAS Step 2

- Portfolio selection
    1. apply TEDAS Gestalt to X_j, obtain $\widehat{w}_t \in \mathbb{R}^k$
    2. determine the realized portfolio wealth for $t+1$, $\widehat{X}_{t+1} \stackrel{\text{def}}{=} (X_{t+1,1}, \dots, X_{t+1,k})^\top$: $W_{t+1} = W_t(1 + \widehat{w}_t^\top \widehat{X}_t)$

Rebalancing of portfolio:

- one of inequalities in step 3 holds
  - sell the core portfolio and buy satellites (step 4) with estimated weights (step 5)
  - stay "in cash" if there are no adversely moving satellites (step 4)
- no one of inequalities holds: invest in the core portfolio
- period (t+1), if no one of inequalities (step 3) holds, we return to the core portfolio

TEDAS Example

1. Suppose $t = 161$ (May. 2011), accumulated wealth $W_{161} = \$2.301$, $Y_{161} = -1.36\% < 0$
2. $\widehat{F}_n(Y_{161}) = 0.35$, so estimate ALQR for $\hat{\beta}_{\lambda,0.35}$
3. ALQR on $X \in \mathbb{R}^{120\times 583}$, $Y \in \mathbb{R}^{120}$ yields $\hat{\beta}_{0.35} = (-1.12, -0.41)^\top$, *Blackrock Eurofund Class I, Pimco Funds Long Term United U.S. States Government Institutional Shares*
4. TEDAS CF-CVaR optimization $\widehat{w}_{161} = (0, 1)^\top$; $\widehat{X}_{162} = (0.014, 0.026)^\top$, $W_{162} = W_{161}(1 + \widehat{w}_{161}^\top \widehat{X}_{161}) = \$2361$

TEDAS Gestalten

TEDAS gestalt	Dynamics modelling	Weights optimization
TEDAS Naïve	NO	Equal weights
TEDAS Hybrid	NO	Mean-variance optimization of weights ▸ Details
TEDAS Basic	DCC volatility ▸ Details	CF-VaR optimization ▸ Details

Small and mid caps German stocks

- MDAX
  - 50 medium-sized German public limited companies and foreign companies primarily active in Germany from traditional sectors
  - Ranks after the DAX30 based on market capitalisation and stock exchange turnover
- SDAX
  - The selection index for smaller companies from traditional sectors
  - 50 stocks from the Prime Standard
- TecDAX
  - Comprises the 30 largest technology stocks below the DAX

Size premium

- Banz (1981) and Reinganum (1981): the US small cap stocks outperformed large-cap stocks (in 1936-1975)
- Fama, French (1992, 1993): a size premium of 0.27% per month in the US over the period 1963-1991
- Results are robust:
  - for stock price momentum by Jegadeesh , Titman (1993) and Carhart (1997)
  - for liquidity by Pastor, Stambaugh (2003) and Ibbotson, Hu (2011)
  - for industry factors, high leverage, low liquidity by Menchero et al. (2008)

Why small and mid cap stocks?

- Strong absolute returns
- Diversification benefits (Eun, Huang, Lai (2006))
- High risk-adjusted returns

Strong absolute returns

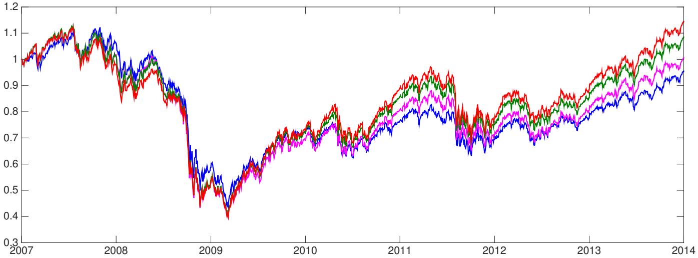


Figure 3: Cumulative index performance: MSCI World Large Cap, MSCI World Mid Cap, MSCI World Small Cap, MSCI World Small and Mid Cap

Diversification

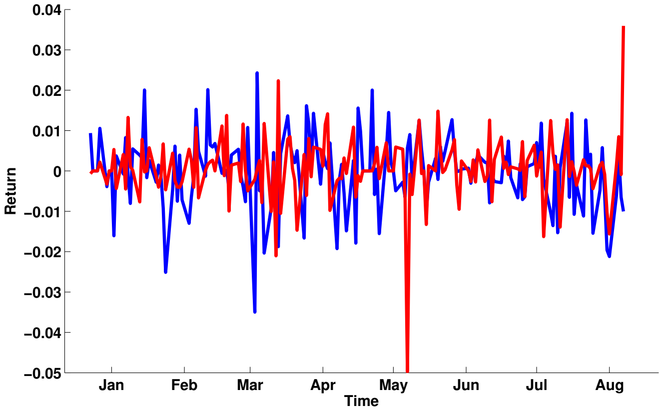
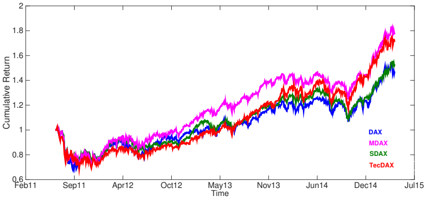


Figure 4: DAX and Hamborner REIT AG daily returns in 20131220-20140831

German equity

- Frankfurt Stock Exchange (Xetra), weekly data
  - 125 stocks - SDAX (48), MDAX (47) and TecDAX (50) as on 20140801
  - DAX index



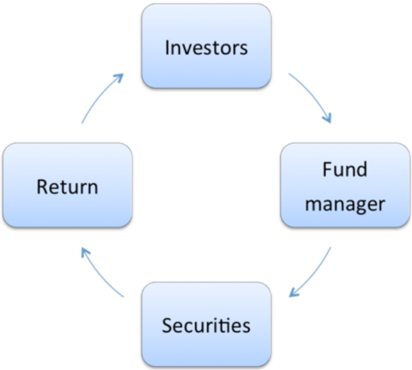
- Span: 20121221 - 20141127 (100 trading weeks)

- Source: Datastream

Mutual Fund

- Open-End: buy and sell the shares, meet the demand for customers
- Unit Investment Trust: exchange-traded fund (ETF), Fixed/ unmanaged Portfolio
- Closed-End: fixed number of shares, not redeemable by the fund, buy and sell on the exchange

Mutual fund flowchart



Why Mutual Funds?

- Importance of MF
  - $30 trillion worldwide, 15 trillion in U.S in 2013
  - 88% investment companies managed asset by holding MF
- Big data: 76 200 MFs worldwide in 2013
- Diversification

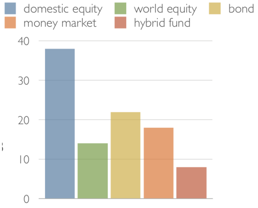


Figure 5: Structure of U.S. Mutual funds, by asset classes

Dynamics of Mutual funds investment

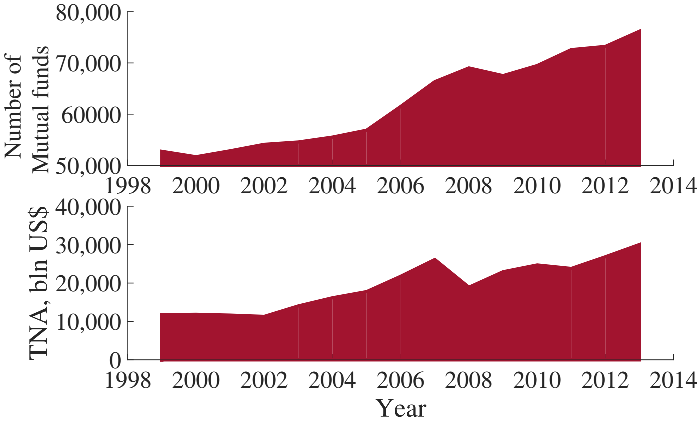
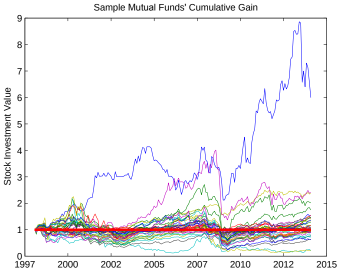


Figure 6: Worldwide Mutual Funds: total number and TNA

Mutual Funds

- Monthly data
  - Core asset (Y): S&P500
  - Satellite assets (X): 583 Mutual funds
- Span: 19980101 - 20131231 (192 months)
- Source: Datastream



Benchmark Strategies

1. RR: dynamic risk-return optimization ▸ Details
2. ERC: Risk-parity portfolio (equal risk contribution) ▸ Details
3. 60/40 portfolio ▸ Details

TEDAS approach:German stocks' results

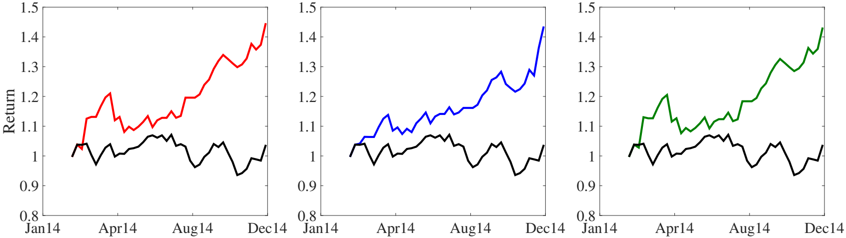


Figure 7: Strategies' cumulative returns' comparison: TEDAS Basic, TEDAS Naïve, TEDAS Hybrid, DAX30

TEDAS_strategies

TEDAS approach:German stocks' results

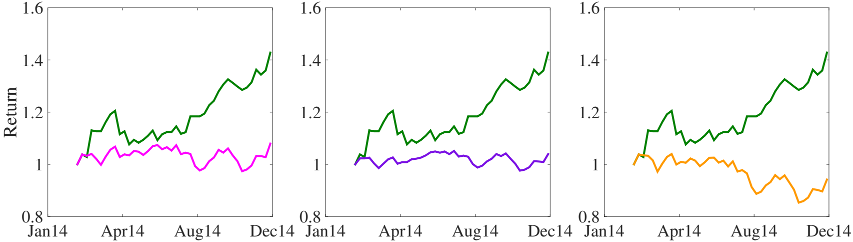


Figure 8: Strategies' cumulative returns' comparison (with transaction costs 1% of portfolio value): TEDAS Hybrid, 60/40, ERC, RR

TEDAS_strategies

Strategies' performance: German stocks

Strategy	Cumulative return	Sharpe ratio	Maximum drawdown
TEDAS Basic	144%	0.3792	0.1069
TEDAS Naïve	143%	0.3184	0.0564
TEDAS Hybrid	143%	0.3079	0.1068
RR	108%	0.0687	0.0934
ERC	129%	-0.0693	0.1792
60/40	121%	0.0306	0.0718
DAX30	103%	0.0210	0.1264

TEDAS approach:German stocks' results

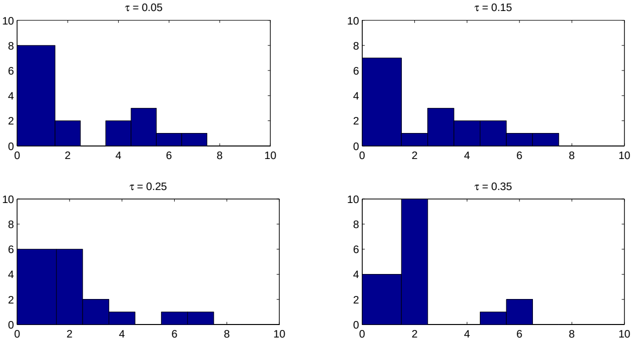


Figure 9: Frequency of the number of selected variables for 4 different τ

TEDAS approach:German stocks results

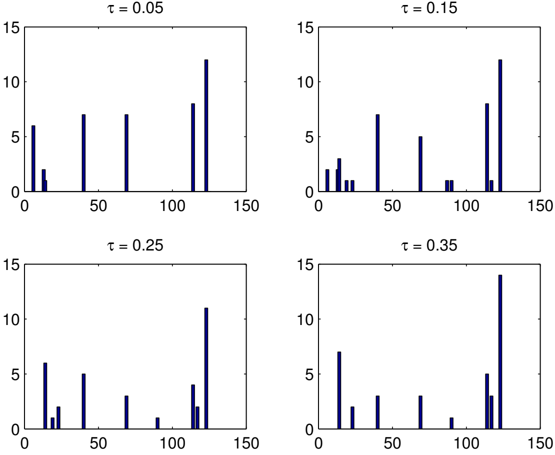


Figure 10: The frequency of stocks

Selected Stocks

Table 1: The selected German Stocks for $\tau = 0.05$

Top 5 influential Stocks	Frequency	Index	Industry
Sartorius Aktiengesellschaft	12	TecDAX	Provision of laboratory and process technologies and equipment
XING AG	8	TecDAX	Online business communication services
Surteco SE	7	SDAX	Household Goods & Home Construction
Kabel Deutschland Holding AG	7	MDAX	Cable-based telecommunication services
Biotest AG	6	MDAX	Producing biological medications

-$\widehat{\beta}$ in each window, $\tau = 0.05$

Figure 11: Different $-\widehat{\beta}$ in application; $\tau = 0.05$ ▸ Selected Stocks

TEDAS approach:Mutual Funds results

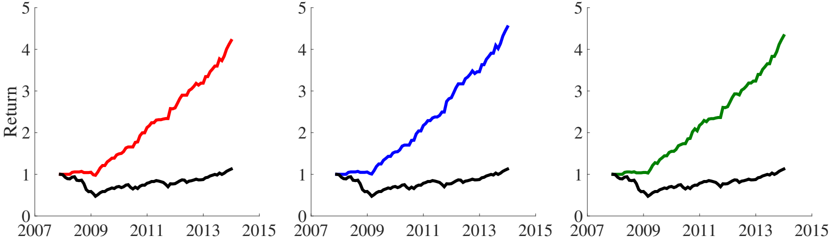


Figure 12: Strategies' cumulative returns' comparison: TEDAS Basic, TEDAS Naïve, TEDAS Hybrid, S&P500

TEDAS_strategies

TEDAS approach:Mutual Funds results

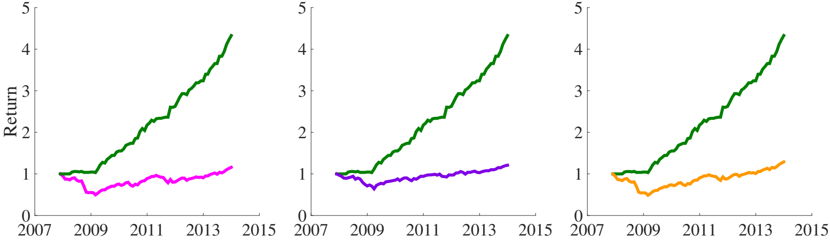


Figure 13: Strategies' cumulative returns' comparison (with transaction costs 1% of portfolio value): TEDAS Hybrid, 60/40, ERC, RR

TEDAS_strategies

Strategies' performance: Mutual funds

Strategy	Cumulative return	Sharpe ratio	Maximum drawdown
TEDAS Basic	421%	0.6393	0.0855
TEDAS Naïve	454%	0.6974	0.0583
TEDAS Hybrid	433%	0.6740	0.0276
RR	116%	0.0214	0.4772
ERC	129%	0.0487	0.4899
60/40	121%	0.0252	0.3473
S&P500	113%	0.0132	0.5037

TEDAS approach:Mutual Funds results

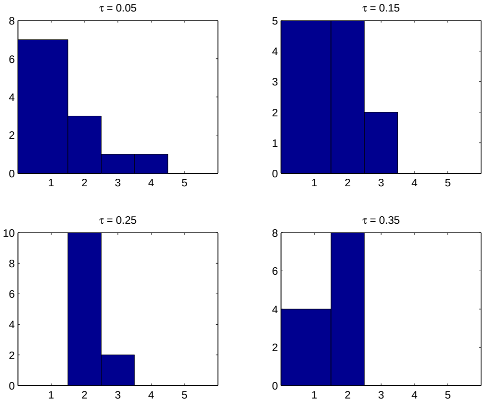


Figure 14: Frequency of the number of selected variables for 4 different τ

TEDAS approach:Mutual Funds results

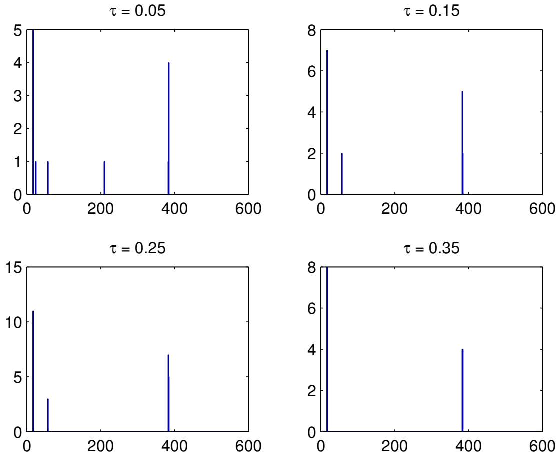


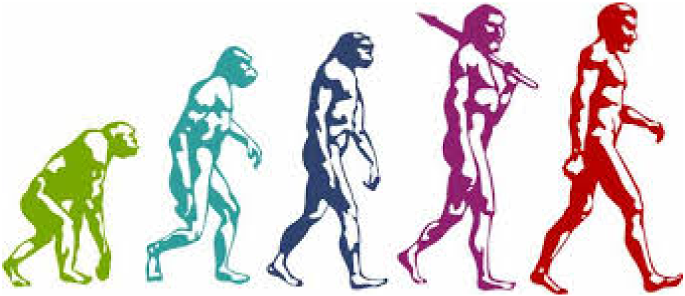
Figure 15: The frequency of mutual funds

Selected Mutual Funds

Table 2: The selected Mutual Funds for $\tau = 0.05$

Top 5 influential Stocks	Frequency	Market
Blackrock Eurofund Class I	12	U.S.
Pimco Funds Long Term United States Government Institutional Shares	8	U.S.
Prudential International Value Fund Class Z	4	U.S.
Artisan International Fund Investor Shares	3	U.S.
American Century 2OTH Century International Growth Investor Class	1	U.S.

How to choose optimal τ-spine?



Generation of different τ-spines

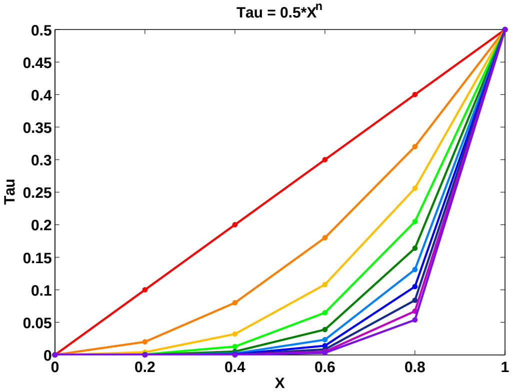


Figure 16: Generation of 10 sets of τ-spines

TEDAS Basic with different τ-spines

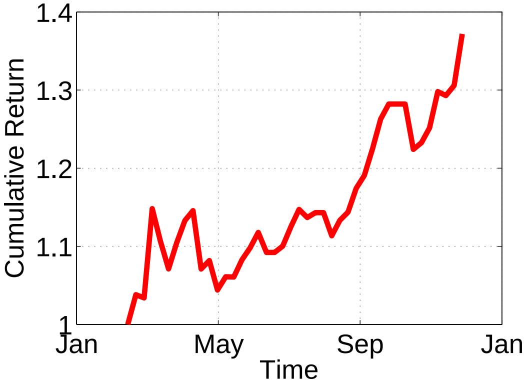


Figure 17: Cumulative return for TEDAS Basic with various τ-spines

TEDAS Basic with different τ-spines

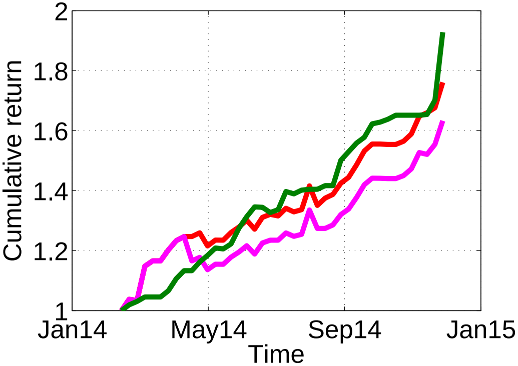


Figure 18: TEDAS Basic cumulative returns' for τ-spines:

$\tau_{j=1,\ldots,5} = (0, 0.002, 0.0233, 0.1311, 0.5)$,

$\tau_{j=1,\ldots,5} = (0.05, 0.15, 0.25, 0.35, 0.5)$,

$\tau_{j=1,\ldots,50} = (0.01, 0.02, 0.03 \ldots 0.49, 0.5)$

What is the best τ-spine?

Monte Carlo simulations

- $Y_i = \hat{q}_{\tau_i}$ $\tau_{j=1,2,3} = (0.05, 0.15, 0.35)$, $n = 100$, $Y_t \sim \mathrm{ALD}(\mu, \sigma, \tau)$; Details
- $X_i \sim \mathrm{N}(0, \Omega)$, $n = 100$ for every τ, $p = 150$, $\beta = (-5, -2, -1, 3, 1, 0.5, 0, ..., 0)$, $\varepsilon_i \sim \mathrm{N}(0, \sigma^2)$; $\lambda_n = 0.25\sqrt{\|\hat{\beta}^{\text{init}}\|_0 \log(n \vee p)}(\log n)^{0.1/2}$, $\hat{\omega}_j = 1/|\hat{\beta}_j^{\text{init}}| \wedge \sqrt{n}$; $\hat{\beta}_j^{\text{init}}$;
- $\Omega_{i,j} = 0.5^{|i-j|}$, $\sigma = 0.1, 0.5, 1$ (three levels of noise);

What is the best τ-spine?

- for $\hat{\beta}^{\text{init}}$ estimator $\hat{\beta}_{\tau,\hat{\lambda}}$ from the model (2) is used, where $\hat{\lambda}$ is chosen according to the BIC criterion

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

- Apply one of TEDAS modification with different τ-spines
- Choose that τ-spine, which gives the highest wealth

$$W_i = \sum_{j=1}^{d} w_j x_{i,\tau} \,,$$

Conclusions

- TEDAS solves TLND challenge
- TEDAS approach performs better than traditional benchmark strategies
- TEDAS outperforms for
  - different regions (global and Germany),
  - various assets
  - alternative time periods (daily, weekly and monthly),
  - big data and small data
- Results for 3 gestalts of TEDAS are robust
- Discussion:
  - How to choose optimal τ-spine?

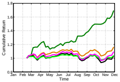
TEDAS - Tail Event Driven ASset Allocation: equity and mutual funds' markets

Wolfgang Karl Härdle

Sergey Nasekin

Alla Petukhina

Ladislaus von Bortkiewicz Chair of Statistics
C.A.S.E. – Center for Applied Statistics
and Economics
Humboldt–Universität zu Berlin
http://lvb.wiwi.hu-berlin.de
http://case.hu-berlin.de



Lasso Shrinkage

Linear model: $Y = X\beta + \varepsilon$; $Y \in \mathbb{R}^n$, $X \in \mathbb{R}^{n \times p}$, $\beta \in \mathbb{R}^p$, $\{\varepsilon_i\}_{i=1}^n$ i.i.d., independent of $\{X_i; i = 1, ..., n\}$

The optimization problem for the lasso estimator:

$$\begin{aligned} \hat{\beta}^{\text{lasso}} = \quad & \arg \min_{\beta \in \mathbb{R}^p} f(\beta) \\ & \text{subject to} \quad g(\beta) \geq 0 \end{aligned} \tag{2}$$

where

$$f(\beta) = \frac{1}{2}(y - X\beta)^\top (y - X\beta)$$
$$g(\beta) = t - \|\beta\|_1$$

where t is the size constraint on $\|\beta\|_1$ Back to "Tail Events"

Lasso Duality

If (1) is convex programming problem, then the Lagrangian is

$$L(\beta,\lambda) = f(\beta) - \lambda g(\beta).$$

and the primal-dual relationship is

$$\underbrace{\underset{\beta}{\text{minimize}} \sup_{\lambda \geq 0} L(\beta,\lambda)}_{\text{primal}} \geq \underbrace{\underset{\lambda \geq 0}{\text{maximize}} \inf_{\beta} L(\beta,\lambda)}_{\text{dual}}$$

Then the dual function $L^*(\lambda) = \inf_{\beta} L(\beta,\lambda)$ is

$$L^*(\lambda) = \frac{1}{2} y^\top y - \frac{1}{2} \hat{\beta}^\top X^\top X \hat{\beta} - t \frac{(y - X\hat{\beta})^\top X\hat{\beta}}{\|\hat{\beta}\|_1}$$

with $(y - X\hat{\beta})^\top X\hat{\beta} / \|\hat{\beta}\|_1 = \lambda$ Back to "Tail Events"

Paths of Lasso Coefficients

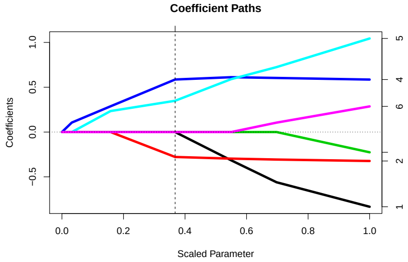


Figure 19: Lasso shrinkage of coefficients in the hedge funds dataset example (6 covariates were chosen for illustration); each curve represents a coefficient as a function of the scaled parameter $\hat{s} = t/\|\beta\|_1$; the dashed line represents the model selected by the BIC information criterion ($\hat{s} = 3.7$)

Example of Lasso Geometry

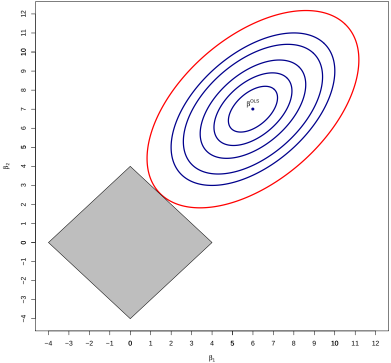


Figure 20: Contour plot of the residual sum of squares objective function centered at the OLS estimate $\widehat{\beta}^{ols} = (6, 7)$ and the constraint region $\sum |\beta_j| \leq t$ MVAlassocontour

Quantile Regression

The loss $\rho_\tau(u) = u\{\tau - \mathbf{I}(u < 0)\}$ gives the (conditional) quantiles $F^{-1}_{y|x}(\tau) \stackrel{\text{def}}{=} q_\tau(x)$.

Minimize

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

Re-write:

$$\underset{(\xi,\zeta)\in\mathbb{R}^{2n}_{+}}{\text{minimize}} \quad \left\{ \tau 1_n^\top \xi + (1-\tau) 1_n^\top \zeta | X\beta + \xi - \zeta = Y \right\}$$

with ξ, ζ are vectors of "slack" variables ▸ Back to "Tail Events"

Non-Positive (NP) Lasso-Penalized QR

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

$$
\begin{aligned}
\underset{(\xi,\zeta,\eta,\tilde{\beta})\in\mathbb{R}_{+}^{2n+p}\times\mathbb{R}^{p}}{\text{minimize}} \quad & \tau 1_n^\top \xi + (1-\tau)1_n^\top \zeta + \lambda 1_n^\top \eta \\
\text{subject to} \quad & \xi - \zeta = Y + X\tilde{\beta}, \\
& \xi \geq 0, \\
& \zeta \geq 0, \\
& \eta \geq \tilde{\beta}, \\
& \eta \geq -\tilde{\beta}, \\
& \tilde{\beta} \geq 0, \qquad \tilde{\beta} \overset{\text{def}}{=} -\beta
\end{aligned}
\tag{3}
$$

Back to "Tail Events"

Solution

Transform into matrix (I_p is $p \times p$ identity matrix; $E_{p\times n} = \begin{pmatrix} I_n \\ 0 \end{pmatrix}$):

$$\begin{aligned} \text{minimize} \quad & c^\top x \\ \text{subject to} \quad & Ax = b, \;\; Bx \leq 0 \end{aligned}$$

where $A = \begin{pmatrix} I_n & -I_n & 0 & X \end{pmatrix}$, $b = Y$, $x = \begin{pmatrix} \xi & \zeta & \eta & \beta \end{pmatrix}^\top$,

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

Back to "Tail Events"

Solution - Continued

The previous problem may be reformulated into *standard form*

$$\begin{array}{ll} \text{minimize} & c^\top x \\ \text{subject to} & Cx = d, \\ & x + s = u, \;\; x \geq 0, s \geq 0 \end{array}$$

and the dual problem is:

$$\begin{array}{ll} \text{maximize} & d^\top y - u^\top w \\ \text{subject to} & C^\top y - w + z = c, \;\; z \geq 0, w \geq 0 \end{array}$$

Back to "Tail Events"

Solution - Continued

The KKT conditions for this linear program are

$$F(x, y, z, s, w) = \left\{ \begin{array}{c} Cx - d \\ x + s - u \\ C^\top y - w + z - c \\ x \circ z \\ s \circ w \end{array} \right\} = 0,$$

with $y \geq 0$, $z \geq 0$ dual slacks, $s \geq 0$ primal slacks, $w \geq 0$ dual variables.

This can be solved by a primal-dual path following algorithm based on the *Newton method*

Back to "Tail Events"

Adaptive Lasso Procedure

Lasso estimates $\hat{\beta}$ can be inconsistent (Zou, 2006) in some scenarios.

Lasso soft-threshold function gives biased results

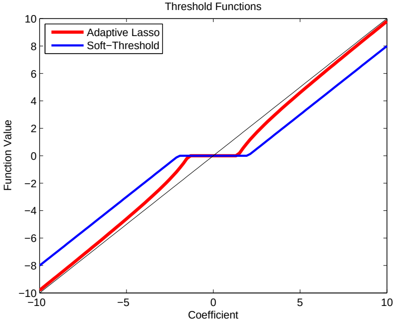


Figure 21: Threshold functions for simple and adaptive Lasso

Adaptive Lasso Procedure

The adaptive Lasso (Zou, 2006) yields a sparser solution and is less biased.

L_1 - penalty replaced by a re-weighted version; $\hat{\omega}_j = 1/|\hat{\beta}_j^{\text{init}}|^{\gamma}$, $\gamma = 1$, $\hat{\beta}^{\text{init}}$ is from (2)

The adaptive lasso estimates are given by:

$$\hat{\beta}_{\lambda}^{\text{adapt}} = \arg \min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} (Y_i - X_i^{\top} \beta)^2 + \lambda \|\hat{\omega}^{\top} \beta\|_1$$

(Bühlmann, van de Geer, 2011): $\hat{\beta}_j^{\text{init}} = 0$, then $\hat{\beta}_j^{\text{adapt}} = 0$

Back to "Tail Events"

Simple and Adaptive Lasso Penalized QR

Simple lasso-penalized QR optimization problem is:

$$\hat{\beta}_{\tau,\lambda} = \arg \min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau(Y_i - X_i^\top \beta) + \lambda \|\beta\|_1 \tag{4}$$

Adaptive lasso-penalized QR model uses the re-weighted penalty:

$$\hat{\beta}_{\tau,\lambda}^{\text{adapt}} = \arg \min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau(Y_i - X_i^\top \beta) + \lambda \|\hat{\omega}^\top \beta\|_1 \tag{5}$$

Adaptive lasso-penalized QR procedure can ensure oracle properties for the estimator Details

Back to "Tail Events"

Algorithm for Adaptive Lasso Penalized QR

The optimization for the adaptive lasso quantile regression can be re-formulated as a lasso problem:

- the covariates are rescaled: $\tilde{X} = (X_1 \circ \hat{\beta}_1^{\text{init}}, \ldots, X_p \circ \hat{\beta}_p^{\text{init}})$;
- the lasso problem (4) is solved:

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

- the coefficients are re-weighted as $\hat{\beta}_{\tau,\lambda}^{\text{adapt}} = \hat{\hat{\beta}}_{\tau,\lambda} \circ \hat{\beta}^{\text{init}}$

Back to "Tail Events"

Oracle Properties of an Estimator

An estimator has oracle properties if (Zheng et al., 2013):

- it selects the correct model with probability converging to 1;
- the model estimates are consistent with an appropriate convergence rate;
- estimates are asymptotically normal with the same asymptotic variance as that knowing the true model

Back to "Simple and Adaptive Lasso Penalized QR"

Oracle Properties for Adaptive Lasso QR

In the linear model, let $Y = X\beta + \varepsilon = X^1\beta^1 + X^2\beta^2 + \varepsilon$, where $X = (X^1, X^2)$, $X^1 \in \mathbb{R}^{n\times q}$, $X^2 \in \mathbb{R}^{n\times(p-q)}$; β^1_q are true nonzero coefficients, $\beta^2_{p-q} = 0$ are noise coefficients; $q = \|\beta\|_0$.

Also assume that $\lambda q/\sqrt{n} \to 0$ and $\lambda/\{\sqrt{q}\log(n \vee p)\} \to \infty$ and certain regularity conditions are satisfied ▸ Details

Oracle Properties for Adaptive Lasso QR

Then the adaptive L_1 QR estimator has the oracle properties (Zheng et al., 2013):

1. Variable selection consistency:

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

2. Estimation consistency: $\|\beta - \hat{\beta}\| = \mathcal{O}_p(\sqrt{q/n})$

3. Asymptotic normality: $u_q^2 \stackrel{\text{def}}{=} \alpha^\top \Sigma_{11} \alpha$, $\forall \alpha \in \mathbb{R}^q$, $\|\alpha\| < \infty$,

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

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

Risk-Return Asset Allocation

Log returns $X_t \in \mathbb{R}^p$:

$$\begin{aligned} \min_{w_t \in \mathbb{R}^p} \quad & \sigma^2_{P,t}(w_t) \stackrel{\text{def}}{=} w_t^\top \Sigma_t w_t \\ \text{s.t.} \quad & \mu_{P,t}(w_t) = r_T, \\ & w_t^\top 1_p = 1, \\ & w_{i,t} \geq 0 \end{aligned} \tag{6}$$

where r_T "target" return, $\Sigma_t \stackrel{\text{def}}{=} E_{t-1}\{(X_t - \mu)(X_t - \mu)^\top\}$, Σ_t is modeled with a GARCH model ▸ Details ▸ Back to "Benchmark Strategies"
▸ Return to "TEDAS Gestalten"

The Orthogonal GARCH Model

- $X_t \in \mathbb{R}^{n \times p}$, $\Gamma_t = B_t \in \mathbb{R}^{p \times p}$ matrix of standardized eigenvectors of $n^{-1} X_t^\top X_t$ ordered according to decreasing magnitude of eigenvalues
- $F_t = P_t \stackrel{\text{def}}{=} X_t \Gamma_t$ PCs of X_t
- factors f, introduce noise u_i, i.e.
  $y_j = b_{j1} f_1 + b_{j2} f_2 + \ldots + b_{jk} f_k + u_i$ or $Y_t = F_t B_t^\top + U_t$
- then $\Sigma_t = \text{Var}(X_t) = \text{Var}(F_t B_t^\top) + \text{Var}(U_t) = B_t \Delta_t B_t^\top + \Omega_t$, $\Delta_t = \text{Var}(F_t)$ diagonal matrix of PC variances at t

Return to "Risk-Return Asset Allocation"

Dynamic Conditional Correlations Model

Assume: $r_t | \mathcal{F}_{t-1} \sim \mathrm{N}(0, D_t R_t D_t)$, $\varepsilon_t = D_t^{-1} r_t$,

$$D_t^2 = \mathrm{diag}(\omega_i) + \mathrm{diag}(\alpha_i) \odot r_{t-1} r_{t-1}^\top + \mathrm{diag}(\beta_i) \odot D_{t-1}^2,$$
$$Q_t = S \odot (11^\top - A - B) + A \odot \{P_{t-1}\varepsilon_{t-1}\varepsilon_{t-1}^\top P_{t-1}\} + B \odot Q_{t-1},$$
$$R_t = \{\mathrm{diag}(Q_t)\}^{-1} Q_t \{\mathrm{diag}(Q_t)\}^{-1}$$

where $r_t \in \mathbb{R}^p$, $D_t = diag(\sigma_{it}) \in \mathbb{R}^{p \times p}$, $\varepsilon_t \in \mathbb{R}^p$ standardized returns with $\varepsilon_{it} \stackrel{\mathrm{def}}{=} r_{it}\sigma_{it}^{-1}$, 1 vector of ones; $P_{t-1} \stackrel{\mathrm{def}}{=} \{\mathrm{diag}(Q_t)\}^{1/2}$, ω_i, α_i, β_i, A, B coefficients, $\odot$ Hadamard (elementwise) product

Return to "TEDAS Gestalten"

The DCC Model - Continued

- correlation targeting: $S = (1/T)\sum_{t=1}^{T} \varepsilon_t \varepsilon_t^\top$
- $Q_0 = \varepsilon_0 \varepsilon_0^\top$ positive definite, each subsequent Q_t also positive definite
- consistent but inefficient estimates: the log-likelihood function

$$L(\theta, \phi) = -\frac{1}{2}\sum_{t=1}^{T}\left\{ n\log(2\pi) + 2\log|D_t| + \log|R_t| + \varepsilon_t^\top R_t^{-1}\varepsilon_t \right\},$$

where θ parameters in D and ϕ additional correlation parameters in R Back

The DCC Model - Continued

Re-write:

$$L(\theta, \phi) = L_V(\theta) + L_C(\theta, \phi),$$

with volatility part $L_V(\theta)$ and correlation part $L_C(\theta, \phi)$,

$$\begin{aligned} L_V(\theta) &= -\frac{1}{2}\sum_{t=1}^{T}\left\{n\log(2\pi) + \log|D_t|^2 + r_t^\top D_t^{-2} r_t\right\} \\ &= -\frac{1}{2}\sum_{t=1}^{T}\sum_{i=1}^{d}\left\{\log(2\pi) + \log(\sigma_{it}^2) + \frac{r_{it}^2}{\sigma_{it}^2}\right\}, \end{aligned}$$

$$L_C(\theta, \phi) = -\frac{1}{2}\sum_{t=1}^{T}\left\{\log|R_t| + \varepsilon_t^\top R_t^{-1}\varepsilon_t - \varepsilon_t^\top \varepsilon_t\right\}.$$

Back

Cornish-Fisher VaR Optimization

Log returns $X_t \in \mathbb{R}^p$:

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

here $W_t \stackrel{\text{def}}{=} W_0 \cdot \prod_{j=1}^{t-1} w_{t-j}^\top (1 + X_{t-j})$, $\tilde{w}$, W_0 initial wealth,

$\sigma_p^2(w) \stackrel{\text{def}}{=} w_t^\top \Sigma_t w_t$,

$$q_\alpha(w) \stackrel{\text{def}}{=} z_\alpha + (z_\alpha^2 - 1)\frac{S_p(w)}{6} + (z_\alpha^3 - 3z_\alpha)\frac{K_p(w)}{24} - (2z_\alpha^3 - 5z_\alpha)\frac{S_p(w)^2}{36},$$

here $S_p(w)$ skewness, $K_p(w)$ kurtosis, z_α is $\mathrm{N}(0,1)$ α-quantile If $S_p(w)$, $K_p(w)$ zero, then obtain Markowitz allocation

▸ Return to "TEDAS Gestalten"

Risk Parity (Equal risk contribution)

Let $\sigma(w) = \sqrt{w^\top \Sigma w}$. Euler decomposition:

$$\sigma(w) \stackrel{\text{def}}{=} \sum_{i=1}^{n} \sigma_i(w) = \sum_{i=1}^{n} w_i \frac{\sigma(w)}{\partial w_i}$$

where $\frac{\sigma(w)}{\partial w_i}$ is the marginal risk contribution and $\sigma_i(w) = w_i \frac{\sigma(w)}{\partial w_i}$ the risk contribution of i-th asset. The idea of ERC strategy is to find risk balanced porfolio, such that:

$$\sigma_i(w) = \sigma_j(w)$$

i.e. the risk contribution is the same for all assets of the portfolio

Return to "Benchmark Strategies"

60/40 allocation strategy

60/40 portfolio allocation strategy implies the investing of 60% of the portfolio value in stocks (often via a broad index such as S&P500) and 40% in government or other high-quality bonds, with regular rebalancing to keep proportions steady.

Return to "Benchmark Strategies"

Regularity Conditions for Adaptive Lasso QR

A1 Sampling and smoothness: $\forall x$ in the support of X_i, $\forall y \in \mathbb{R}$, $f_{Y_i|X_i}(y|x)$, $f \in \mathcal{C}^k(\mathbb{R})$, $|f_{Y_i|X_i}(y|x)| < \overline{f}$, $|f'_{Y_i|X_i}(y|x)| < \overline{f'}$; $\exists \underline{f}$, such that $f_{Y_i|X_i}(x^\top \beta_\tau | x) > \underline{f} > 0$

A2 Restricted identifiability and nonlinearity: let $\delta \in \mathbb{R}^p$, $T \subset \{0, 1, ..., p\}$, δ_T such that $\delta_{Tj} = \delta_j$ if $j \in T$, $\delta_{Tj} = 0$ if $j \notin T$; $T = \{0, 1, ..., s\}$, $\overline{T}(\delta, m) \subset \{0, 1, ..., p\} \backslash T$, then $\exists m \geq 0$, $c \geq 0$ such that

$$\inf_{\boldsymbol{\delta} \in A, \boldsymbol{\delta} \neq 0} \frac{\delta^\mathsf{T} \operatorname{E}(X_i X_i^\top)\delta}{\|\delta_{T \cup \overline{T}(\delta, m)}\|^2} > 0, \qquad \frac{3\underline{f}^{3/2}}{8\overline{f}'} \inf_{\delta \in A, \delta \neq 0} \frac{\operatorname{E}[|X_i^\mathsf{T}\delta|^2]^{3/2}}{\operatorname{E}[|X_i^\mathsf{T}\delta|^3]} > 0,$$

where $A \overset{\text{def}}{=} \{\delta \in \mathbb{R}^p : \|\delta_{T^c}\|_1 \leq c\|\delta_T\|_1, \|\delta_{T^c}\|_0 \leq n\}$

Back

Regularity Conditions - Continued

A3 Growth rate of covariates:

$$\frac{q^3\{\log(n \vee p)\}^{2+\eta}}{n} \to 0, \eta > 0$$

A4 Moments of covariates: Cramér condition

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

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

A5 Well-separated regression coefficients: $\exists b_0 > 0$, such that $\forall j \leq q$, $|\hat{\beta}_j| > b_0$

Asymetric Laplace Distribution

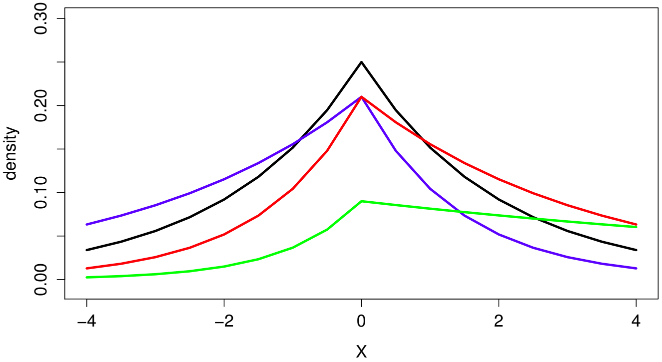


Figure 22: Standart ALD: $\tau = 0.3$,$\tau = 0.5$,$\tau = 0.7$, $\tau = 0.1$

Quantile regression using ALD

- Yu & Moyeed(2001)
  $Y_i \sim \text{ALD}(\mu, \sigma, \tau)$, if its pdf is given by

$$f(y|\mu, \sigma, \tau) = \frac{\tau(1-\tau)}{\sigma} exp \left\{ \rho_\tau \frac{(y-\mu)}{\sigma} \right\}$$

  where μ is location, σ - scale and τ-skewness parameters, and loss function $\rho_\tau(u) = u\{\tau - \mathbf{I}(u < 0)\}$

- Sanches et. al (2013)

$$y_i = x_i^\top \beta_\tau + \varepsilon_i, \; i = 1, \dots, n$$

  Re-write:

$$Y_i | U_i = u_i \sim \text{N}(x_i \beta_\tau + \theta u_i, p_\tau^2 \sigma u_i)$$
$$U_i \sim \text{Exp}(\sigma), \; i = 1, \dots, n$$

  here $\theta = \frac{1-2\tau}{\tau(1-\tau)}$ and $p_\tau^2 = \frac{2}{\tau(1-\tau)}$ Back to "Choice of τ-spine"

References

W. K. Härdle, S. Nasekin, D. K. C. Lee , K. F. Phoon

TEDAS - Tail Event Driven Asset Allocation

SFB 649 Discussion Paper 2014-032: Jun, 2014

L. B. Sanchez V. H. Lachos, F. V. Labra

Likelihood Based Inference for Quantile Regression Using the Asymmetric Laplace Distribution

Discussion Paper, Universidade Estadual de Campinas: Sep, 2013