Regular(ized) Hedge Fund Clones

Daniel Giamouridis
Dept. of Accounting and Finance, Athens Univ. of Economics and Business,GRE

Sandra Paterlini
Dept. of Economics, Univ. of Modena and Reggio E., IT
Center for Quantitative Risk Analysis, Dept. of Statistics, LMU Munich, GER

support from CAREFIN-Bocconi Centre for Applied Research is gratefully acknowledged

Outline

- Why Synthetic Hedge Funds?
- How can we build syntetic hedge funds?
  - Factor Models
- Regularization Methods
  - Ridge Regression
  - LASSO
- Empirical Analysis
- Conclusion and further research

Introduction

Why Hedge Funds (HF) are special?

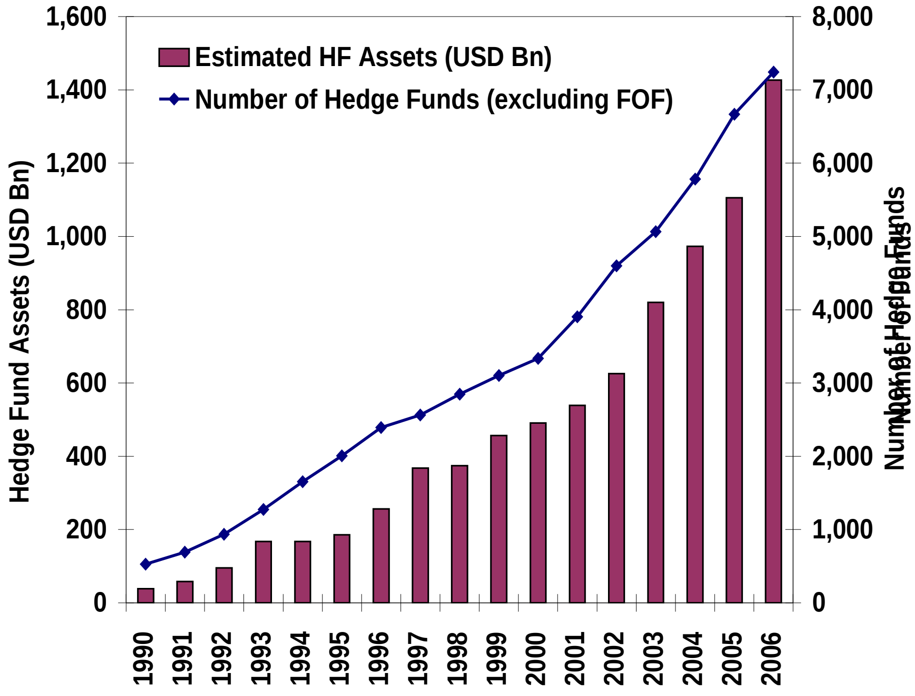
- Flexible w.r.t. the types of securities they hold and the type of positions they take
- Not subject to public disclosure of their activities
- Not evaluated against passive benchmark
- HF managers construct highly dynamic, complex trading strategies
- Returns on alternative investments has shown little correlation with returns on traditional assets
- In many cases, HF returns have also been significantly high

Introduction

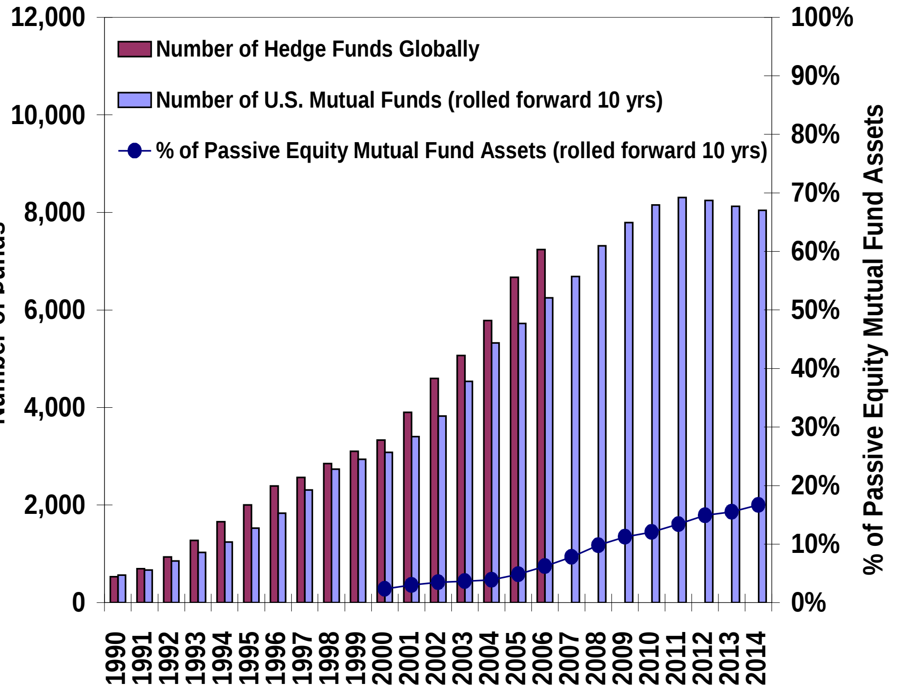
A. Rapid growth in the HF Industry

B. Similar trend in the long-only industry gave rise to passive management

(A)



(B)

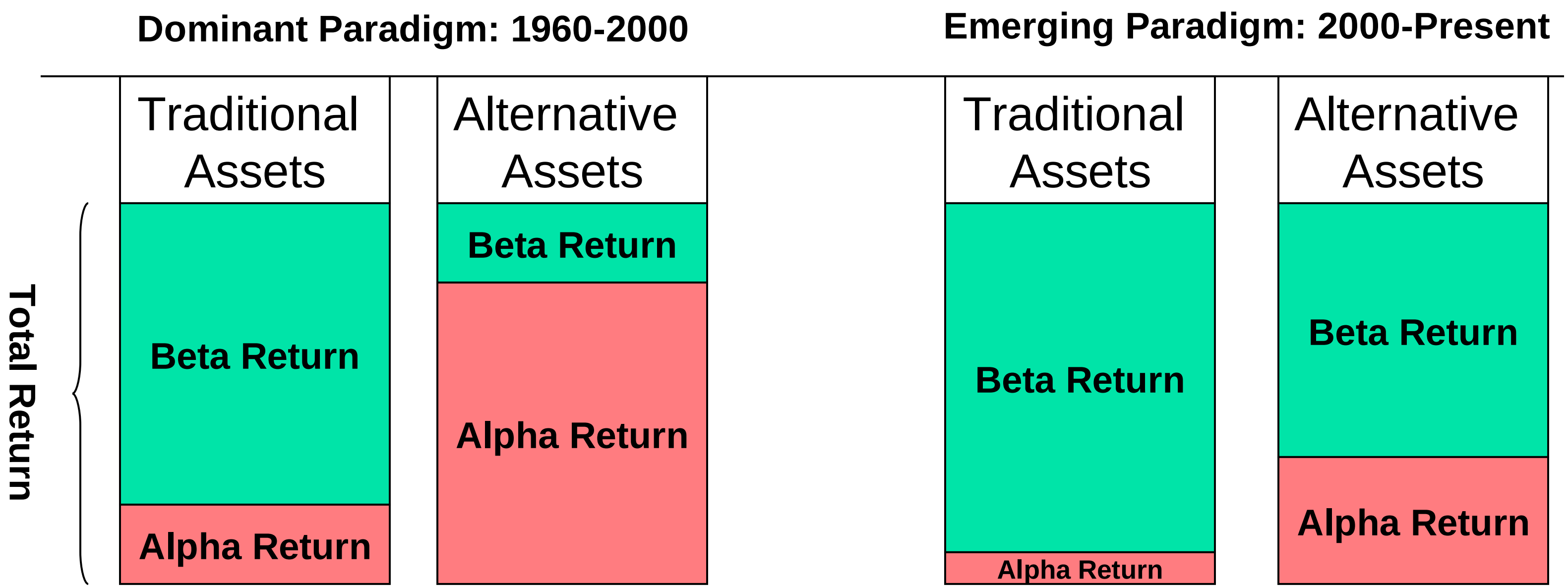


Sources: HFR Industry Reports © HFR Inc. Q4 2006, www.hedgefundresearch.com, ICI, Lipper

Introduction

Betas also drive returns

- Schneeweis and Kazemi (2007) argue that:



Introduction

What is "Alternative Beta"?

- Alternative Beta refers to alternative systematic risks (including credit risk, volatility risk, the small company effect, and so on) in the context of Markowitz' portfolio theory. Systematic risks are compensated through risk premia (the expected rate of returns above the risk-free interest rate)
- The return of an investment can be decomposed into the contribution from risk exposure (risk premium) and one resulting from skill-based investing (alpha or outperformance)
- Alternative beta extends the idea of traditional passive investing into the alternative investment space

Introduction

What is "Alternative Beta"?

- Need to redefine the meaning of beta to anything that can be generated using a rules-based methodology
- Then, if it doesn't require an active manager to continually make investment decisions, it is ***beta***, not ***alpha***
- One of the key implications of "alternative beta" theory is that it is possible to create "synthetic hedge funds", which replicate the alternative beta exposures of the hedge fund universe

Introduction

Skill or alternative betas?

- Agarwal and Naik (2000) and Fung et al. (2006) conclude that the alpha of the average hedge fund manager is very poor and not persistent

- Hedge fund performance - on average – is primarly attributed to 'alternative betas' rather than skill

- Can synthetic hedge funds be created at a lower cost for investors?

Why not clone HF returns?

If successful…

- Benefits from lower cost and greater transparency
- Liquidity
- No barriers to entry
- Permits separation of hedge fund alpha from hedge fund beta (avoid paying alpha fees for beta returns)
- Elimination of single-manager risk and style drift

Hedge Fund Clones

In practice…

- Two styles of passive hedge fund strategies are emerging
  - **<u>Factor Approach</u>**:

    combinations of liquid assets designed to replicate broad hedge fund benchmarks

  - **<u>Mechanical Trade Replication</u>**: systematically replicating popular hedge-fund strategies such as merger, convertible bond, or volatility arbitrage

Factor Approach

The linear model

Let:

$\mathbf{r}^{HF}$	*(Tx1 vector)* hedge fund´s return TS
F	*(Txp matrix)* factors returns TS
β	*(px1 vector)* hedge fund’s clone portfolio weights
$\mathbf{r}^{HFClone}=\mathbf{F}\boldsymbol{\beta}$	*(Tx1 vector)* hedge fund´s clone return TS

$$\underset{\boldsymbol{\beta}}{minimize} \quad f(\boldsymbol{\beta}) = \sum_{t=1}^{T} (\mathbf{r}_t^{HF} - \mathbf{r}_t^{HFClone})^2$$

$$subject\ to \ \sum_{i=1}^{n} \beta_i = 1$$

Factor Approach

Open issues

- How many factors? Which factors?
- How do we deal with correlated factors?
- Can we improve the out-of- sample performance?

Linear Models and Regularization Methods

Factor Models

Model Selection and Estimation

- Ideally, we would like a factor model, which
  - provides accurate prediction
  - has a small number of factors, easily interpretable
  - provides stable estimates

Factor Models

The "Simple" OLS Linear Regression

Let:

$\mathbf{y} = [y_1,..y_n]$: (nx1 vector) dependent variable (i=,…,n)

$\mathbf{X} = [x_{ik}]$: (nxp matrix) independent variables (i=1,..,n,k=1,..,p)

$\boldsymbol{\beta} = [\beta_{ik}]$: parameters to be estimated

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{minimize} \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{p} x_{ij} \beta_{\mathbf{j}} \right)^2 = \underset{\boldsymbol{\beta}}{minimize} (\mathbf{y} - \mathbf{x}\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{x}\boldsymbol{\beta})$$

$$\Rightarrow \hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$$

Gauss - Markov: **B**est **L**inear **U**nbiased **E**stimator

Factor Models

Problems

- What happens if **X** are highly correlated?
  - The matrix $\mathbf{X}^T\mathbf{X}$ may be nearly singular (ill-conditioned problem)

- Have OLS estimates good prediction accurancy?
  - They have low bias but high variance

- How many factors? Which factors?
  - Model selection via AIC, BIC, stepwise, etc.

Factor Models

A possible solution

- L*q*-regularization methods or continuous shrinkage estimators
- Idea: control the variability of the estimator by regularization. We shrink the betas by trading variance for bias, such that we get a sparse model – a model with few independent variables

Regularization Methods

Ridge Regression (Tikhonov, 1963, Hoerl and Kennard, 1970)

- Add L_2-penalty

$$\boldsymbol{\beta}^{Ridge} = \underset{\boldsymbol{\beta}}{minimize} \sum_{i=1}^{n} \left(y_i - \sum_{k=1}^{p} x_{i,k} \beta_k \right)^2 + \lambda \sum_{k=1}^{p} \beta_k^{\ 2} \quad \text{where } \lambda \geq 0$$

which is equivalent to

$$\boldsymbol{\beta}^{Ridge} = \underset{\boldsymbol{\beta}}{minimize} \sum_{i=1}^{n} \left(y_i - \sum_{k=1}^{p} x_{i,k} \beta_k \right)^2 \quad \text{subject to} \sum_{k=1}^{p} \beta_k^{\ 2} \leq t$$

- Idea: add a positive constant to the diagonals of $\mathbf{X}^T\mathbf{X}$ such that the matrix $\mathbf{X}^T\mathbf{X}$ is non-singular

  ➡ computing the inverse does not affect the stability of the minimization problem

Regularization Methods

Ridge Regression

- Solution:

$$\hat{\boldsymbol{\beta}}^{ridge} = \left(\mathbf{X}^{\mathrm{T}}\mathbf{X} + \lambda\mathbf{I}\right)^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}$$

  where **I** is the *pxp* identity matrix
- The optimization problem is still a continuously differentiable optimization problems and it can be easily tackled by conventional techniques
- Ridge regression provides more stable and accurate estimates than standard residual sum of squares minimization

Regularization Methods

Ridge Regression

- Ridge regression shrinks the coordinates with respect to the orthonormal basis formed by the principal components
- The coordinate with respect to the principal component with a smaller variance is shrunk more. The underlying hypotesis is that **y** varies most in the directions of high variance

Regularization Methods

Ridge Regression

Pro:

- can be interpreted as prior knowledge that **β** should not be too large (statistics)
- can be interpreted as balancing large values of the variables compared to meeting the target (optimal design)
- can improve the prediction accuracy of the model (large **β** values can cause large variation in $\mathbf{X}$ and may not achieve high predictive performance) (prediction)
- gives a compromise between solving the system and having a small **β** (optimization)

Regularization Methods

Ridge Regression

<u>Cons</u>:

- Ridge regressions does not set any coefficient to zero=> No easily interpretable model

Regularization Methods

LASSO (Tibshirani, 1996)

- Add L1-penalty

$$\boldsymbol{\beta}^{Lasso} = \underset{\boldsymbol{\beta}}{minimize} \sum_{i=1}^{n} \left(y_i - \sum_{k=1}^{p} x_{i,k} \beta_k \right)^2 + \lambda \sum_{k=1}^{p} |\beta_k| \text{ where } \lambda \geq 0$$

which is equivalent to

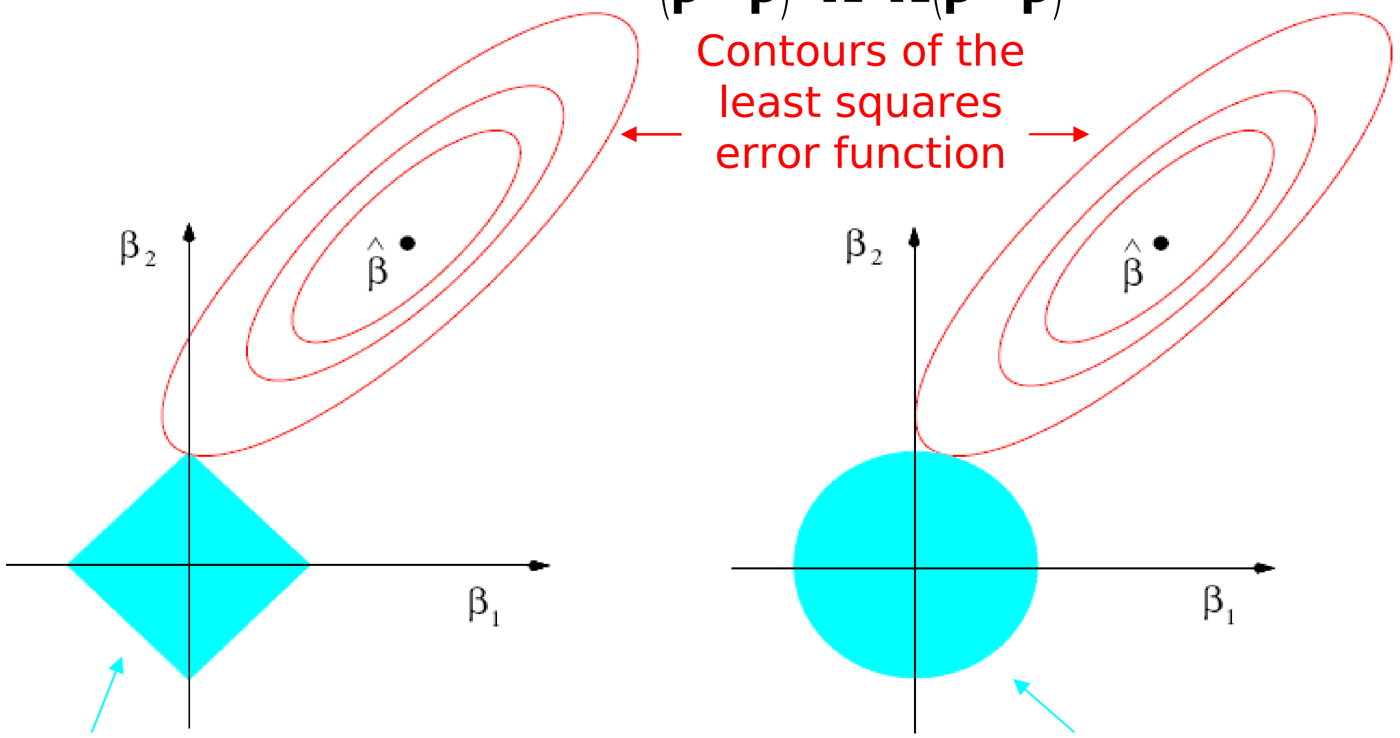
$$\boldsymbol{\beta}^{Lasso} = \underset{\boldsymbol{\beta}}{minimize} \sum_{i=1}^{n} \left(y_i - \sum_{k=1}^{p} x_{i,k} \beta_k \right)^2 \text{ subject to } \sum_{k=1}^{p} |\beta_k| \leq t$$

Regularization Methods

LASSO (L_1)

Ridge (L_2)

$$(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})^T \mathbf{X}^T \mathbf{X}(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})$$



$$|\beta_1| + |\beta_2| \le t$$

Constraint Regions

$$\beta_1^2 + \beta_2^2 \le t$$

Regularization Methods

LASSO

<u>Pro:</u>

- L_1-penalty is still a continuous function of **β**
- When λ is large, some β_k will be ***exactly*** 0 (no approx 0) => model selection and estimation in a single step
- LASSO tends to promote **sparse** and **stable** models that can be more easily interpretable
- Nice asymptotic properties

Regularization Methods

LASSO

<u>Cons</u>:

- The optimization problem is non-differentiable and in general there is no closed form solution for the global minimum
- The choice of λ can be difficult
- Inconstency for variable selection in some instances (adaptive lasso)

LASSO

How do we estimates β?

- Even if the problem is non-differentiable, there are fast and accurate solvers. (see Schmidt et al. 2007 for comparisons and MATLAB code)

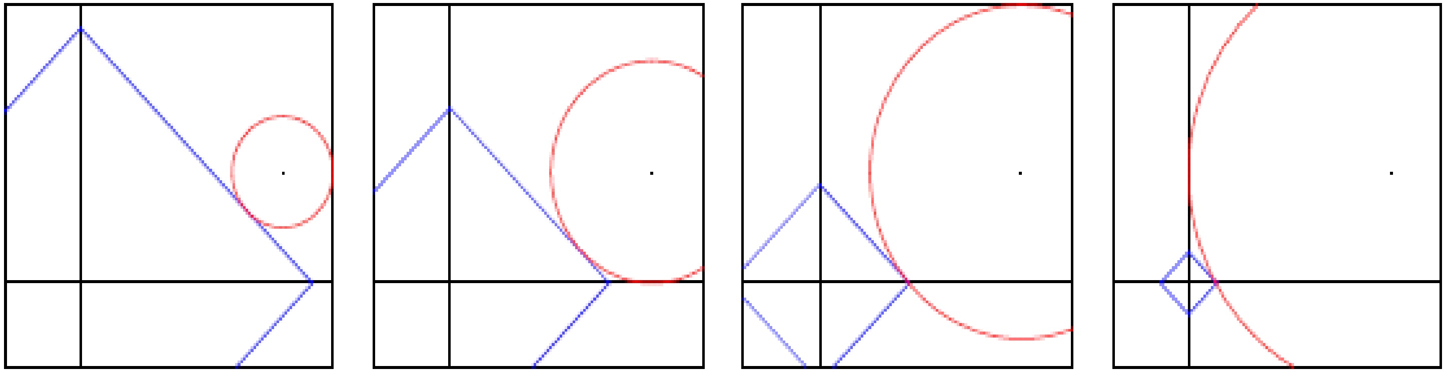
  - Quadratic Programming , non negative variables
  - LARS
  - Iterative Soft Thresholding
  - Block Coordinate
  - Gauss-Seidel
  - Grafting
  - Iterate Ridge
  - …

http://www.cs.ubc.ca/~schmidtm/Software/lasso.html

LASSO

What is the effect of λ?

$$\boldsymbol{\beta}^{Lasso} = \underset{\boldsymbol{\beta}}{minimize} \left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right)^T \left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right) + \lambda \,|\, \boldsymbol{\beta} \,|_1$$



(a) Very small ℓ_1 penalty: tangency on edge.

(b) Small ℓ_1 penalty: tangency on edge, nearing vertex.

(c) Moderately sized ℓ_1 penalty: tangency reaches vertex.

(d) Large ℓ_1 penalty: tangency remains at vertex, moves toward origin along axis.

Source: De Mol et al. 2008

- The larger λ, the sparser the model is (the smaller the number of selected independent variables)

LASSO

How do we choose λ?

- Cross-validation
- Generalized Cross-Validation
- Analytical risk
- ….
- Depending on the problem settings and targets
  - e.g. choose λ to maximise last period return

LASSO in Finance

Adding the L_1 penalty…

- promotes SPARSITY: few factors are selected
- stabilizes the problem => we reduces the sensitivity of the optimization to possible collinearities between the independent variables

LASSO in Finance

Adding the L_1 penalty…

- allows to account for transaction costs in a natural way

$$\boldsymbol{\beta}^{Lasso} = \underset{\boldsymbol{\beta}}{minimize} \sum_{i=1}^{n} \left(y_i - \sum_{k=1}^{p} x_{i,k} \beta_k \right)^2 + \lambda \sum_{k=1}^{p} s_k \, | \beta_k | \; \leq t$$

where s_k is the bid-ask spread for the *k*-th security

- If we assume the bid-ask spread is the same for all assets and constant for a wide range of transactiom size, then the L_1-penalty account for the transaction cost
- Otherwise, if the bid-ask spread are different, we can deal with the weighted penalty and the optimization will tend to prefer to invest in more liquid securities (low s_k), while avoiding less liquid ones (high s_k)

LASSO in Finance

Adding the L_1 penalty…

- regulates the total amount of shorting

$$\boldsymbol{\beta}^{Lasso} = \underset{\boldsymbol{\beta}}{minimize} \sum_{i=1}^{n} \left(y_i - \sum_{k=1}^{p} x_{i,k} \beta_k \right)^2 \text{ subject to } \sum_{k=1}^{p} | \beta_k | \le t$$

$$\text{and subject to } \sum_{k=1}^{p} \beta_k = 1$$

Let
$$\begin{cases} \beta_i^+ = \beta_i \;\; and \;\; \beta_i^- = 0 \;\; if \;\; \beta_i > 0 \\ \beta_i^+ = 0 \; and \;\; \beta_i^- = \beta_i \;\; if \;\; \beta_i < 0 \\ \beta_i^+ = 0 \; and \;\; \beta_i^- = 0 \;\; if \;\; \beta_i = 0 \end{cases}$$

Hence, $\sum_{k=1}^{p} | \beta_k | \le t$ is equivalent to $-\sum_{i=1}^{N} \beta^- \le \frac{t-1}{2}$

LASSO in Finance

No short sellling …

- Notice that imposing no short selling with budget constraint (Sharpe's style analysis) is a special case of LASSO

$$\boldsymbol{\beta}^{Lasso} = \underset{\boldsymbol{\beta}}{minimize} \sum_{i=1}^{n} \left(y_i - \sum_{k=1}^{p} x_{i,k} \beta_k \right)^2$$

subject to

$$\beta_k \geq 0$$

$$\sum_{k=1}^{p} \beta_k = 1$$

Jagannathan, R., and T. Ma, 2003, Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps," Journal of Finance, 58, 1651-1684.

LASSO in Finance…

Previous Work

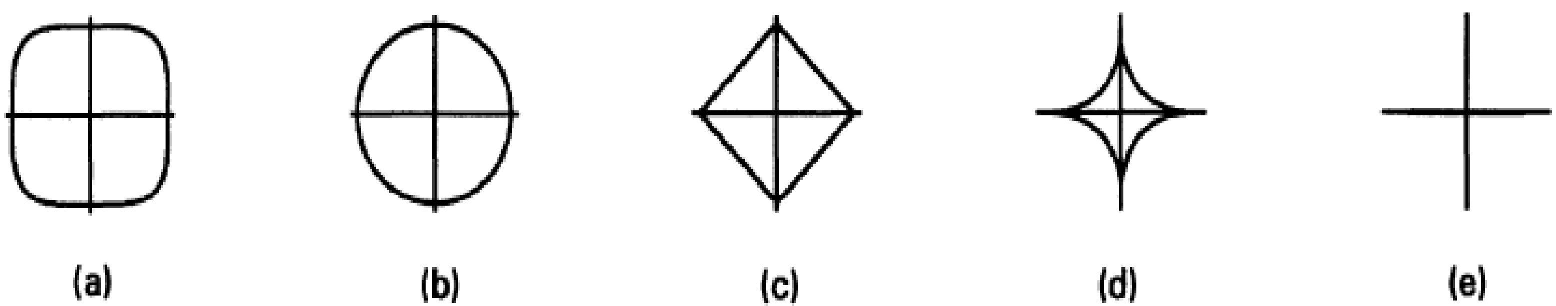
- Mean-Variance Framework
  - Brodie, J., I. Daubechies, C. De Mol, and D. Giannone, 2007, Sparse and Stable Markowitz Portfolios, ECORE Discussion paper 2007/61.
  - DeMiguel, V., L. Garlappi, J. Nogales, and R. Uppal, 2007, A Generalized Approach to Portfolio Optimization: Improving Performance By Constraining Portfolio Norms, Working Paper.
  - Welsch R.E. and X. Zhou, 2007, Application of Robust Statistics to asset allocation models, REVSTAT – Statistical Journal, 5, 1, 97–114.
- Transaction costs
  - Lobo, M.S., Fazel, M. and Boyd, S. (in press), Portfolio optimization with linear and fixed transaction costs, Annals of Operation Research, DOI 10.1007/s10479-006-0145-1.

Regularization Methods

(B)ridge Regression (Frank and Friedman, 1993)

- Add L_q-penalty proportional to the q-norm of β

$$\boldsymbol{\beta}^{(B)ridge} = \underset{\boldsymbol{\beta}}{minimize} \sum_{i=1}^{n}\left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 + \lambda|\boldsymbol{\beta}|_q \text{ where } |\boldsymbol{\beta}|_q = \left(\sum_{k=1}^{p}|\beta_k|^q\right)^{\frac{1}{q}}$$



$q = 4$ $q = 2$ $q = 1$ $q = 0.5$ $q = 0.1$

Regularization Methods

Further research…

- Very active research field nowadays
  - Adaptive lasso
  - Elastic Net
  - SCAD (Smoothly Clipped Absolute Deviation Penalty)
  - …

Empirical Results

The Data

- Seven monthly hedge fund index return times series from the Hedge Fund Research Database. We have analysed more indexes
- 20 factors from Bloomberg and Datastream
- 4 commercial products
  - MLHFFCEU Index - Merril Lynch
  - DBARB Index - Deutsche Bank
  - CAPHTRU Index – Citigroup
  - ARTIUSD Index – Goldman Sachs
- Monthly data (Period: 28/02/1990-30/06/2008)

The HFR Indexes

EH	Equity Hedge	Investment Managers who maintain positions both long and short in primarily equity and equity derivative securities.
ED	Event-Driven	Investment Managers who maintain positions in securities of companies currently or prospectively involved in corporate transactions of a wide variety
M	Macro (Total) Index	Investment Managers which execute a broad range of strategies in which the investment process is predicated on movements in underlying economic variables and the impact these have on equity, fixed income, currency and commodity markets.
RV	Relative Value	Investment Managers who maintain positions in which the investment thesis is predicated on realization of a valuation discrepancy in the relationship between multiple securities
EM	Emerging Markets	The constituents of the HFRI Emerging Markets Indices are selected according to their Regional Investment Focus only (Asia ex-Japan Index, Global Index, Russia/Eastern Europe Index, Latin-America Index). There is no Investment Strategy criteria for inclusion in these indices.
FC	Fund Weighted Composite Index	Includes over 2000 constituent funds, includes both domestic and offshore funds, no Fund of Funds included in Index
FF	Fund of Funds Composite Index	Includes over 800 constituent funds, includes both domestic and offshore funds, only Fund of Funds included in Index

The Factors

Equity	Bonds	Commodities	Currency	Real Estate/Equity
Standard & Poors 500 Total Return	Citigroup Treasury	S&P GSCI	Dollar Index	FTSE EPRA/NAREIT Global Real Estate
Russell 2000	Citigroup World Government 1-3Y	S&P GSCI Gold		
Russell 2000 Total Return	Citigroup World Government 7-10Y	S&P GSCI Crude Oil		
Dow Jones Industrial Average	Broad Investment-Grade Bond			
MSCI World Equity	Cash			
MSCI EAFE Total Return				
MSCI Emerging Markets				
Bovespa Brazil Ibovespa				
Hang Seng				
CBOE Implied Volatility				

Experimental Set-up

The models

- The simple linear model
- The HF clone model

$$\underset{\boldsymbol{\beta}}{minimize} \quad f(\boldsymbol{\beta}) = \sum_{t=1}^{T} (\mathbf{r}_t^{HF} - \mathbf{r}_t^{HFClone})^2 + \lambda \sum_{i=1}^{n} |\beta_i| \quad \boldsymbol{\beta} \in \Re^n$$

subject to

$$(1) \ \sum_{i=1}^{n} \beta_i = 1$$

$$(2) \ -1 \leq \beta_i \leq 1$$

$$where \ \mathbf{r}^{HFClone} = \mathbf{F}\boldsymbol{\beta}$$

The Simple Linear Model

The Simple Linear Model

The Simple Linear Model

- Rolling Window of {60, 120} obs, stepsize=1 obs
- Compare
  - OLS Regression
  - Ridge Regression
  - LASSO (λ=0.5, λ=2, λ chosen by cross-validation)
  - Exhaustive Model Selection via AIC or BIC

The Simple Linear Model

Main Results

- All the methods have comparable out-of-sample mean squared errors
- LASSO selects models comparable in size with AIC, BIC detects more parsimonious models
- LASSO provides more stable and less extreme estimates
- LASSO with λ chosen by cross-validation increases the instability in the estimates, without improving the accuracy and parsimony of the model

The Simple Linear Model

The ten “most important” factors

- Russell 2000 Total Return Index,
- MSCI Emerging Markets Index,
- the Dollar Index
- S&P GSCI Crude Oil Index
- CBOE Implied Volatility Index
- Bovespa Brazil Ibovespa Index
- Hang Seng Index
- Cash
- S&P GSCI Gold Index
- FTSE EPRA/NAREIT Global Real Estate Index

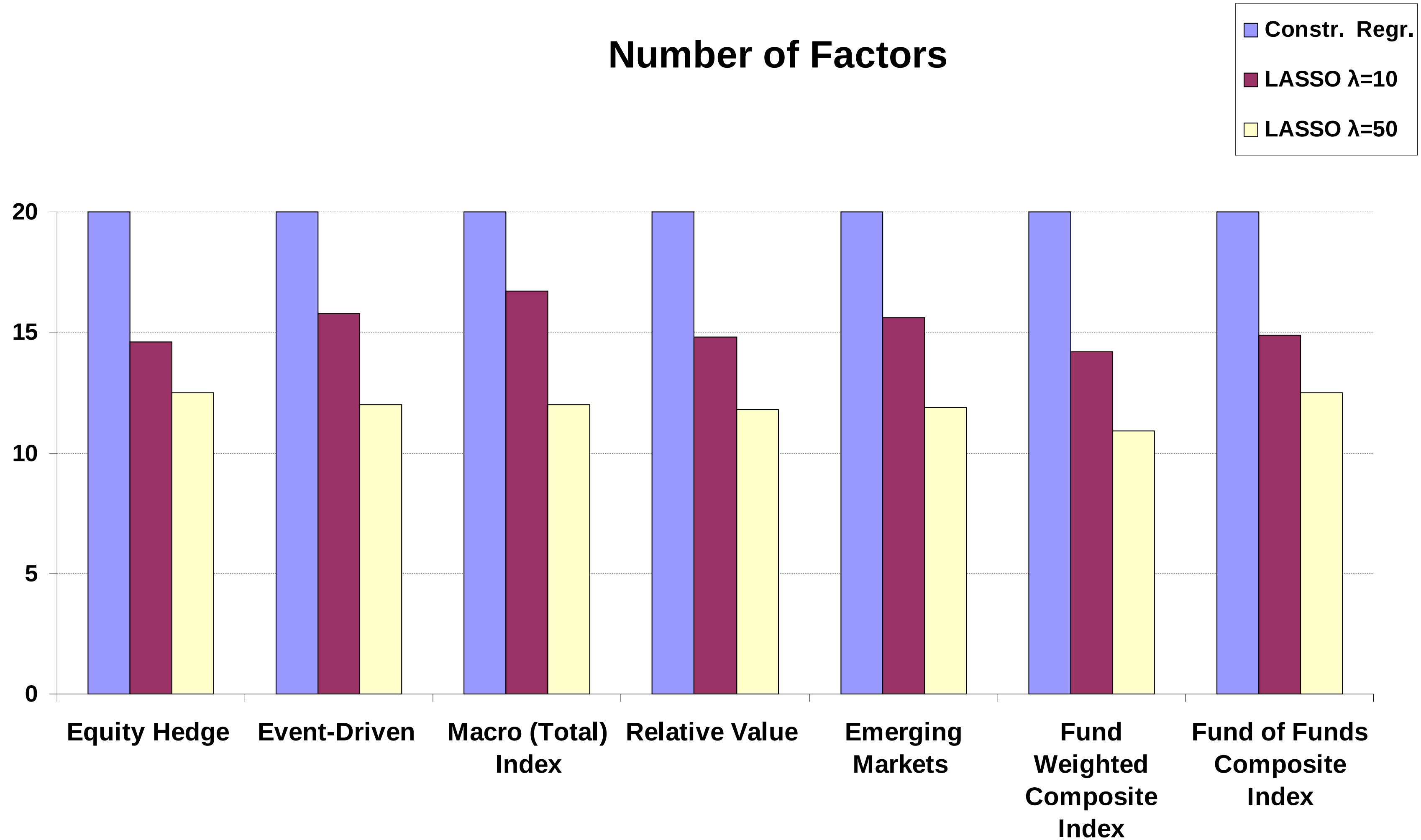
The HF Clone Model

The HF Clone model

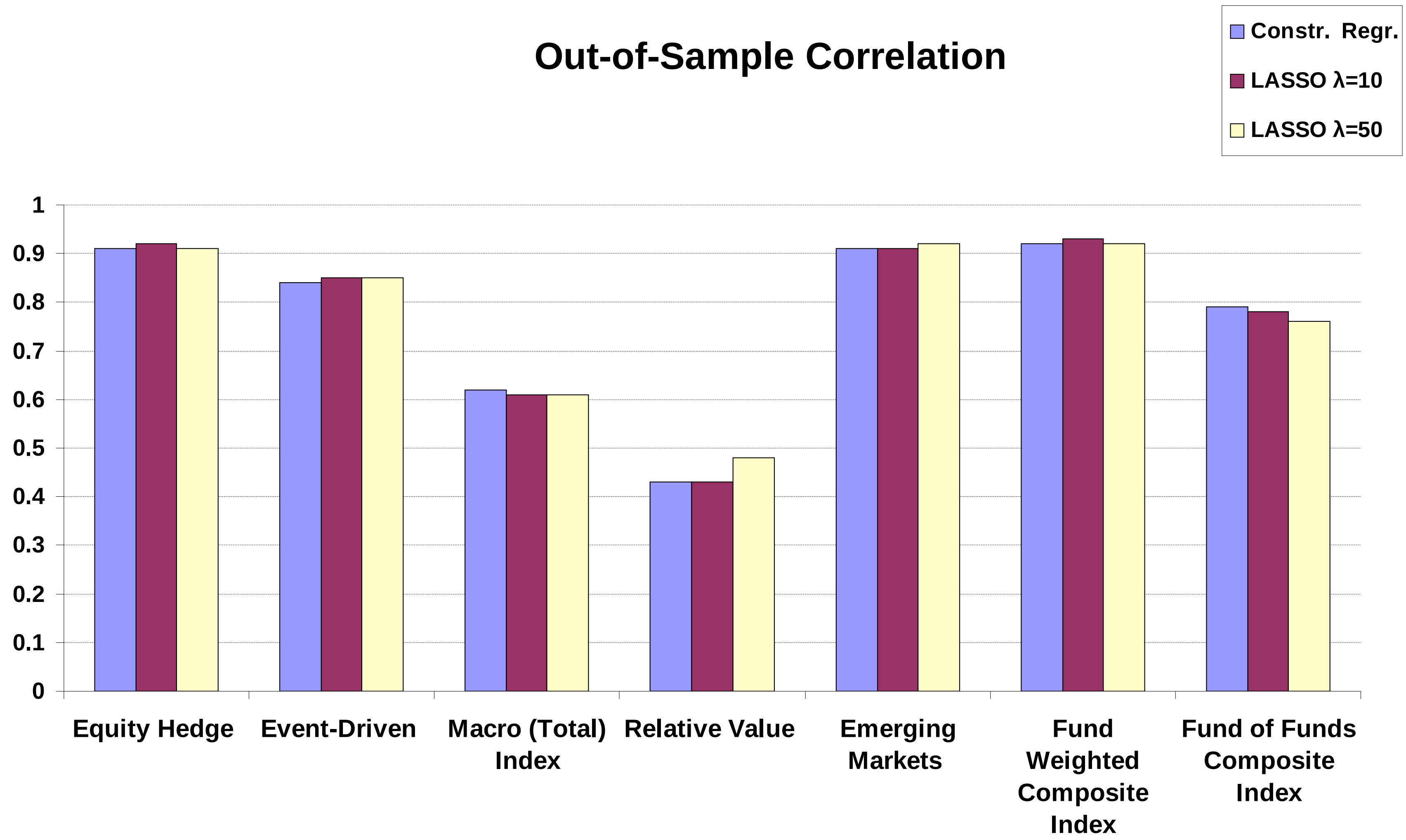
The HF Clone Model

- Rolling Window of {60, 120} obs, stepsize=1 obs
- Results for $\lambda=10$ (low), $\lambda=50$ (high), λ chosen by cross-validation
- Comparison with constrained model ($-1 \leq \beta \leq 1$, $\Sigma\beta=1$) without penalty on the norm of the asset weights (constrained regression - QP)
- Comparison with
  - MLHFFCEU Index - Merril Lynch (replicate FC)
  - DBARB Index - Deutsche Bank (replicate FF)
  - CAPHTRU Index – Citigroup (replicate FF)
  - ARTIUSD Index – Goldman Sachs (unknown)

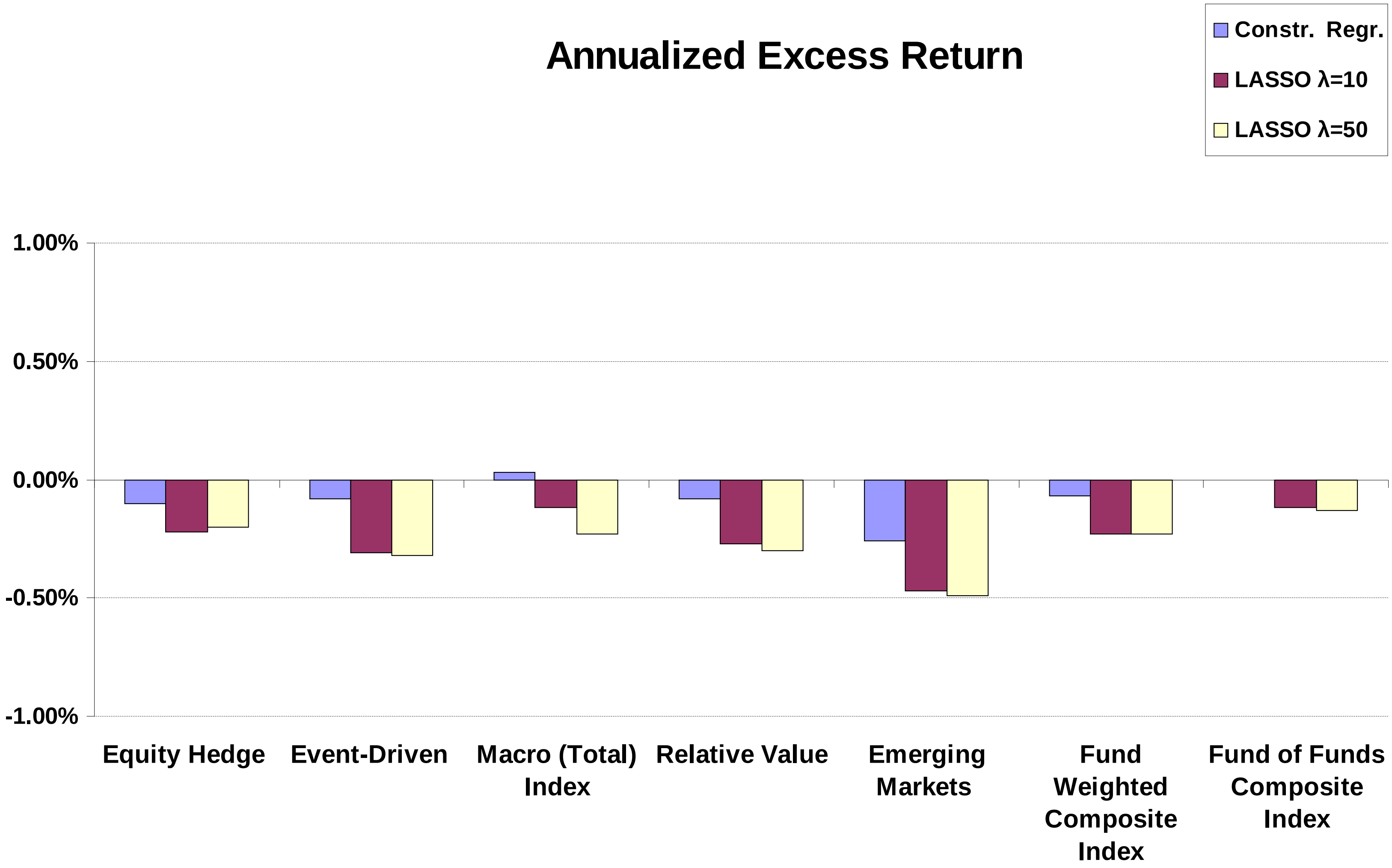
Number of Factors



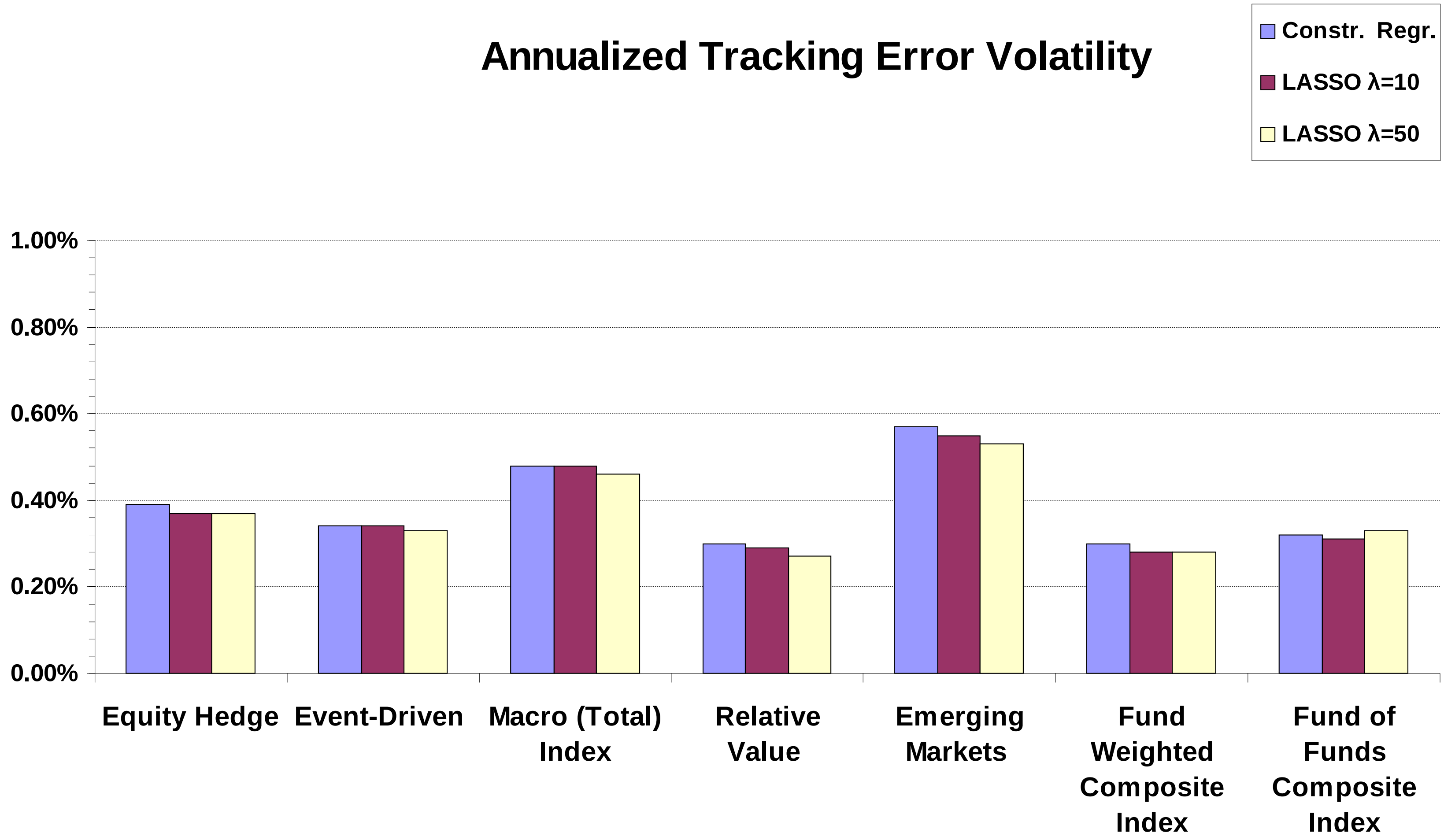
Out-of-Sample Correlation



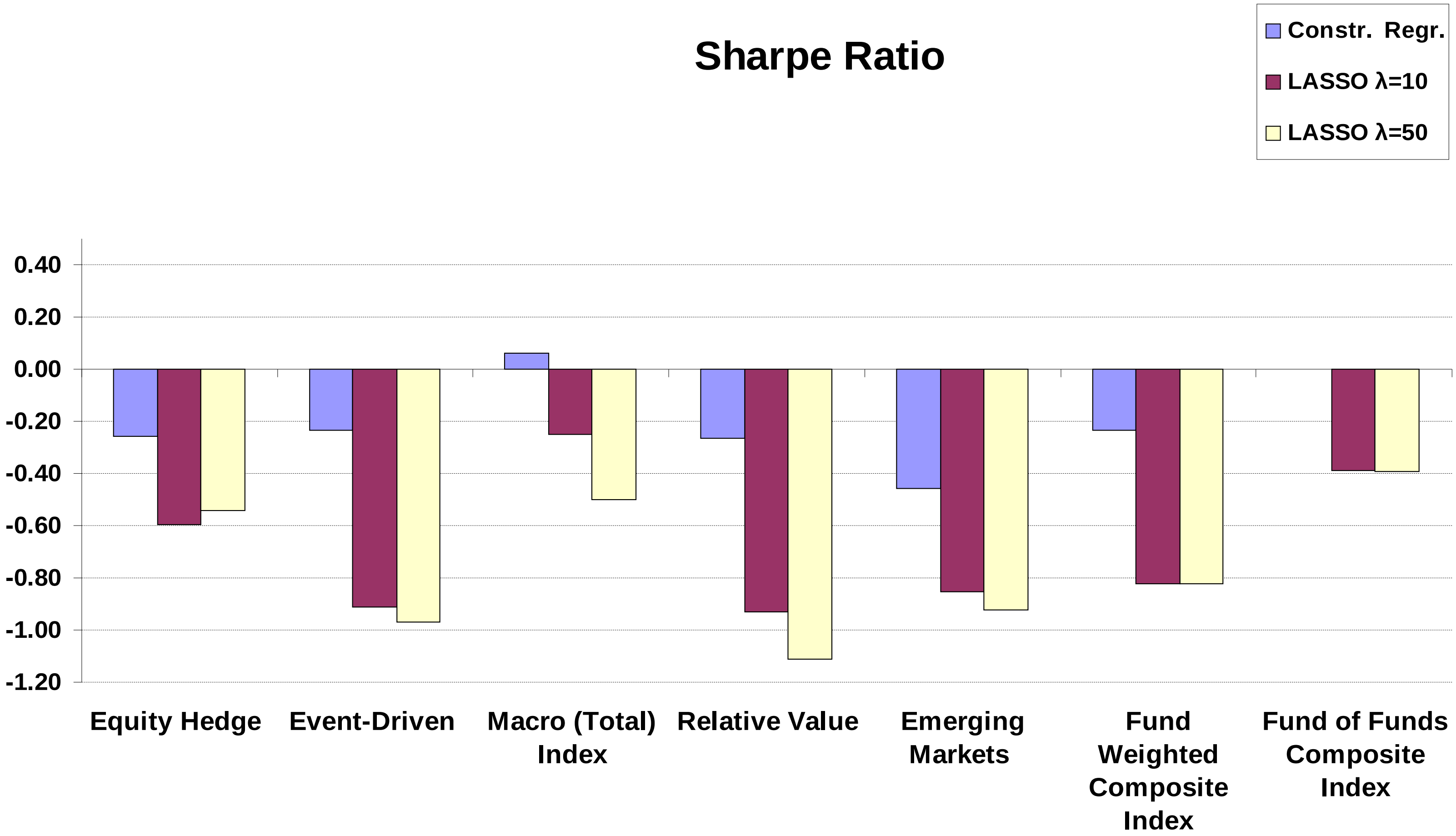
Out-of-Sample Excess Return



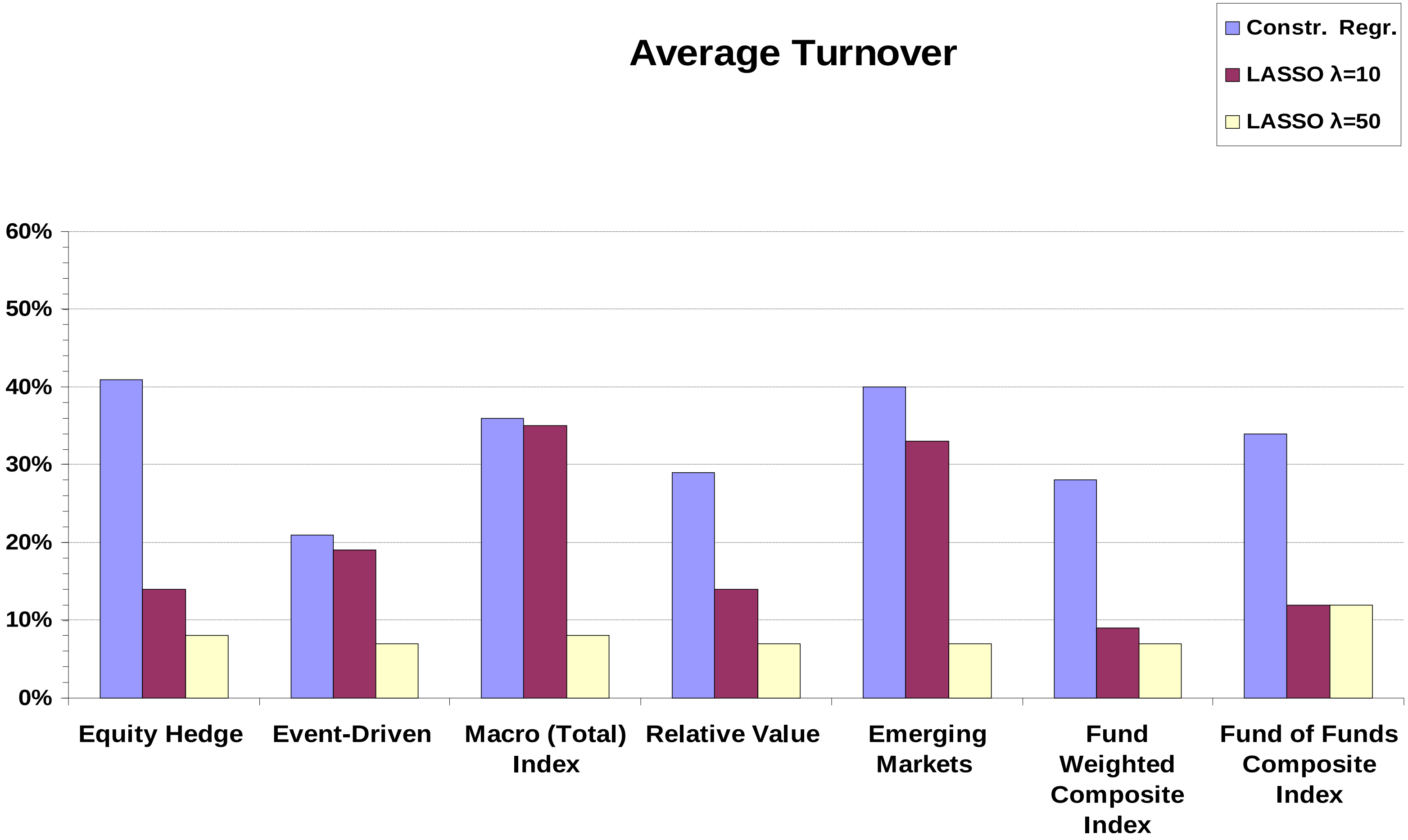
Out-of-Sample Tracking Error Volatility



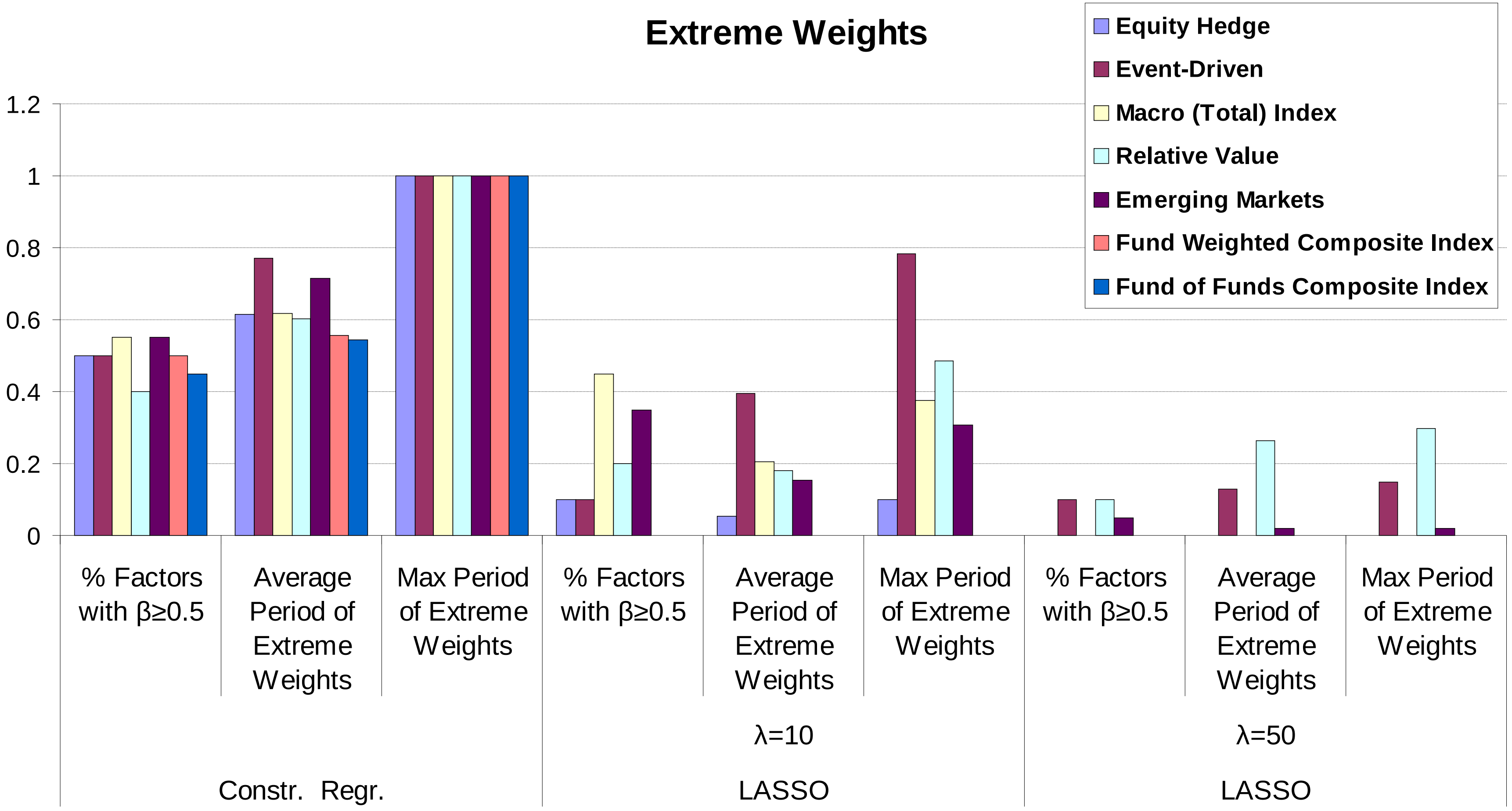
Out-of-Sample Sharpe Ratio



Average Turnover



Extreme Weights



A comparison with other products...

	Fund Weighted Composite Index	Constr. Regr.	LASSO $\lambda=10$	LASSO $\lambda=50$	Merril Lynch	Goldman Sachs
Mean return (%)	0.86	0.86	0.72	0.68	0.61	1.03
Std (%)	1.42	1.54	1.51	1.51	1.44	1.67
Mean return/ Std	0.61	0.56	0.47	0.45	0.43	0.62
VaR95 (%)	1.72	1.79	1.88	2.03	1.57	2.30
VaR99 (%)	2.62	3.15	3.17	3.04	5.20	3.94
CVaR95 (%)	2.38	2.58	2.73	2.65	3.51	3.11
CVaR99 (%)	2.69	3.30	3.27	3.15	5.64	4.17

A comparison with other products...

	Fund of Funds Composite Index	Constr. Regr.	LASSO		Citigroup	Deutsche Bank	Goldman Sachs
			λ=10	λ=50			
Mean return (%)	0.64	0.71	0.60	0.57	0.83	0.76	1.03
Std (%)	1.29	1.16	1.11	1.09	1.77	2.08	1.67
Mean return/ Std	0.50	0.61	0.54	0.52	0.47	0.36	0.62
VaR95 (%)	1.97	1.40	1.47	1.38	2.00	2.57	2.30
VaR99 (%)	2.87	2.63	2.43	2.25	3.38	8.00	3.94
CVaR95 (%)	2.59	2.13	2.11	2.00	2.86	5.72	3.11
CVaR99 (%)	2.90	2.75	2.49	2.30	3.43	8.33	4.17

Conclusion

Summing up

- Simple HF Clone methodology
- Constraining portfolio norms can help in:
  - selecting a subset of factors
  - avoiding extreme weights
  - controlling turnover and transaction cost
  - decreasing extreme risk of the portfolio
  - …

Conclusion

Further research

- Include transaction costs
- Choice of lambda
- *q*-norm penalty with q<1 (nonconcave penalized likelihood)
- …

References

Finance

- Agarwal, V. and N.Y. Naik, 2000, Generalized style analysis of Hedge Funds, Journal of Asset Management, 1, 93-109.
- Fung, W. and D.A. Hsieh, 2004, Extracting portable alphas from Equity Long/Short Hedge Funds, Journal of Asset Management, 2, 4,1-19.
- Hasanhodzic, J. and A. W. Lo, 2007, Can hedge-fund returns be replicated?: The linear case." Journal of Investment Management, 5, 2, 5-45.
- Jagannathan, R., and T. Ma, 2003, Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps," Journal of Finance, 58, 1651-1684.

Stats/Math

- Hoerl, A. E., and R. Kennard, 1970, Ridge Regression: Biased Estimation for Nonorthogonal Problems, Technometrics, 8, 27-51 .
- Frank, I. and J. Friedman, A statistical view of some chemometrics regression tools (with discussion), Tecnometrics, 35, 109-148.
- Schmidt, M., G. Fung, and R. Rosales, 2007, Fast Optimization Methods for L1-Regularization: A Comparative Study and 2 New Approaches, ECML'07.
- Tibshirani, R. 1996, Regression shrinkage and selection via the lasso, J. Royal. Statist. Soc B., 58, 267-288.
- Tikhonov, A.N., 1963, Solution of incorrectly formulated problems and the regularization method, Soviet Math Dokl 4, 1035-1038 English translation of Dokl Akad Nauk SSSR 151, 1963, 501-504.

Appendix

Choice of λ

LASSO

How do we choose λ? Cross-Validation

For lambda={0.5, 1,2,…}

- Given T= 120 sample returns,

  For t = 1 : T

  - Remove t^{th} return from sample
  - Estimate **β_hat** from other returns
  - Compute **y_hat=X* β_hat**
  - Compute "out-of-sample" return using t^{th} sample return
  - Compute "out of sample" Mean Squared Error (MSE)

  End

- Compute "out of sample" Mean Squared Error (MSE)

End

Choose lambda to minimize "out-of-sample" MSE

Literature Review

Hedge Fund Clones

In practice…

- Four styles of passive hedge fund strategies are emerging
  - **Factor Approach** (Hasanhodzic and Lo, 2007): combinations of liquid assets designed to replicate broad hedge fund benchmarks
  - **Mechanical Trade Replication**: systematically replicating popular hedge-fund strategies such as merger, convertible bond, or volatility arbitrage
  - **Moment-matching** (Kat and Palaro, 2005): match the statistical properties of a fund of index
  - **Security based replication** (Kat and Palaro, 2005): implementation of a generic version of a given strategy

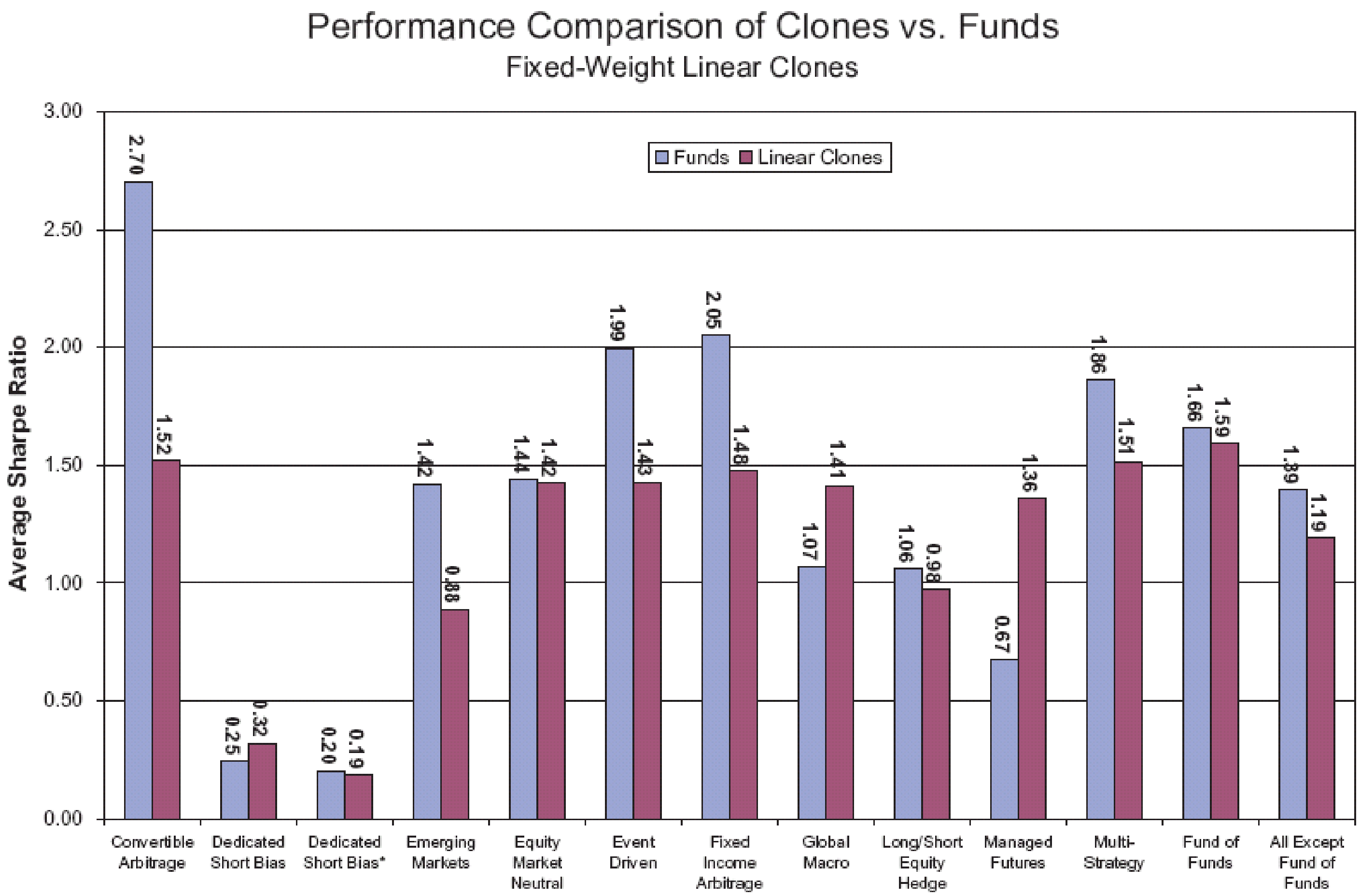
Factor Approach

An interesting investigation… Hasanhodzic and Lo (2007)

- TASS Database: 1610 hedge funds
- Period: 02/1996- 09/2005
- 6 factors
  - U.S. Dollar Index return (currency)
  - Lehman Corporate AA Intermediate Bond Index return (credit)
  - spread between the Lehman BAA Corporate Bond Index and the Lehman Treasury Index
  - S&P 500 total return Goldman Sachs Commodity Index
  - first-difference of the end-of-month value of VIX
- Linear regression with budget constraint
- Fixed Window and Rolling Window Estimates

Factor Approach

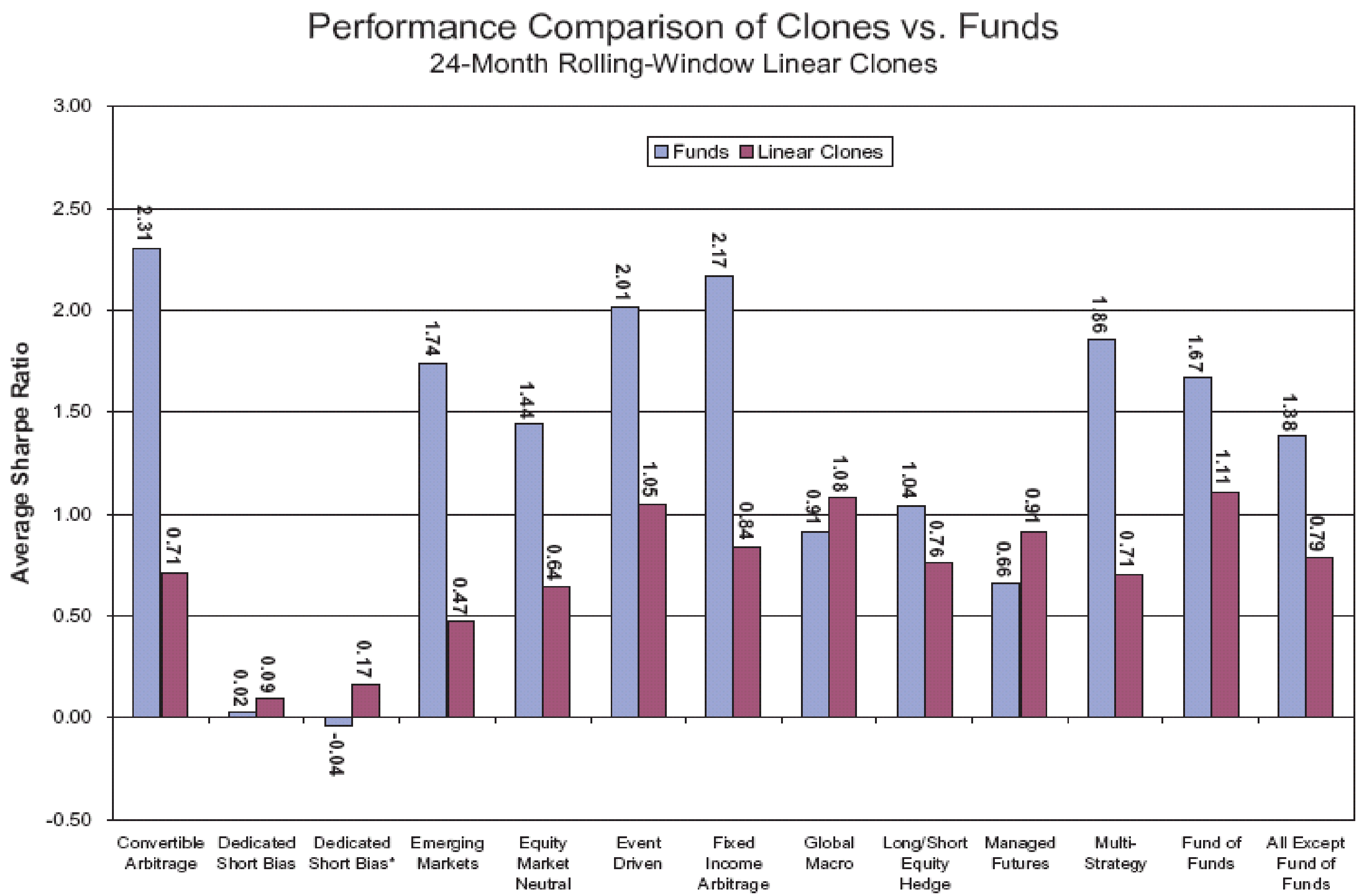
An interesting investigation…



Hasanhodzic and Lo (2007)

Factor Approach

An interesting investigation…



Hasanhodzic and Lo (2007)

Moment matching

Moment matching

- Representative study: Kat and Palaro (2005)
- ‘Mach the risk profile and you will match the return’
  - Achieve desired volatility, skewness, kurtosis, correlation to a specified portfolio
  - Does not target the mean return of the original fund
- Trading strategy involves holdings on:
  - Reserve portfolio: the return driver, always a long position, i.e. S&P500, GSCI, Russell2000, 4 X Intermediate Bonds, 5 X EurodollarFutures
  - Investor’s portfolio: reference for the correlation target, potential short position required, i.e. 50/50 mix of S&P500 and Long Bonds

Moment matching

Moment matching

- Almost 80% HFs and FoFs generated after-fee returns that do not exceed the returns from the replicating strategies
- Critics (answered in Kat and Palaro, 2007) include:
  - High transaction costs
  - The time series of returns is not matched
  - Average returns depend on the reserve asset that is chosen by the user
  - Investors need to have long time horizon to realize the benefit
  - The risk profile is targeted, but not the expected return
  - The approach is similar to mean-variance optimization although much more complicated
  - The dependence with other assets than the reference portfolio is not targeted
  - Black-Scholes framework is not an appropriate set up

Security based replication

Security based replication

- Implementation of a generic version of a given strategy
- Trend-following CTAs: long/short positions in commodity and financial futures using moving average signals.
- Merger Arbitrage: long positions in target firms and short positions in acquiring firms (Mitchell and Pulvino, 2001)
- Convertible Arbitrage: long positions in convertible bonds and short positions in equities (Agarwal et al., 2006)
- Fixed income arbitrage (Duarte et al. 2006)
- Not well studied, rarely cited

LARS

Least Angle Regression

LARS (Efron et al. 2004)

- Variable Selection in Regression
  - Many approaches: stagewise, LASSO, regularization, boosting…
- Least Angle Regression unifies Lasso and Forward Stagewise, provides a fast implementation, and a way to choose the number of terms to include

Least Angle Regression

- **Stepwise regression**:
  - Pick predictor most correlated **with y**
  - Bring predictor completely into model (full LS fit)
- **Forward stagewise:**
  - Pick predictor most correlated with **y**
  - Increment coefficient for predictor
- **Least Angle Regression:**
  - Pick predictor most correlated with y
  - Bring predictor into model only to extent it is better than others
  - Move in least-squares direction until another variable is as correlated

Least Angle Regression

LARS (Efron et al. 2004)