Statistical Arbitrage in US Equities Markets

April 25, 2018

Risk and Portfolio Management with Econometrics

Statistical Arbitrage

Trading strategy consisting of investing in stocks, both long and short, with a mean-reversion theme

-- Pairs trading: practiced since the 1990s, invented by Morgan Stanley's prop team, the ``grand-daddy'' of stat-arb

-- Factor neutral trading: generalizes pairs trading

-- Relative-value with sector ETFs trading: the ultimate subject of this presentation

Idea: measure statistically a time-series of spread between two or more correlated stocks or indices and invest in overbought/undersold spreads.

Trading Universe:
Stocks of more than 1BB cap

Sector	ETF	Num of Stocks	Market Cap		unit: 1M/usd
			Average	Max	Min
Internet	**HHH**	22	10,350	104,500	1,047
Real Estate	**IYR**	87	4,789	47,030	1,059
Transportation	**IYT**	46	4,575	49,910	1,089
Oil Exploration	**OIH**	42	7,059	71,660	1,010
Regional Banks	**RKH**	69	23,080	271,500	1,037
Retail	**RTH**	60	13,290	198,200	1,022
Semiconductors	**SMH**	55	7,303	117,300	1,033
Utilities	**UTH**	75	7,320	41,890	1,049
Energy	**XLE**	75	17,800	432,200	1,035
Financial	**XLF**	210	9,960	187,600	1,000
Industrial	**XLI**	141	10,770	391,400	1,034
Technology	**XLK**	158	12,750	293,500	1,008
Consumer Staples	**XLP**	61	17,730	204,500	1,016
Healthcare	**XLV**	109	14,390	192,500	1,025
Consumer discretionary	**XLY**	207	8,204	104,500	1,007
Total		**1417**	**11,291**	**432,200**	**1,000**

January, 2007

Investing: a quant perspective

Separate the **systematic** components of stock returns from the company-specific, or **idiosyncratic** components

$$R_i = \beta_i R_{Mkt} + \varepsilon_i$$

Project returns on single Market Factor (CAPM)

$$R_i = \sum_{j=1}^{m} \beta_{ij} F_j + \varepsilon_i$$

Project returns on Multiple (sector, size) Factors (APT)

In principle, market-neutral portfolios should have no exposure to market factors ("defactoring")

Defactoring: the Correlation Matrix Approach

$$R_{it} = \text{ daily stock returns in panel form}$$

$$i = 1, \ldots, N, \; t = 1, \ldots, T$$

$$\overline{R_i} = \frac{1}{T}\sum_{i=1}^{T} R_{it}, \qquad \overline{\sigma_i}^{2} = \frac{1}{T-1}\sum_i \left(R_{it} - \overline{R_i}\right)^2$$

$$\overline{\rho}_{ij} = \frac{1}{T-1}\sum_i \frac{\left(R_{it} - \overline{R_i}\right)\left(R_{jt} - \overline{R_j}\right)}{\overline{\sigma_i}\,\overline{\sigma_j}}$$

Principal Component Analysis

$$\lambda_1 > \lambda_2 \geq \lambda_3 \geq \ldots \geq \lambda_N$$

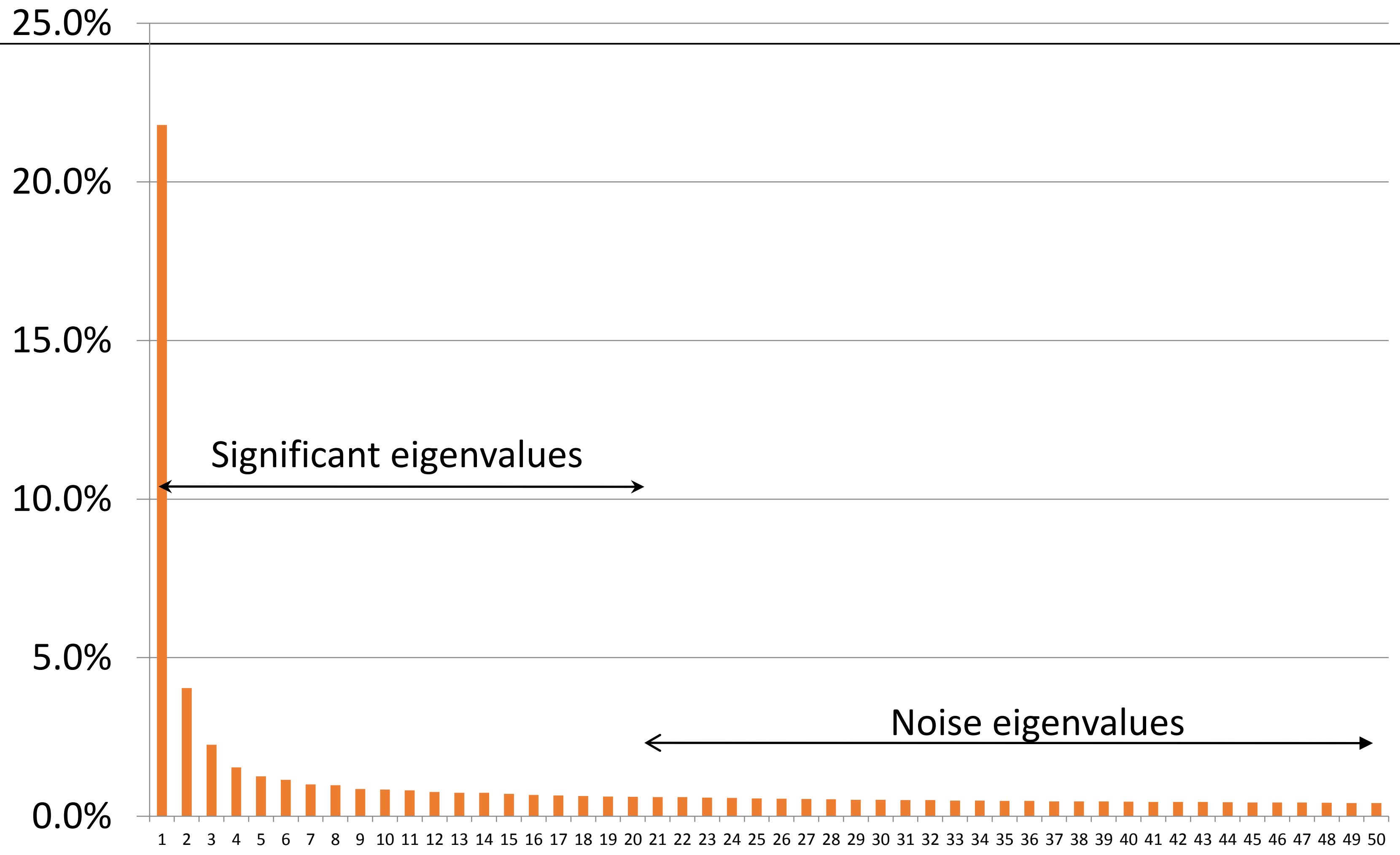
Eigenvalues are all non-negative

$$V^{(j)} = \left(v_1^{(j)}, v_2^{(j)}, \ldots, v_N^{(j)}\right)$$

Orthogonal eigenvectors

Stock market fluctuations can be characterized as moves along the eigenvector directions. We seek to extract mathematical factors from the PCA analysis.

PCA: Explained variance from the viewpoint of eigenvalues



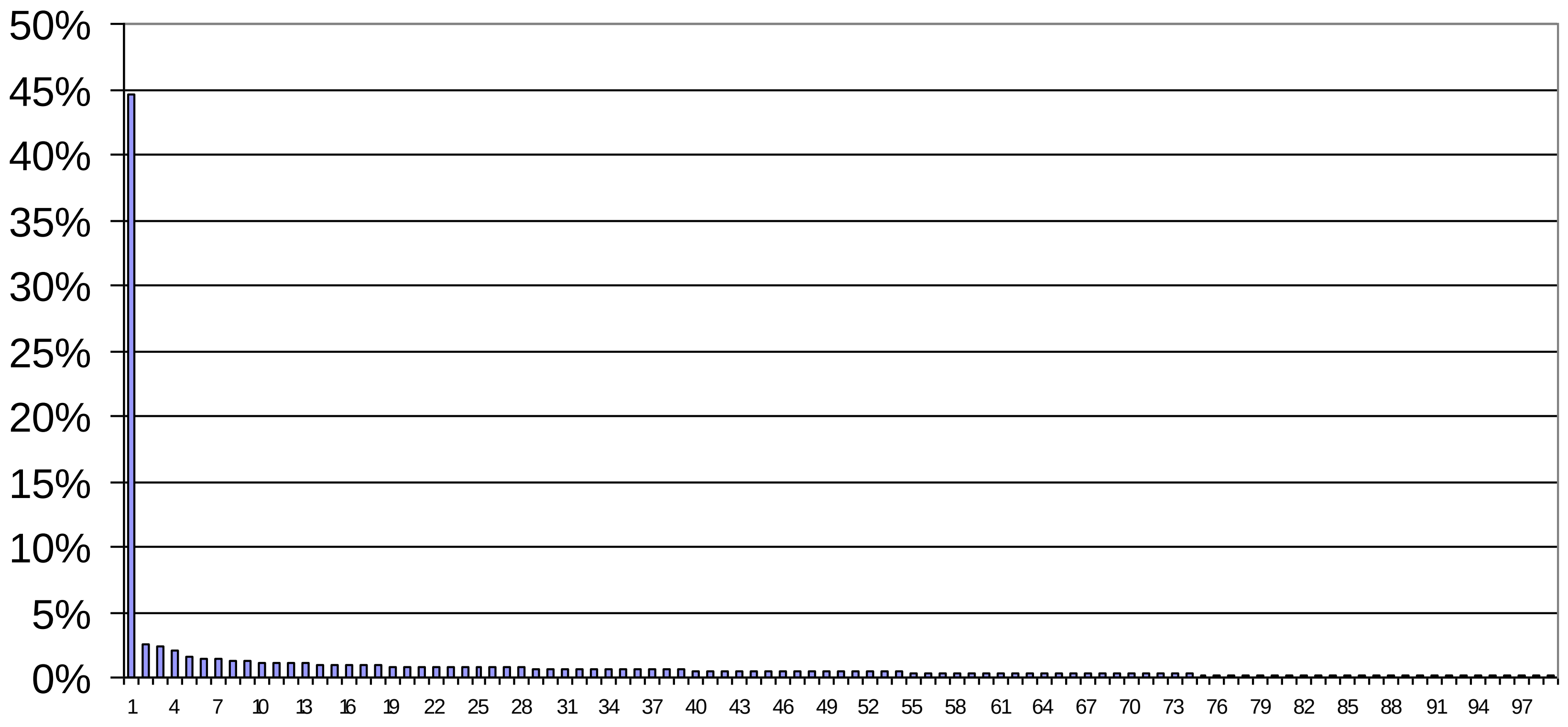
Big universe: Jan 2007-Dec 2007

Nasdaq-100
Components of NDX/QQQQ

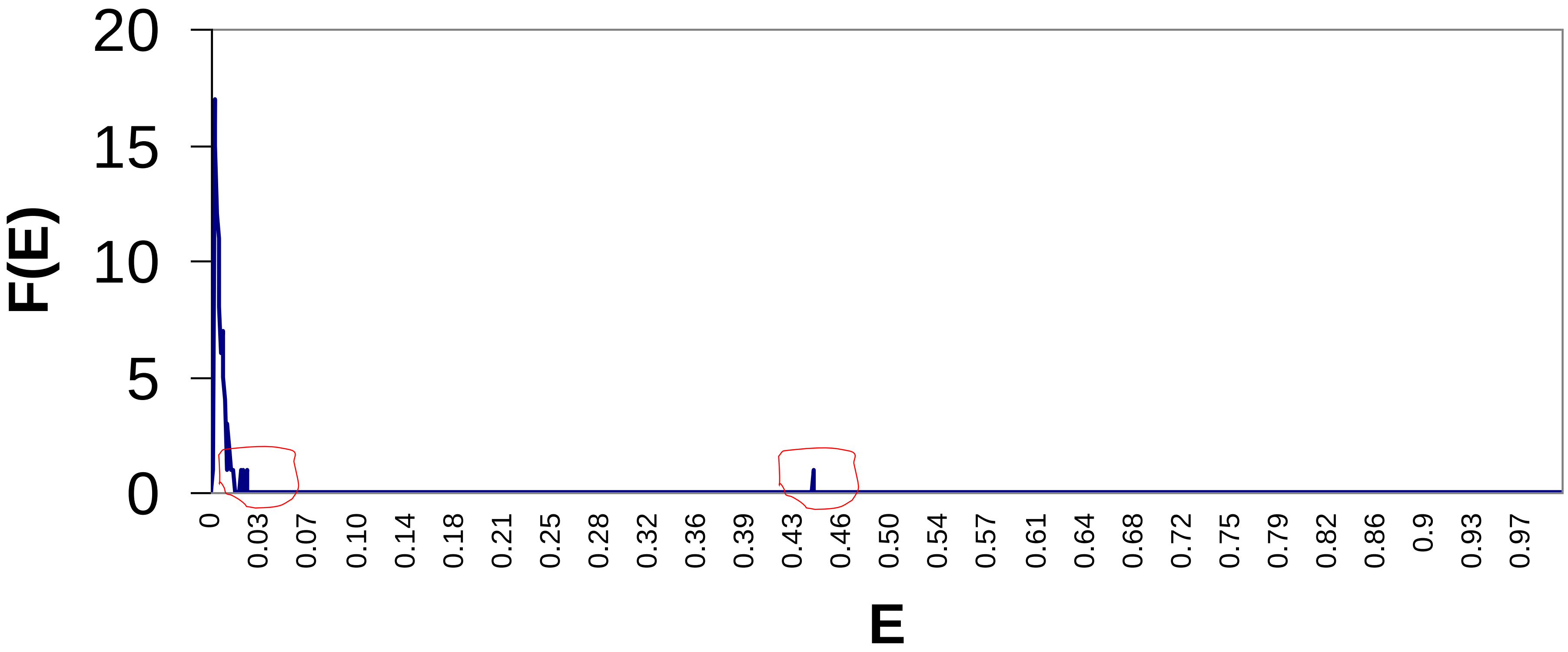
Data: Jan 30, 2007 to Jan 23, 2009

502 dates, 501 periods

99 Stocks (1 removed) MNST (Monster.com), now listed in NYSE

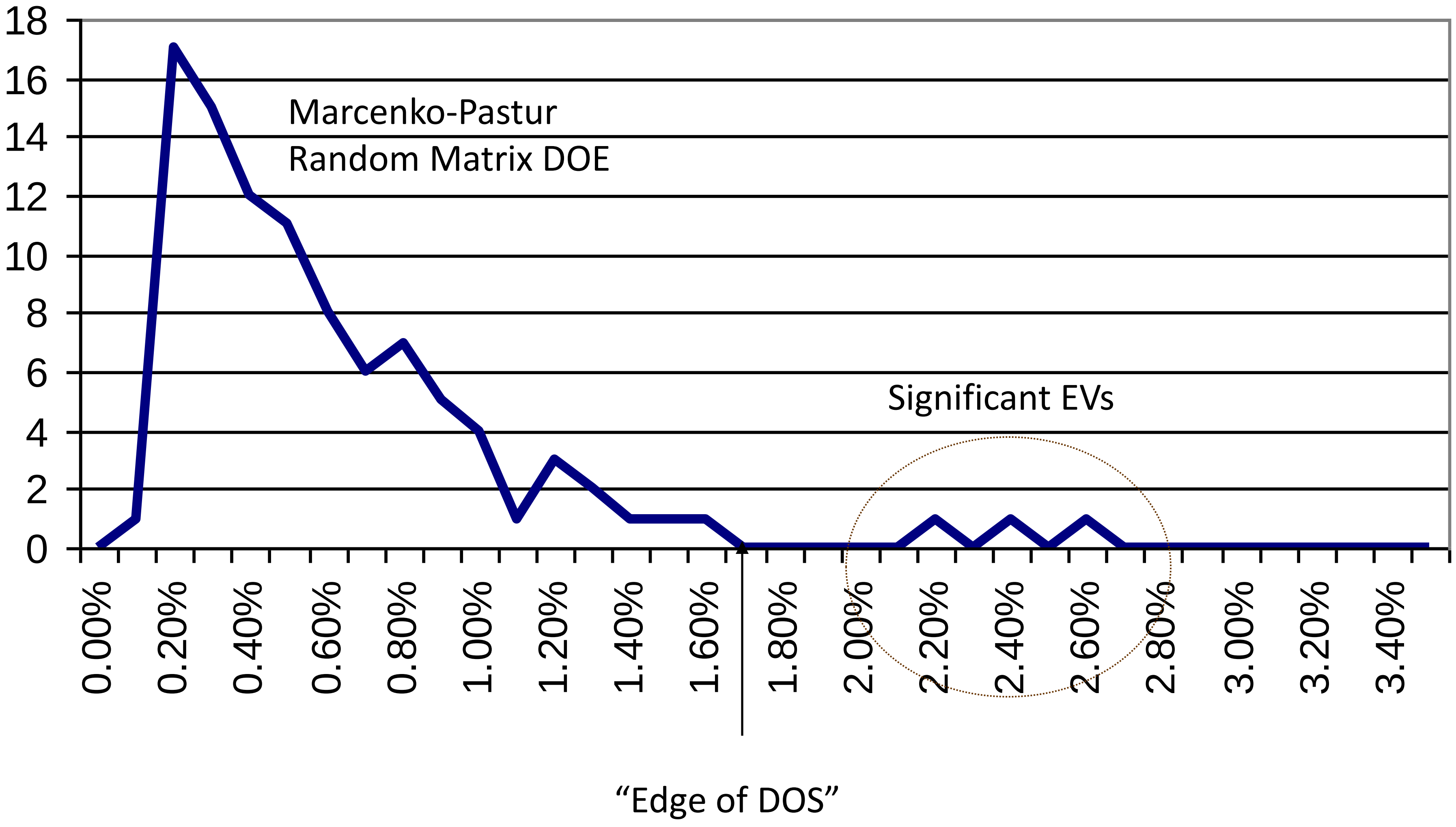


Density of States (from previous data)



One large mass at 0.44,
Some masses near 0.025
Nearly continuous density for lower levels

Zoom of the DOS for low eigenvalues



Factors & Eigenportfolios

For each eigenvector, build a portfolio which is weighted proportionally to the coefficient of each stock and inversely proportionally to its volatility

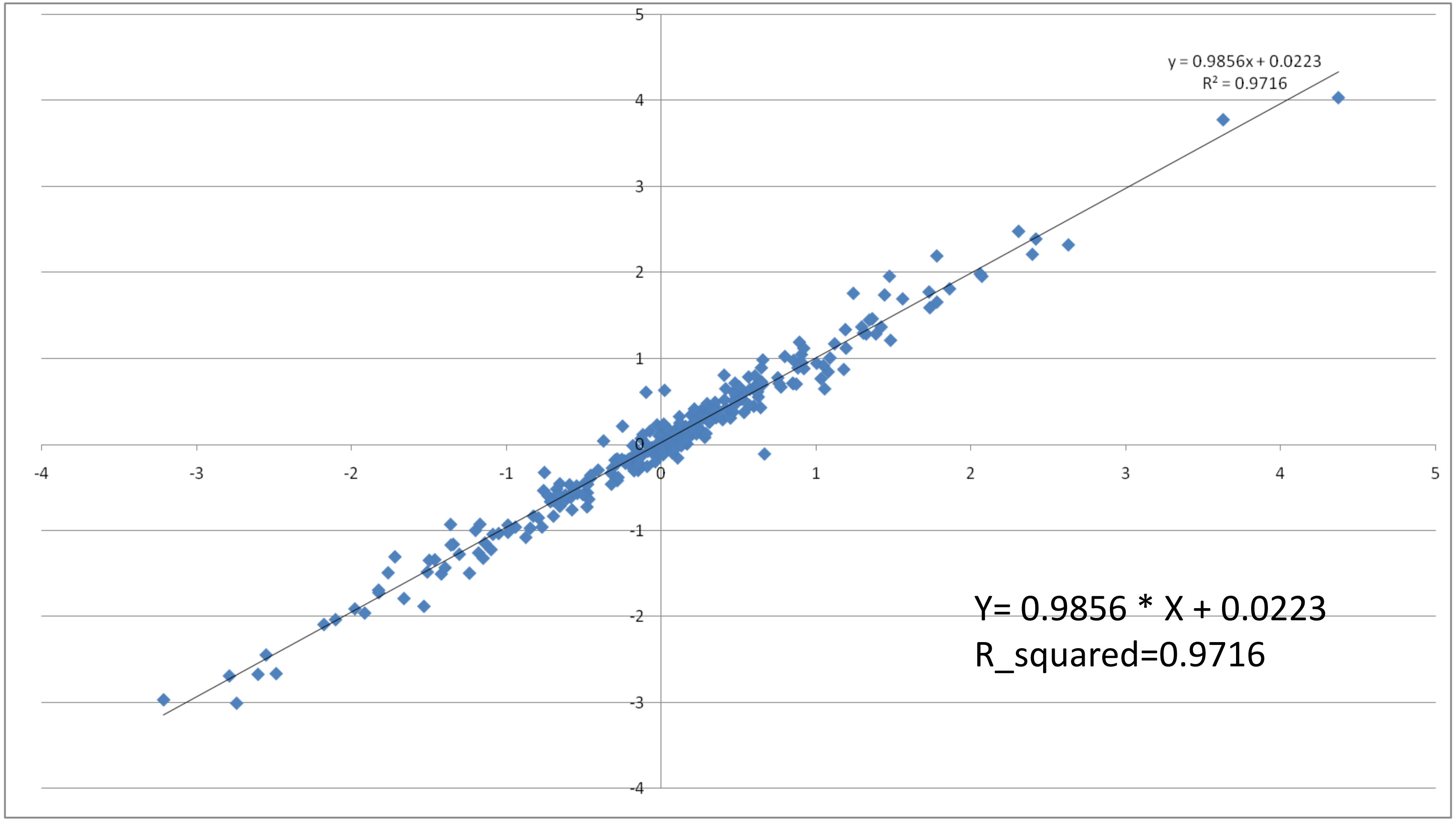
$$Q_i^{(j)} = \frac{v_i^{(j)}}{\overline{\sigma}_i}$$

Portfolio weight of j-th eigenportfolio

$$F_j = \sum_{i=1}^{N} Q_i^{(j)} R_i = \sum_{i=1}^{N} \left(\frac{v_i^{(j)}}{\overline{\sigma}_i} \right) R_i$$

J-th factor is the return of the j-th eigenportfolio

First eigenportfolio returns compared with S&P 500 returns (1/5/2009 to 1/29/2010)



How many eigenportfolios are significant?

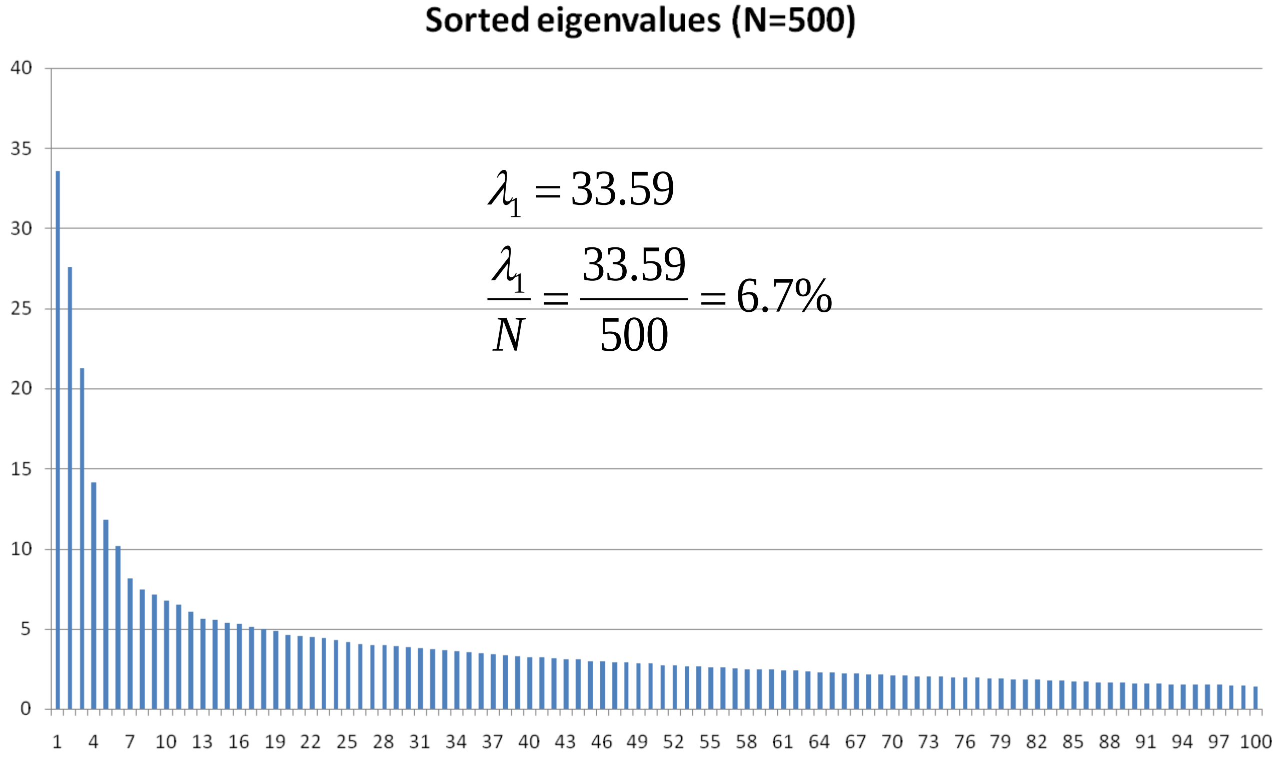
- Perform PCA on the empirical correlation matrix with 1 year moving window

- Consider the correlation matrix of the residuals after removing 1, 2 ,3… eigenportfolios

- Compare the DOS of the correlation of the residuals with the spectrum of the correlation matrix of pure noise (Marcenko-Pastur)

-The number of significant factors corresponds to the **first m for which the matrix of residuals is close to MP** (e.g. in the sense of hypothesis testing)

Eigenvalues of the correlation matrix of residuals (m=1)



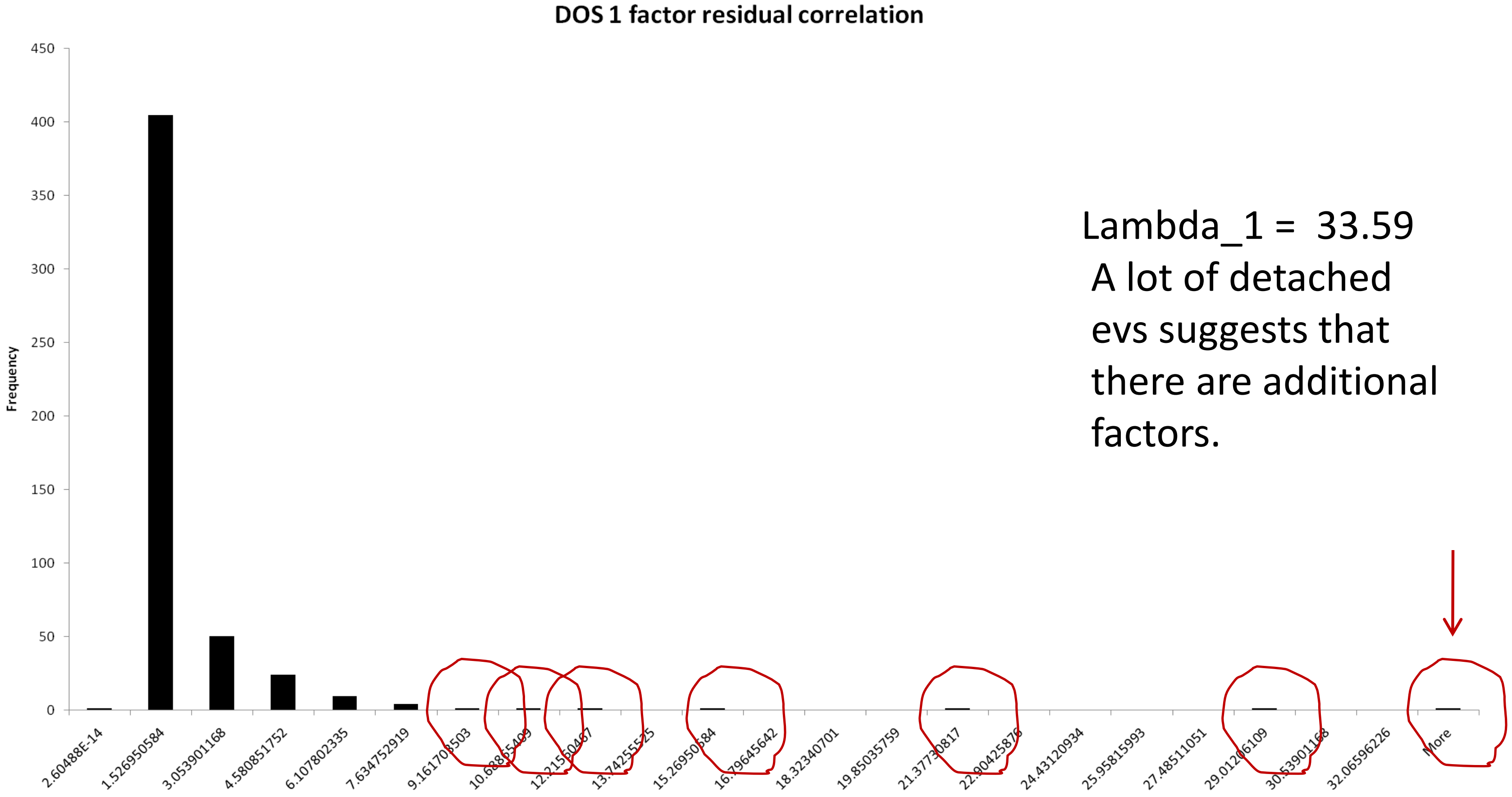
$$\lambda_1 = 33.59$$

$$\frac{\lambda_1}{N} = \frac{33.59}{500} = 6.7\%$$

Recall that lambda_1 for the original correlation matrix was ~ 220, so the residuals matrix has ``smaller'' correlations.

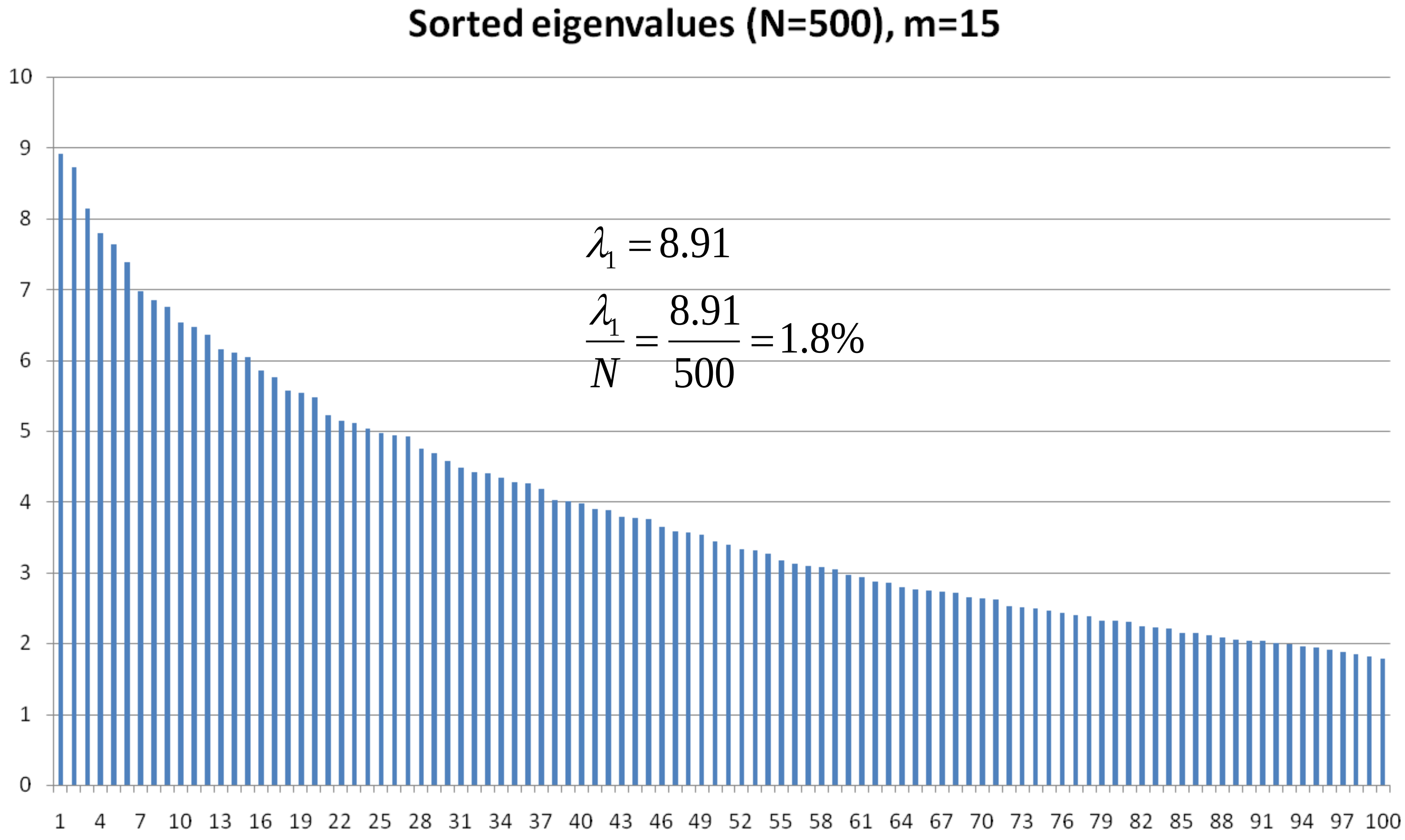
Ratio lambda/N is a proxy for the average correlation.

Density of States, or Histogram, of Eigenvalues for m=1 (correlation of returns with first factor removed)



Lambda_1 = 33.59
A lot of detached evs suggests that there are additional factors.

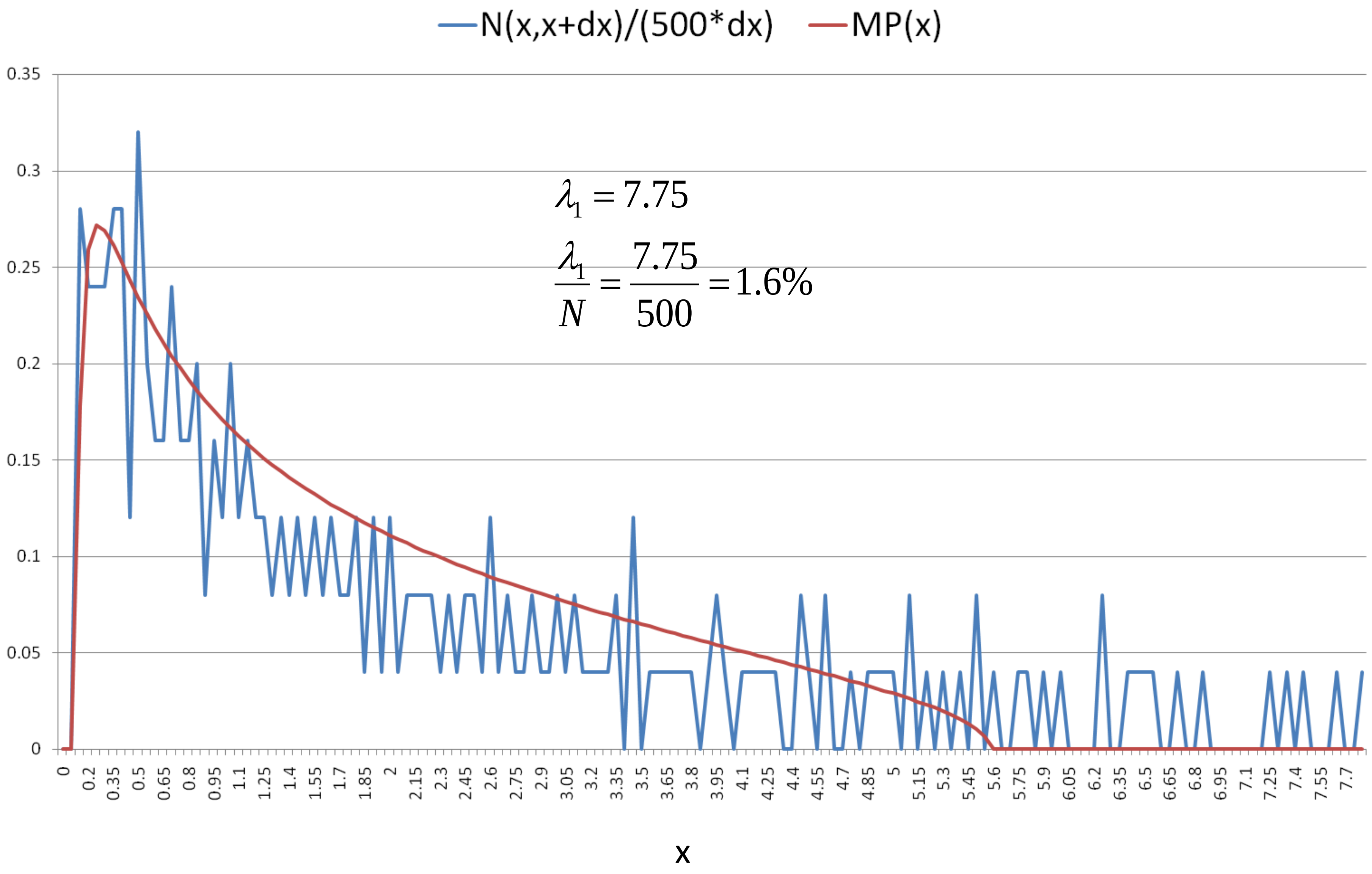
Top 100 eigenvalues of the correlation matrix of residuals (m=15)



$$\lambda_1 = 8.91$$

$$\frac{\lambda_1}{N} = \frac{8.91}{500} = 1.8\%$$

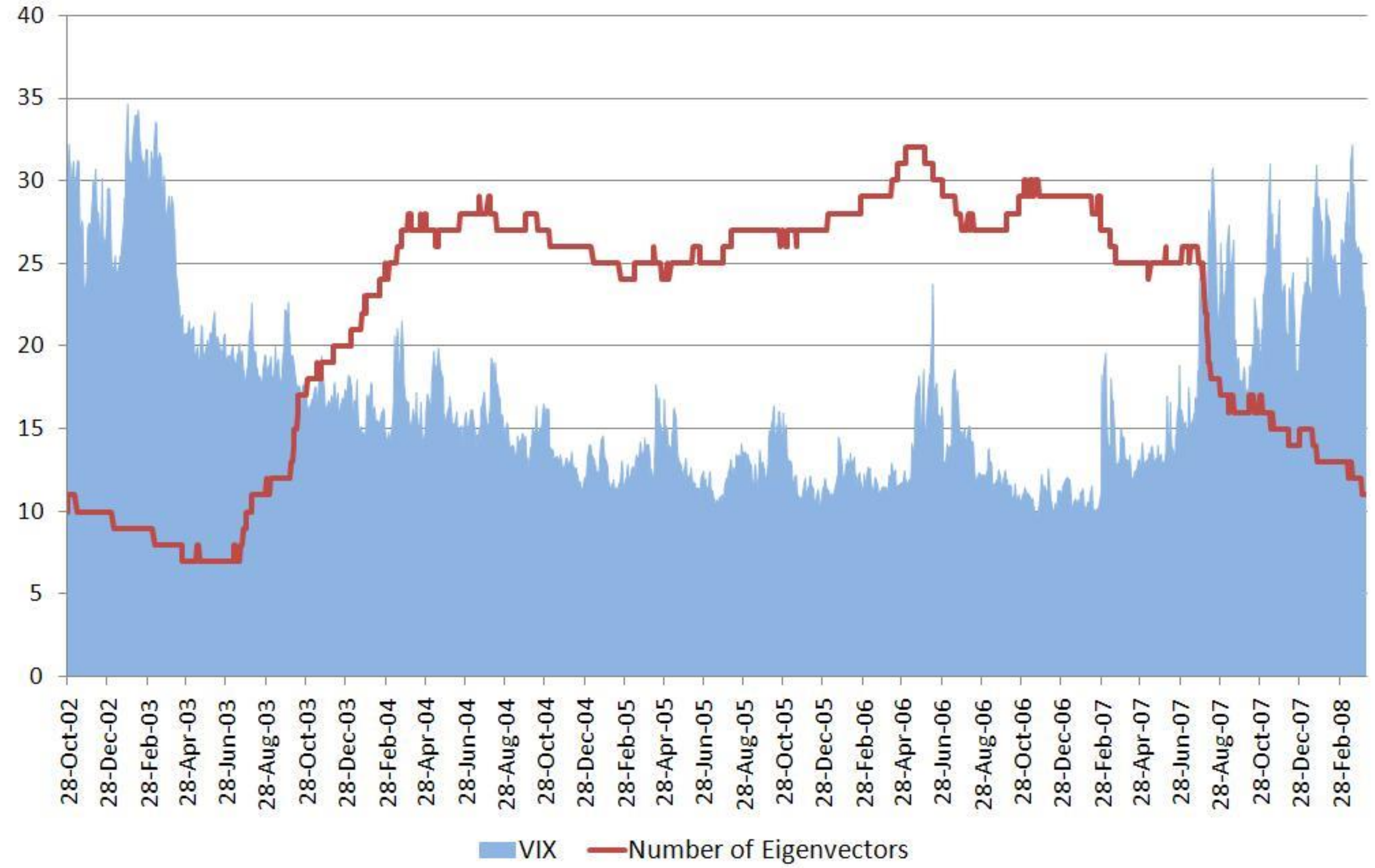
Marcenko Pastur compared to data with m=30



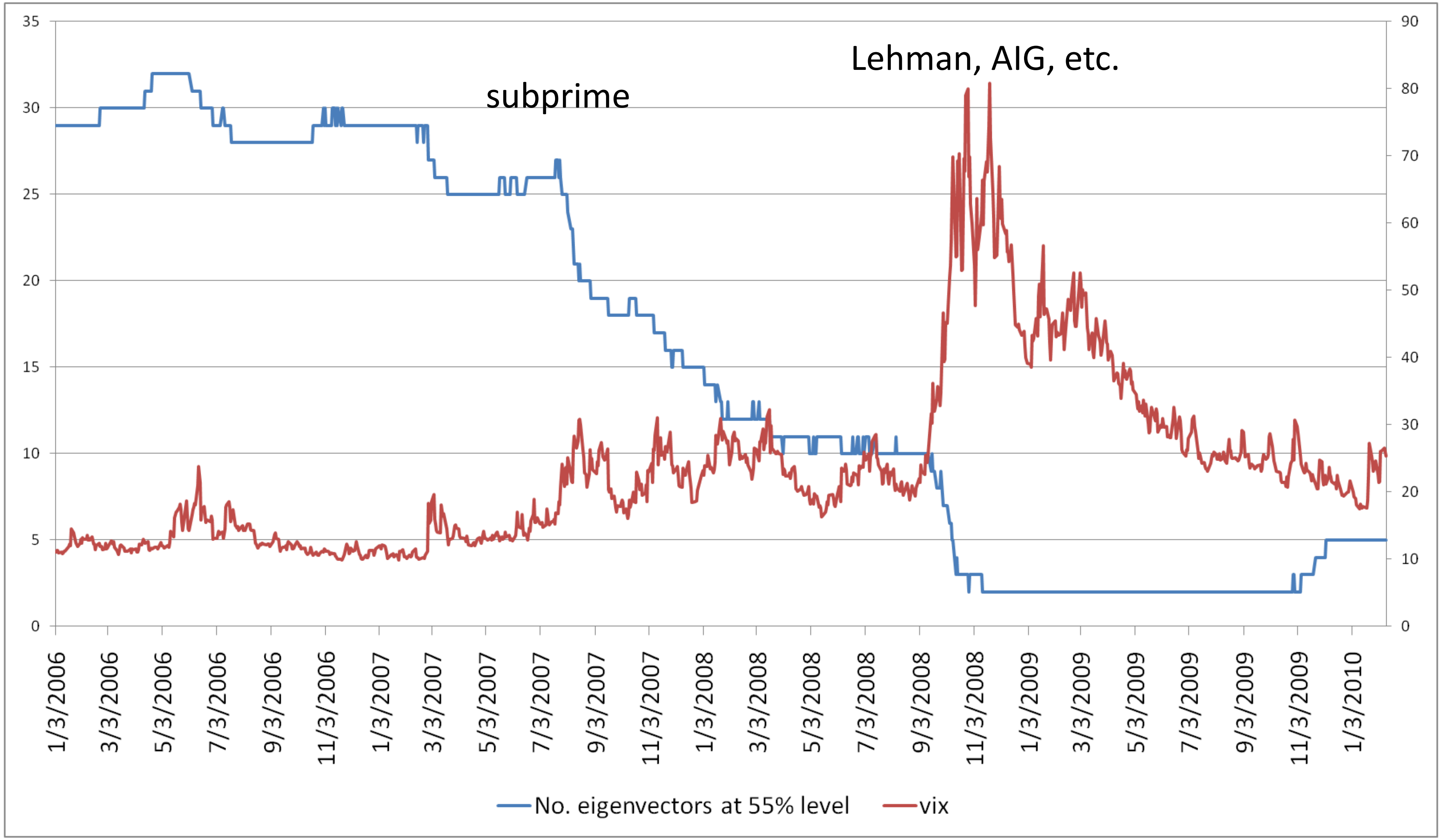
$$\lambda_1 = 7.75$$

$$\frac{\lambda_1}{N} = \frac{7.75}{500} = 1.6\%$$

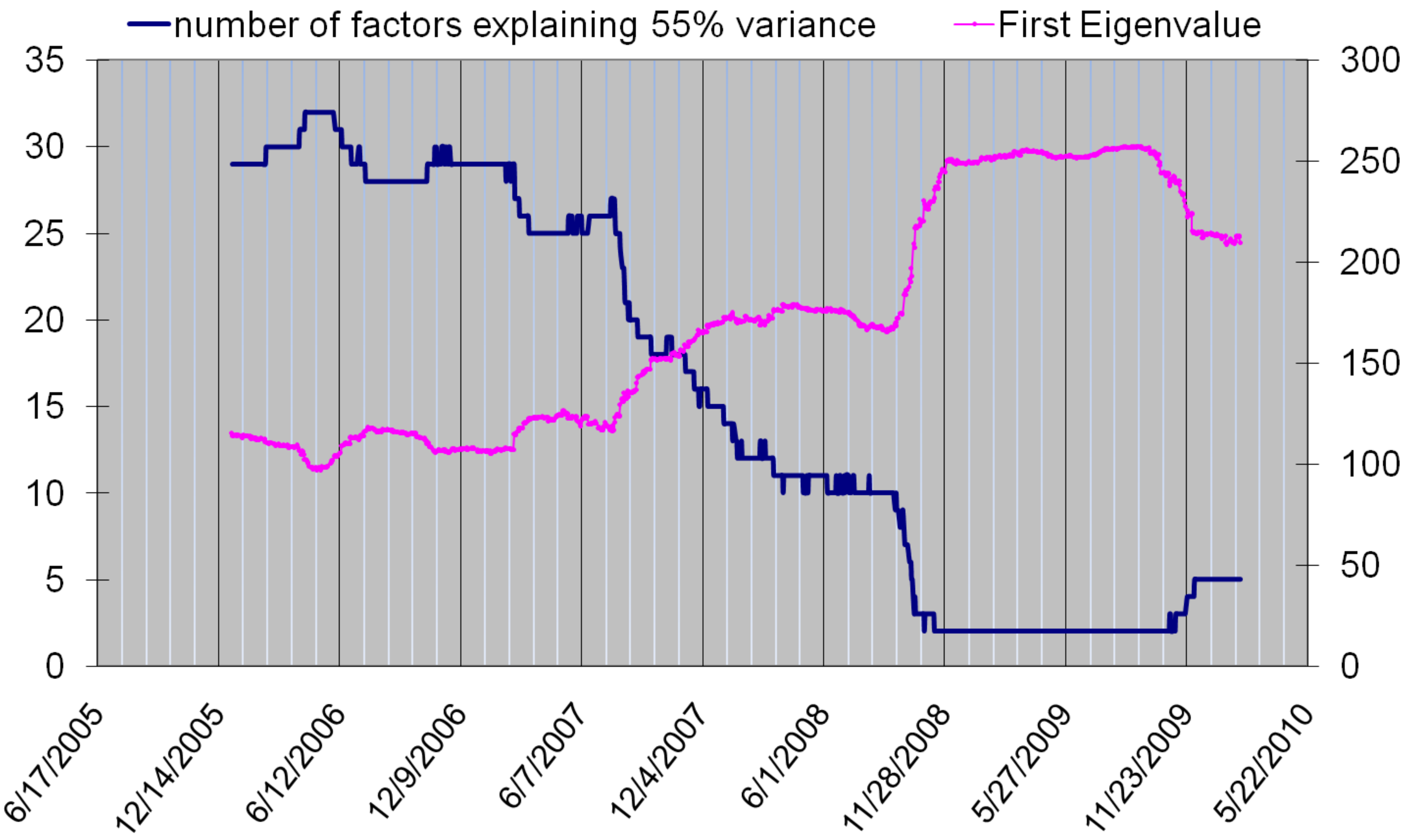
Number of factors explaining 55% variance vs. VIX (2002-2008)



Number of significant EVs versus VIX (1/2006-2/2010)



Number of significant factors vs. first eigenvalue of correlation matrix



Using ETFs as factors and relative value indicators (e.g. trade EBAY versus QQQQ)

Idea: rather than decomposing risk into abstract factors, consider a stock and the sector ETF corresponding to its industry

Pro: the relative analysis of a stock compared to its peers is done comparing with a tradable index. We deal with one stock-specific risk at a time, not 2 or more.

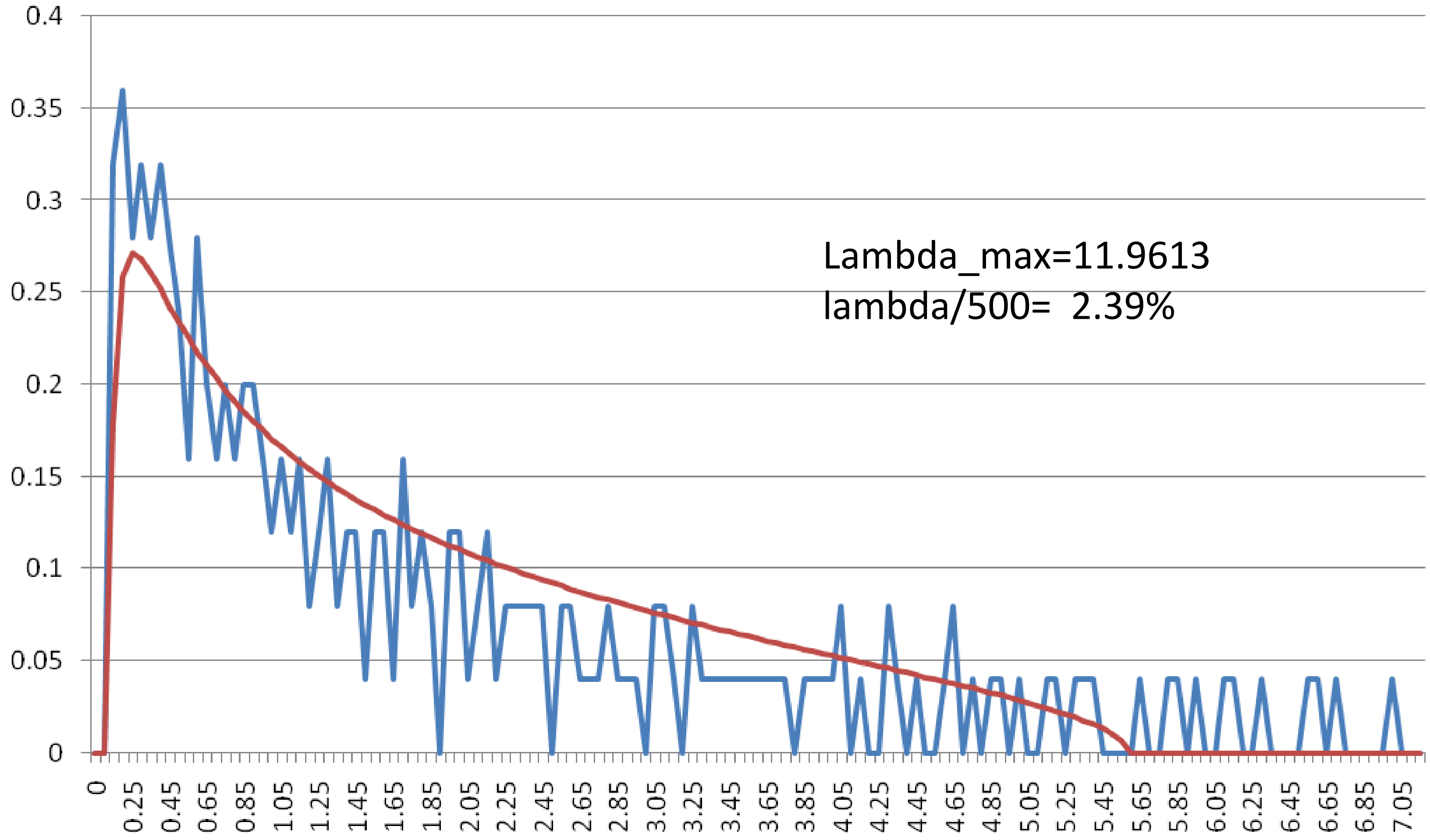
Pro: this is better than using an abstract factor from PCA because the latter are more likely to be unstable

Pro: the complexity of the signal generation is significantly reduced and the origin of excess returns is simple to explain

Con: some large cap stocks track too closely the industry, so no signals are available. Equivalently, the strategy can be viewed as somewhat biased to small caps

Con: Some stocks do not fit well the ETF identikit. We leave those to fundamental analysts.

Defactoring with ETFs as factors (19 ETFs)
(Red line=Marcenko-Pastur, universe =S&P 500)



Modeling the Evolution of Stock Residuals

$$\frac{dS_i(t)}{S_i(t)} = \beta_i \frac{dI(t)}{I(t)} + \varepsilon_i(t)$$

Stock returns a sum of sector-explained return and a residual process

$$\varepsilon_i(t) = \alpha_i dt + dX_i(t)$$

Residual= drift component (expected excess return above mkt.) + increment of a stationary process

$$dX_i(t) = \kappa_i (m_i - X_i(t))dt + \sigma_i dW_i(t)$$

Ornstein-Ulembeck
AR-1 process

Statistical Estimation Window=3 months (~ 60 business days)

Estimation of Ornstein-Uhlenbeck models

$$X_{t+\Delta t} = e^{-k\Delta t} X_t + m\left(1 - e^{-k\Delta t}\right) + \sigma \int_t^{t+\Delta t} e^{-k(t-s)} dW_s$$

$$X_{n+1} = a + bX_n + \varepsilon_{n+1} \qquad \{\varepsilon_n\} \text{ i.i.d. } N\left(0, \sigma^2\left(\frac{1-e^{-2k\Delta t}}{2k}\right)\right)$$

$$b = \mathrm{SLOPE}\left(\left(X_{n-l},\ldots,X_n\right);\left(X_{n-l-1},\ldots,X_{n-1}\right)\right),$$
$$a = \mathrm{INTERCEPT}\left(\left(X_{n-l},\ldots,X_n\right);\left(X_{n-l-1},\ldots,X_{n-1}\right)\right)$$
$$k = \frac{1}{\Delta t}\ln\left(\frac{1}{b}\right), \quad m = \frac{a}{1-b}, \quad \sigma = \frac{\mathrm{STDEV}\left(X_{n+1} - bX_n - a\right)}{\sqrt{1-b^2}}\sqrt{2\frac{1}{\Delta t}\ln\left(\frac{1}{b}\right)}$$

Statistics on the Estimated OU Parameters

ETF	Abs(Alpha)	Beta	Kappa	Reversion days	EquiVol	Abs(m)
HHH	0.20%	0.69	38	7	4%	3.3%
IYR	0.11%	0.90	39	6	2%	1.8%
IYT	0.18%	0.97	41	6	4%	3.0%
RKH	0.10%	0.98	39	6	2%	1.7%
RTH	0.17%	1.02	39	6	3%	2.7%
SMH	0.19%	1.01	40	6	4%	3.2%
UTH	0.09%	0.81	42	6	2%	1.4%
XLF	0.11%	0.83	42	6	2%	1.8%
XLI	0.15%	1.15	42	6	3%	2.4%
XLK	0.17%	1.03	42	6	3%	2.7%
XLP	0.12%	1.01	42	6	2%	2.0%
XLV	0.14%	1.05	38	7	3%	2.5%
XLY	0.16%	1.03	39	6	3%	2.5%
Total	0.15%	0.96	40	6	3%	2.4%

Average over 2006-2007

Trading Signals

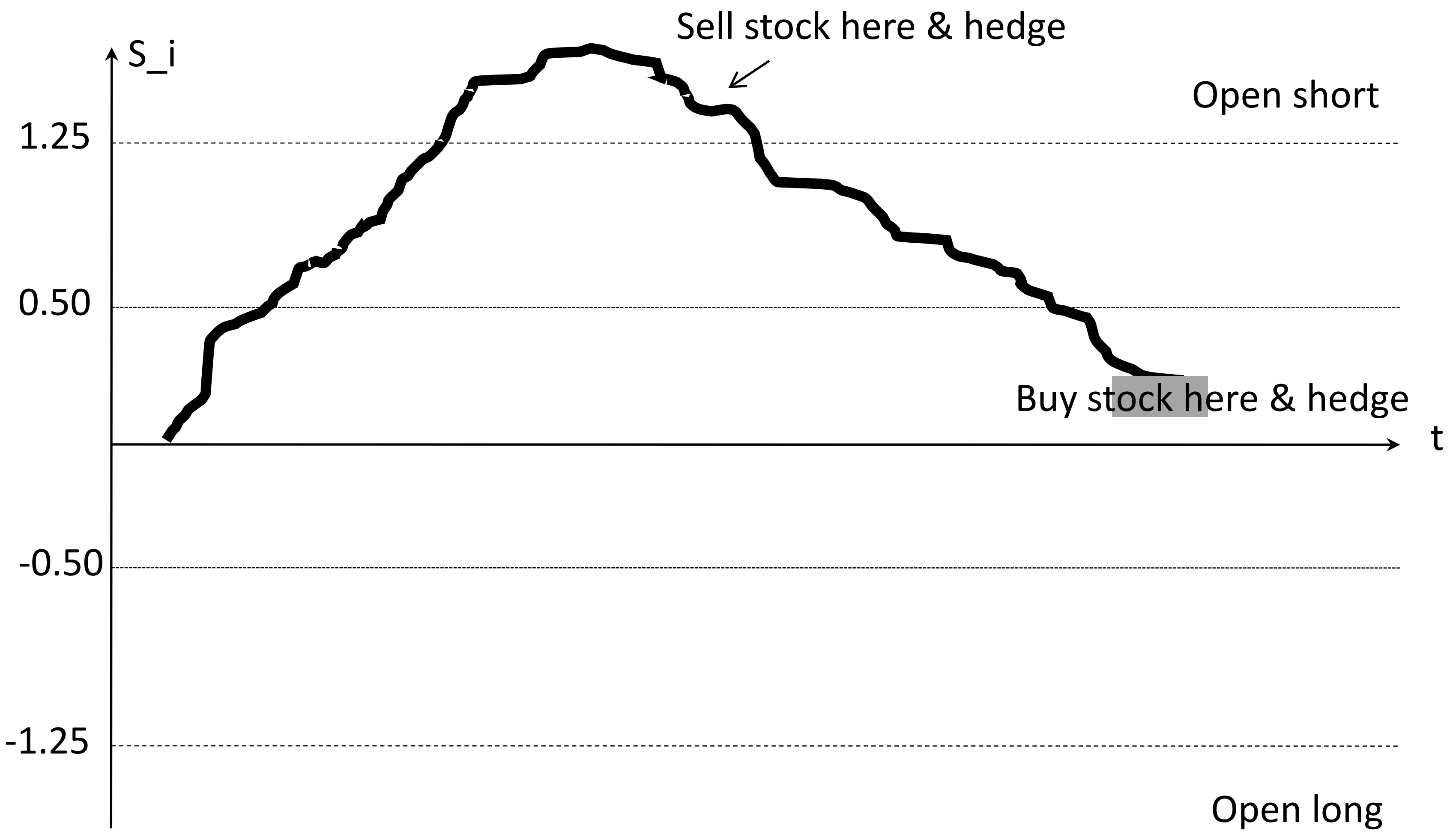
We introduce the **s-score** for each stock:

$$s_i(t) = \frac{X_i(t) - m_i}{\sigma_{eq,i}}$$

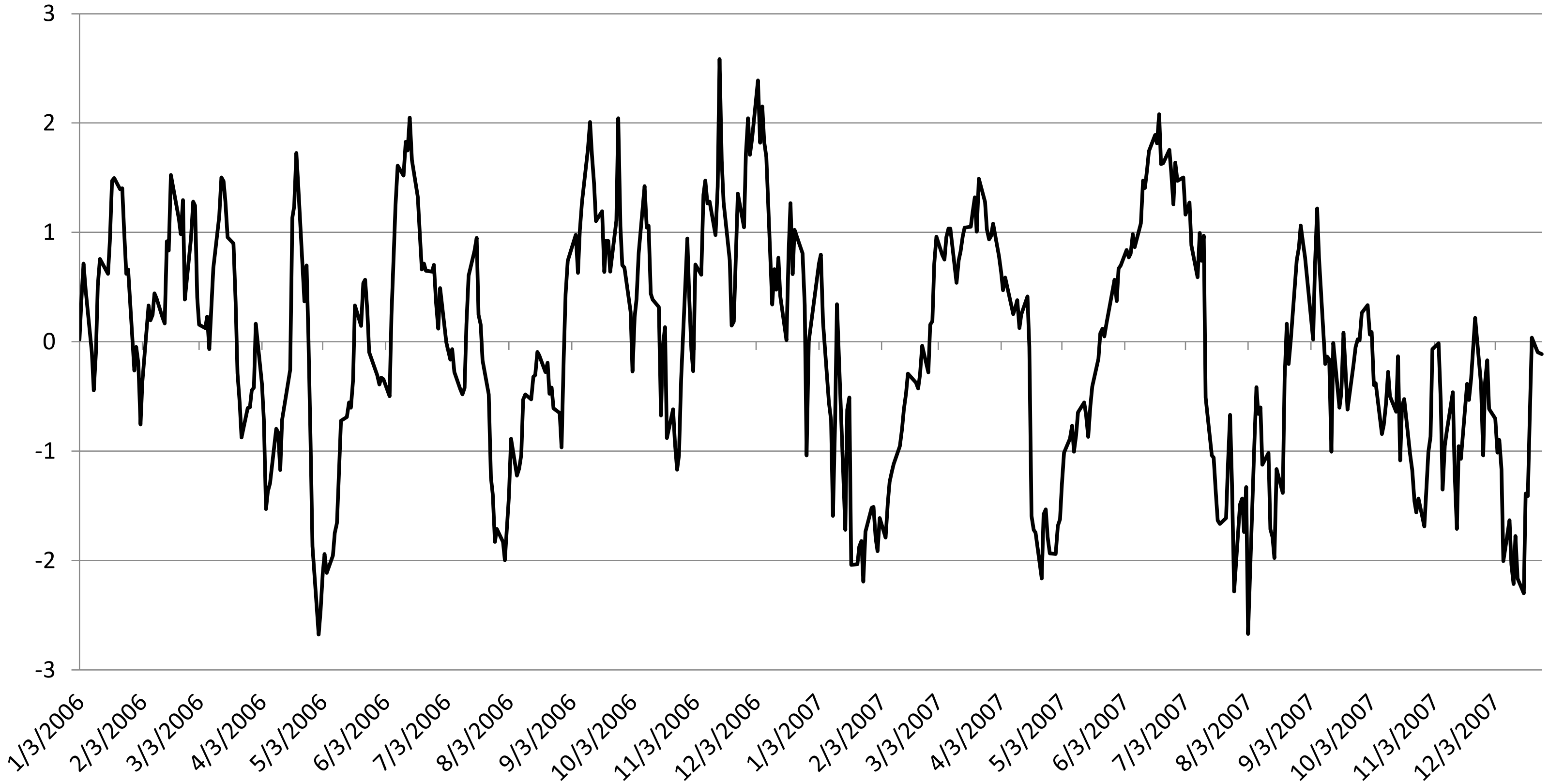
$$\begin{array}{ll} \text{Open long position if} & s_i < -1.25 \\ \text{Open short position if} & s_i > +1.25 \\ \text{Close long position if} & s_i > -0.50 \\ \text{Close short position if} & s_i < +0.50 \end{array}$$

Very naïve, but testable and reproducible.

Schematic view of mean-reversion trading



S-score of JPM (vs. XLF)



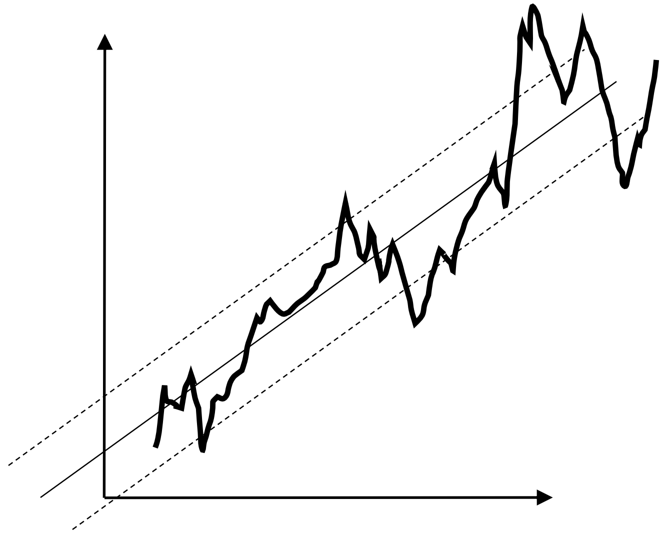
Including the drift in signal-generation

$$\mu_i = \alpha_i + \frac{E\{dX_i(t) \mid X_i(t)\}}{dt}$$

$$= \alpha_i + \kappa_i \left(m_i - X_i(t) \right)$$

$$= \alpha_i - \kappa_i \sigma_{eq,i} s_i$$

$$= -\kappa_i \sigma_{eq,i} \left(s_i - \frac{\alpha_i}{\kappa_i \sigma_{eq,i}} \right)$$



$$\therefore \qquad s_{\mathrm{mod},i} = s_i - \frac{\alpha_i}{\kappa_i \sigma_{eq,i}} = s_i - \frac{\alpha_i \tau_i}{\sigma_{eq,i}}$$

Stock-ETF trading example : ACGL/XLF trades in 2009
Arch Capital Group Ltd. Market cap: 4.3 BB

ticker	trade_date	close_date	days	stock_PNL	etf_PNL	Total_PNL	etf_ticker	size
ACGL	20090206	20090212	4	3,598.79	2,038.35	5,637.14	XLF	75,579.64
ACGL	20090226	20090302	2	(1,416.04)	4,036.11	2,620.08	XLF	78,906.91
ACGL	20090309	20090319	6	15,016.04	(11,763.40)	3,252.64	XLF	78,469.98
ACGL	20090420	20090428	6	4,643.71	(1,046.79)	3,596.92	XLF	79,350.04
ACGL	20090505	20090512	3	3,989.70	(1,289.76)	2,699.94	XLF	79,815.06
ACGL	20090519	20090602	9	2,257.45	24.34	2,281.79	XLF	79,201.01
ACGL	20090608	20090618	6	(1,784.06)	1,563.52	(220.54)	XLF	79,210.95
ACGL	20090806	20090812	2	955.02	0.00	955.02	XLF	81,552.52

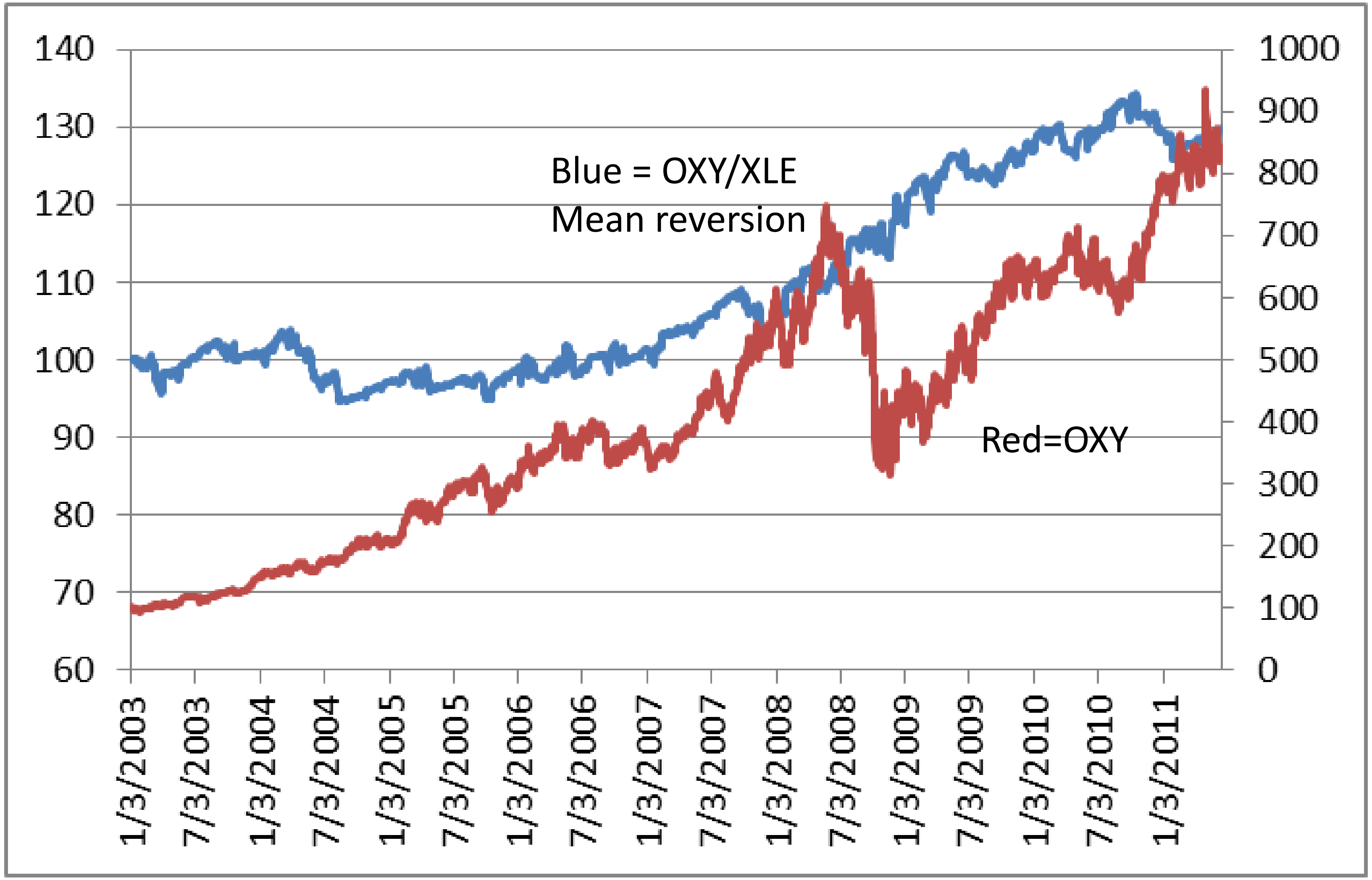
stock_cost	stock_shares	stock_exe_price	etf_cost	etf_shares	etf_exe_price	status	open_s_score	close_s_score
60.13	1,257.00	62.99	9.83	(2,811.52)	9.10	1.00	(1.83)	(0.08)
53.94	1,463.00	52.97	8.18	(3,603.67)	7.06	1.00	(1.30)	(0.49)
45.10	1,740.00	53.73	6.28	(4,801.39)	8.73	1.00	(1.77)	(0.33)
54.02	1,469.00	57.18	10.09	(2,754.71)	10.47	1.00	(1.29)	0.94
56.85	1,404.00	59.69	11.66	(2,528.93)	12.17	1.00	(1.96)	(0.37)
58.28	1,359.00	59.94	12.23	(2,433.83)	12.22	1.00	(1.68)	(0.12)
58.03	1,365.00	56.72	12.44	(2,605.86)	11.84	1.00	(1.69)	(0.87)
61.97	1,316.00	62.70	14.01	(2,293.69)	14.01	1.00	(1.25)	(0.01)

Stock-ETF trading example : EBAY/QQQQ trades in 2009
eBay Inc. Market cap= USD 38 BB

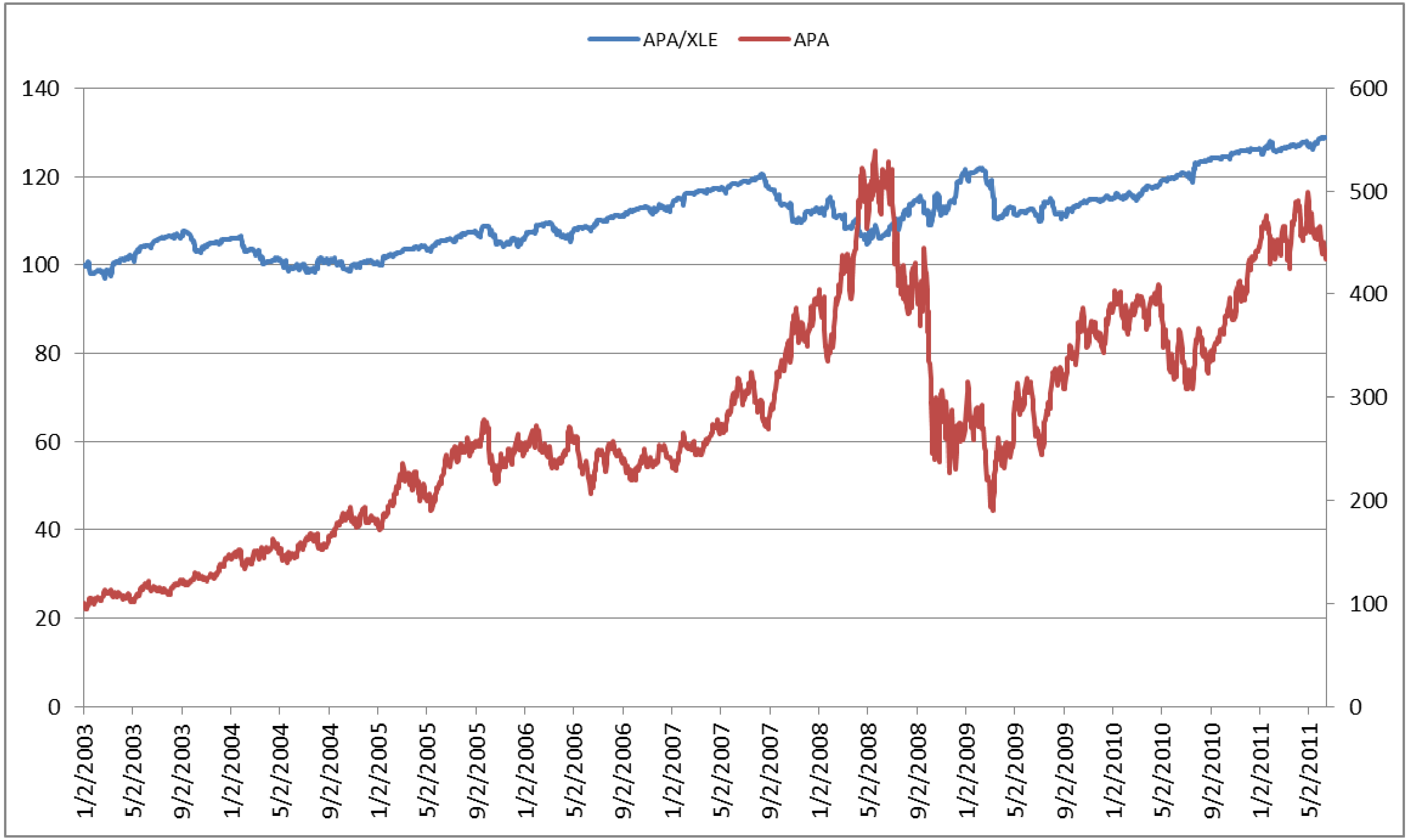
ticker	trade_date	close_date	days	stock_PNL	etf_PNL	Total_PNL	etf_ticker	size
EBAY	10/9/2009	11/9/2009	16	(3,884.70)	(2,447.11)	(6,331.81)	QQQQ	82,152.96
EBAY	5/26/2009	6/8/2009	9	497.31	(5,278.60)	(4,781.29)	QQQQ	79,388.76
EBAY	5/13/2009	5/14/2009	0	2,973.75	(1,391.58)	1,582.17	QQQQ	79,023.75
EBAY	4/16/2009	4/20/2009	1	(1,981.80)	3,673.72	1,691.92	QQQQ	79,547.25
EBAY	9/23/2009	10/7/2009	9	1,304.16	1,409.00	2,713.16	QQQQ	82,127.76
EBAY	9/8/2009	9/10/2009	1	6,439.61	(2,043.87)	4,395.73	QQQQ	81,613.02

stock_cost	stock_shares	stock_exe_price	etf_cost	etf_shares	etf_exe_price	open_s_score	close_s_score	beta
24.32	3378	23.17	42.43	(2,562.42)	43.39	(2.24)	(0.89)	0.756
17.56	4521	17.67	34.69	(2,559.32)	36.75	(1.84)	(1.00)	0.894
16.21	4875	16.82	33.00	(2,676.12)	33.52	(1.53)	(0.41)	0.895
14.45	5505	14.09	33.35	(3,251.08)	32.22	(1.27)	(0.43)	0.734
23.93	3432	24.31	42.58	(2,471.93)	42.01	(2.35)	0.00	0.780
21.81	3742	23.53	40.62	(2,579.02)	41.41	(2.26)	0.13	0.779

OXY/XLE 2003-2011
(20% stop loss)



Apache Corp. vs. XLE (20% stop loss)



Constructing market-neutral portfolios from the signals

-- Large, diversified trading universe of equities (~ 1000 names)

-- Select within the trading universe those stocks that have a trading signal (s-score) and open trades

-- Monitor for closing trades through s-score as well

-- Keep all sectors beta-neutral by using ETFs to balance the portfolio and maintain sector-neutrality

-- Leverage = 2+2 (i.e. $2 long, $2 short for $1 of capital)

-- Volatility < 7% annualized (< 50 bps/day).

The back-testing equations

$$E_{n+1} = E_n + rE_n\Delta t + \sum_i Q_{i,n}R_{i,n} - \left(\sum_i Q_{i,n}\right)r\Delta t - \sum_i \left|Q_{i,n+1} - Q_{i,n}\right|\chi$$

$$Q_{i,n} = \Lambda E_n \xi_{i,n}$$

E_n = equity in the account

r = interest rate (FF)

Δt = 1/252

$Q_{i,n}$ = dollars invested in ith stock at the beginning of period n (long or short)

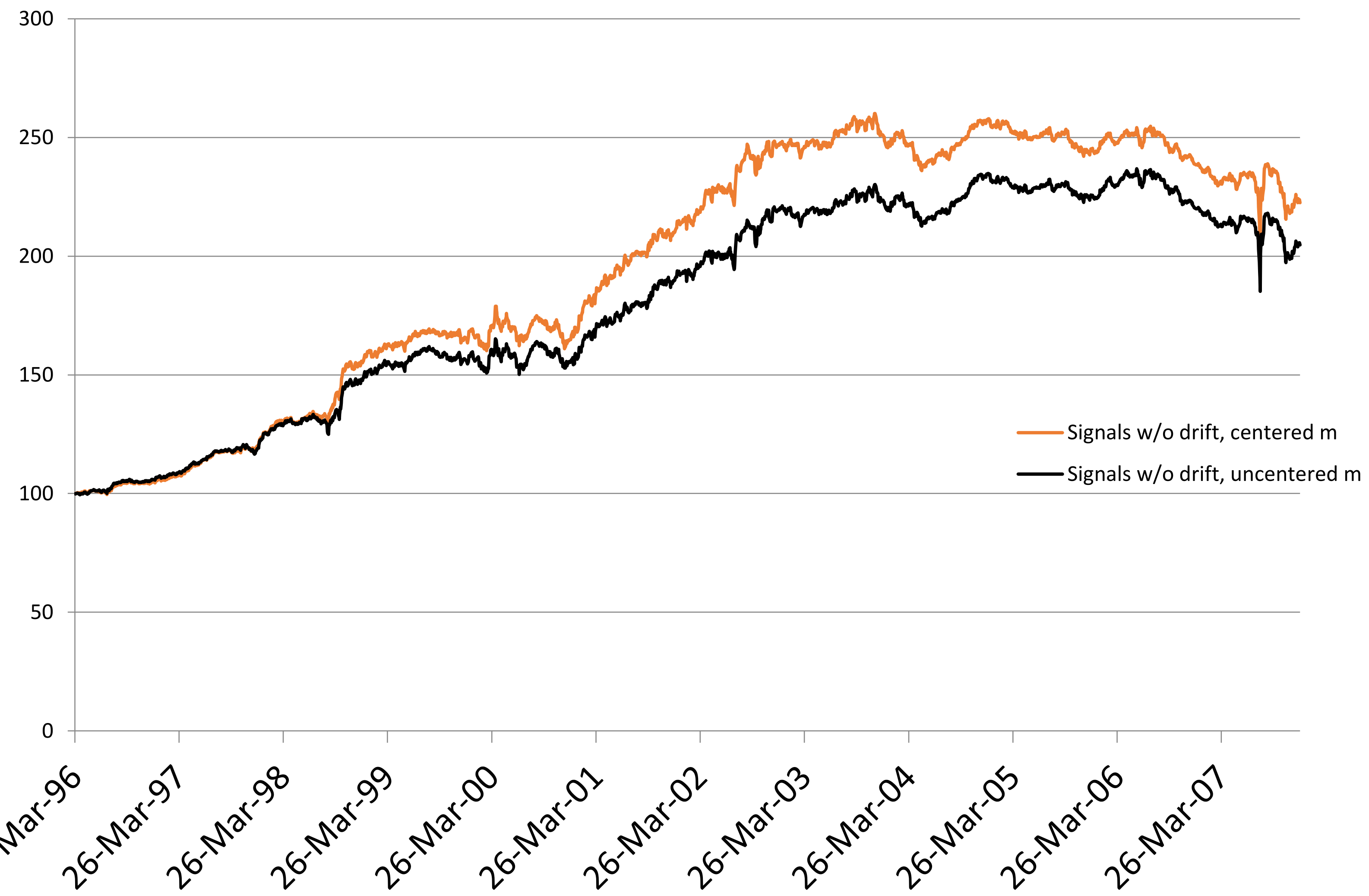
$R_{i,n}$ = return on stock i over the period, including dividends

χ = 0.0005 (5 bps), estimated slippage + transaction costs

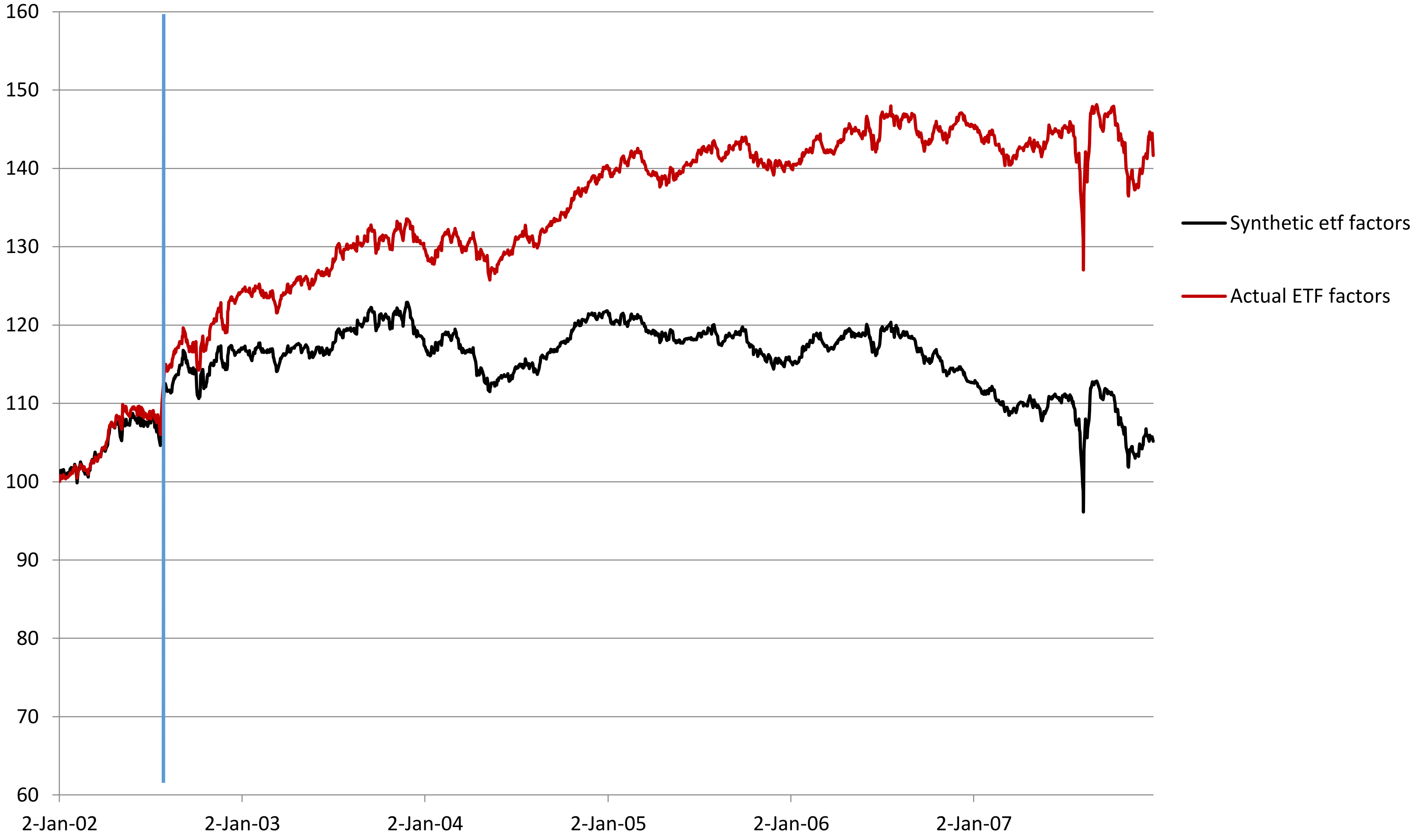
Λ = 2×0.01 (approximately), leverage coefficient

$\xi_{i,n}$ = 0, ±1, according to s-score signal

Back-testing results: **Synthetic ETFs** as factors



Signals based on **actual ETFs** are better than signals on synthetic ETFs



Annualized Returns (synthetic ETF)

	HHH	IYR	IYT	OIH	RKH	RTH	SMH	UTH	XLE	XLF	XLI	XLK	XLP	XLV	XLY	Porfolio
1996	9%	10%	-13%	8%	7%	-11%	10%	12%	3%	14%	6%	17%	-5%	15%	4%	6%
1997	1%	14%	0%	25%	16%	19%	46%	10%	-8%	23%	8%	16%	7%	19%	35%	15%
1998	15%	-7%	-8%	9%	62%	55%	47%	24%	10%	30%	18%	59%	1%	-2%	63%	25%
1999	-31%	-13%	22%	-13%	-13%	8%	37%	-14%	10%	27%	18%	41%	-21%	2%	13%	5%
2000	-18%	10%	17%	-7%	55%	3%	-35%	12%	10%	4%	-15%	36%	4%	-9%	-32%	2%
2001	-32%	35%	12%	7%	72%	42%	-40%	10%	17%	1%	45%	68%	12%	37%	84%	25%
2002	30%	20%	-1%	9%	50%	20%	-14%	-22%	-12%	28%	17%	41%	31%	1%	46%	16%
2003	9%	1%	-6%	-3%	46%	26%	-27%	-11%	1%	-5%	-11%	38%	-11%	-19%	-22%	0%
2004	10%	1%	19%	2%	19%	-7%	2%	-10%	6%	2%	11%	15%	0%	-12%	0%	4%
2005	1%	-29%	-4%	-5%	-1%	4%	6%	-18%	0%	-9%	-1%	10%	-15%	-8%	-9%	-5%
2006	-9%	-24%	-1%	-3%	22%	-8%	-3%	2%	-9%	-12%	12%	9%	-12%	-17%	17%	-2%
2007	27%	-46%	16%	-19%	-32%	-27%	3%	-19%	-11%	-25%	19%	0%	0%	-12%	31%	-6%

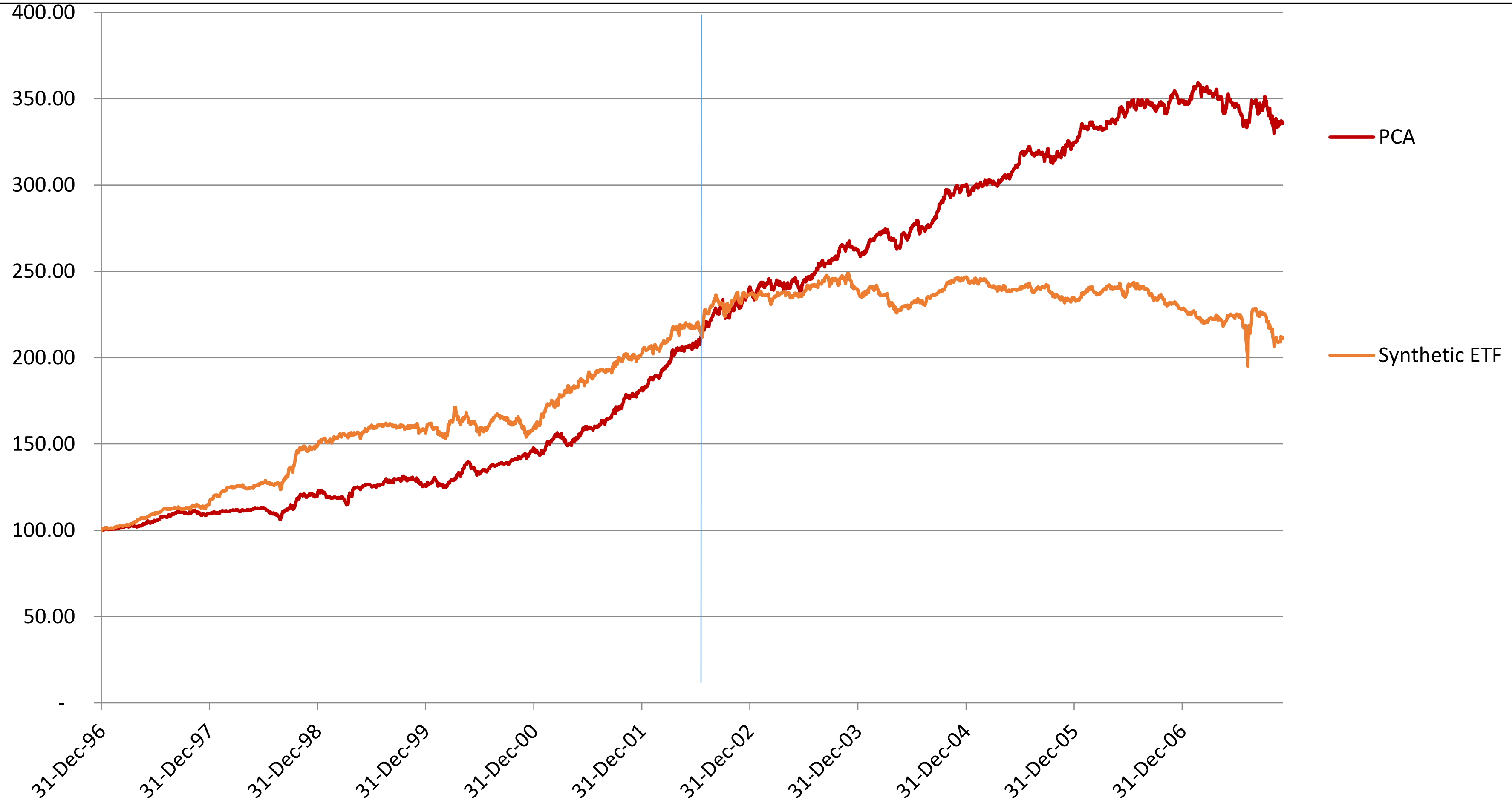
Sharpe ratio: synthetic ETFs

	HHH	IYR	IYT	OIH	RKH	RTH	SMH	UTH	XLE	XLF	XLI	XLK	XLP	XLV	XLY	Porfolio
1996	1.7	1.7	(1.2)	1.0	0.8	(0.6)	0.6	1.4	0.6	2.3	0.5	1.5	(0.5)	1.1	0.4	1.7
1997	0.1	1.5	(0.0)	2.5	1.2	1.1	2.2	1.1	(1.0)	2.3	0.6	1.1	0.4	1.5	2.9	3.6
1998	0.9	(0.5)	(0.5)	0.8	2.5	1.8	2.4	2.0	1.1	2.1	0.8	3.0	0.1	(0.1)	2.8	3.4
1999	(1.0)	(1.3)	1.5	(1.3)	(0.7)	0.3	1.2	(1.2)	1.4	1.9	1.1	1.9	(1.1)	0.1	0.6	0.8
2000	(0.4)	1.0	1.2	(0.6)	2.1	0.1	(0.7)	0.7	1.0	0.2	(0.8)	0.9	0.1	(0.5)	(1.1)	0.3
2001	(0.9)	2.8	0.7	0.6	2.7	1.5	(0.9)	0.6	1.6	0.1	1.9	1.9	0.6	1.4	3.3	2.9
2002	1.9	1.5	(0.1)	1.0	2.1	0.7	(0.5)	(1.1)	(1.3)	1.6	0.8	2.0	1.3	0.0	1.8	2.0
2003	0.5	0.0	(0.4)	(0.4)	2.6	1.3	(1.3)	(0.9)	0.1	(0.4)	(0.8)	2.5	(0.6)	(1.0)	(1.1)	0.1
2004	0.7	0.1	1.2	0.3	1.3	(0.4)	0.1	(1.1)	0.6	0.1	1.1	1.2	(0.0)	(0.8)	(0.0)	0.8
2005	0.1	(2.1)	(0.3)	(0.8)	(0.1)	0.2	0.5	(2.1)	0.0	(0.8)	(0.1)	1.0	(1.1)	(0.6)	(0.5)	(1.3)
2006	(0.7)	(1.8)	(0.1)	(0.3)	1.6	(0.4)	(0.2)	0.3	(0.7)	(1.1)	0.9	0.7	(0.9)	(1.0)	1.1	(0.5)
2007	2.1	(2.1)	0.6	(1.4)	(1.1)	(0.9)	0.1	(1.1)	(0.8)	(1.0)	1.0	(0.0)	0.0	(0.6)	1.1	(0.5)
	0.4	0.1	0.2	0.1	1.2	0.4	0.3	(0.1)	0.2	0.6	0.6	1.5	(0.2)	(0.0)	0.9	1.1

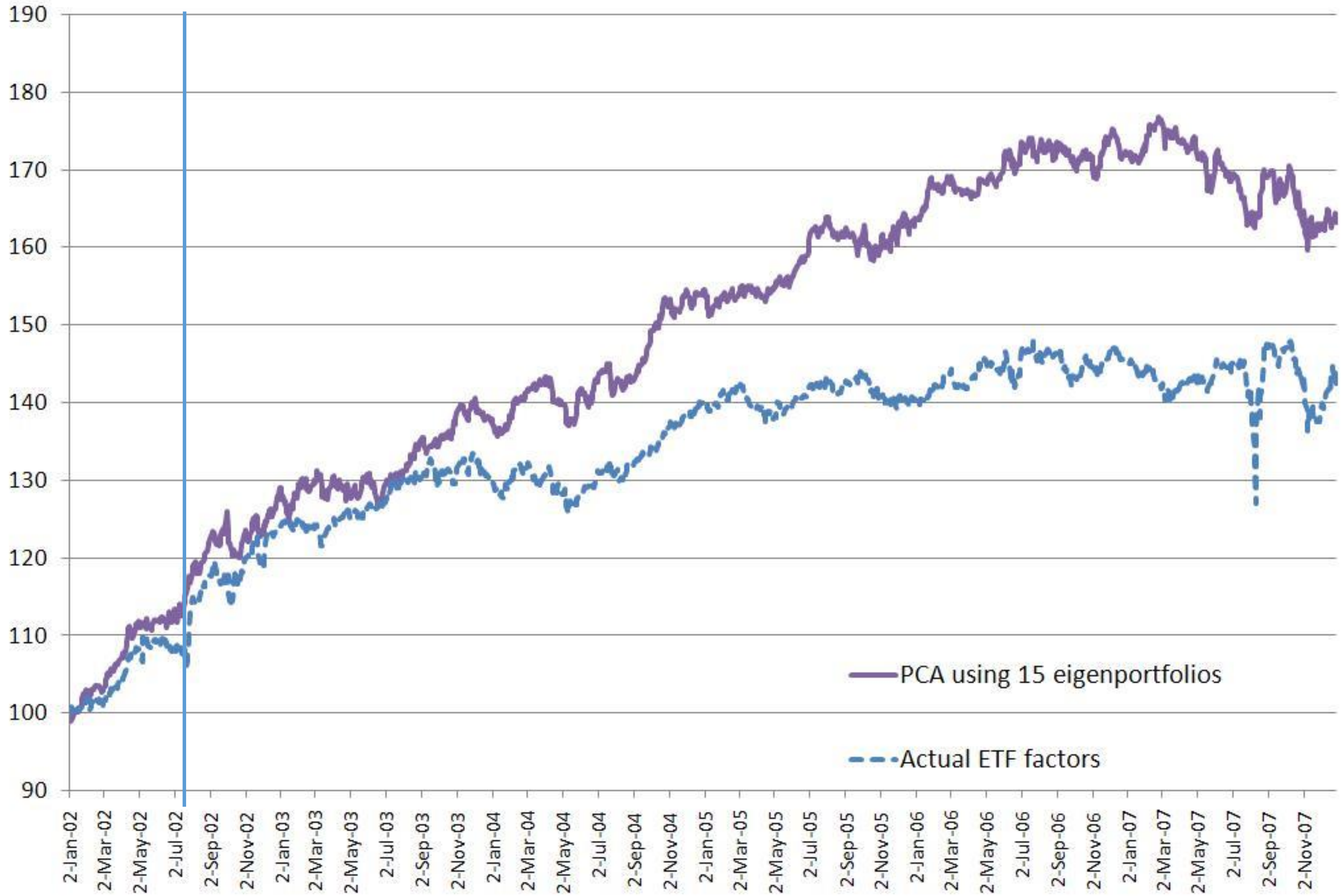
Sharpe Ratio: Actual ETFs

	HHH	IYR	IYT	OIH	RKH	RTH	SMH	UTH	XLE	XLF	XLI	XLK	XLP	XLV	XLY	Porfolio
2002	1.9	2.1	1.4	0.6	2.4	1.5	(0.7)	(0.2)	(0.2)	1.8	0.7	1.5	1.8	(0.1)	2.4	2.7
2003	(0.2)	0.8	(0.3)	(0.5)	1.4	1.1	(1.0)	(0.1)	0.5	0.6	(0.6)	2.6	0.3	(0.4)	(0.4)	0.8
2004	0.9	1.6	(0.7)	0.4	0.5	0.1	0.2	(0.4)	0.6	0.6	1.4	1.9	0.5	(0.6)	0.3	1.6
2005	0.3	(1.5)	0.8	(0.6)	0.3	0.5	0.5	(1.1)	(0.1)	0.9	0.6	1.3	(0.7)	0.2	0.0	0.1
2006	(0.2)	(1.3)	0.0	(0.2)	0.9	(0.1)	0.5	1.7	(0.5)	(0.6)	1.7	1.7	(0.0)	(0.4)	2.0	0.7
2007	(0.4)	(0.3)	0.0	(1.3)	(1.2)	(0.7)	0.9	(0.7)	(1.0)	(0.6)	1.1	0.6	0.4	(0.5)	1.3	(0.2)
	0.4	0.2	0.2	(0.3)	0.7	0.4	0.1	(0.1)	(0.1)	0.5	0.8	1.6	0.4	(0.3)	0.9	0.9

Signals based on 15 PCA factors outperform synthetic ETFs (1997-2007)



PCA 15 Factors vs. Actual ETFs (2002-2007)



15 PCA factors: performance measures

year	AnnRetPL	AnnVolPL	SharpePL
1997	9%	5%	1.73
1998	11%	6%	1.71
1999	3%	7%	0.49
2000	16%	7%	2.44
2001	22%	8%	2.86
2002	28%	7%	3.73
2003	8%	7%	1.23
2004	14%	5%	2.56
2005	8%	5%	1.53
2006	7%	5%	1.42
2007	-6%	16%	(0.36)

Sensitivity to number of PCA factors

-- The separation between systematic and idiosyncratic components is a theoretical construct. **It is not observable!**

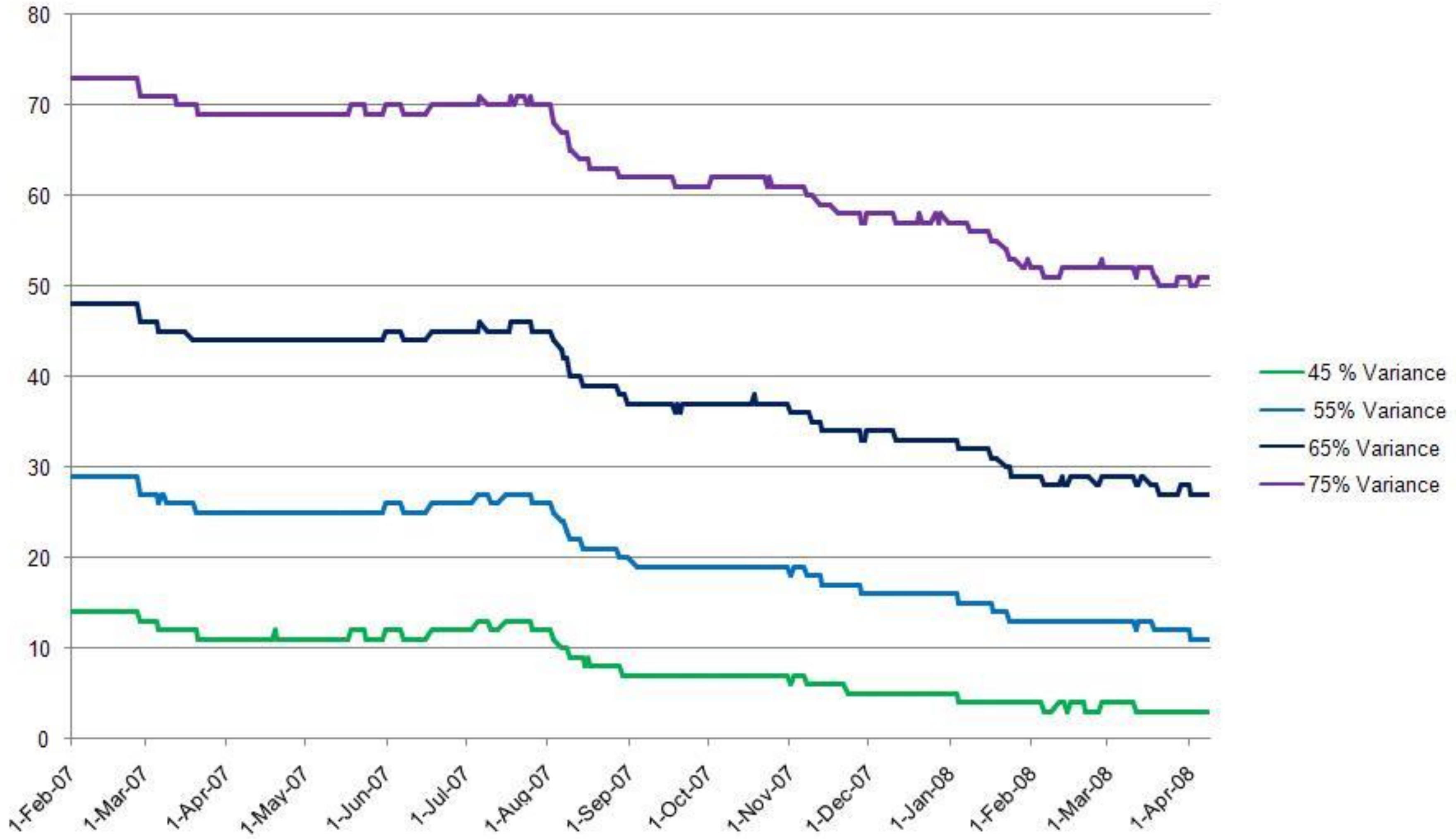
-- Tradeoff between ``systematic'' and ``idiosyncratic'' model of stock returns

-- **Too few factors** (think CAPM) lead to large variance of residuals and low sensitivity to size, industry, etc.

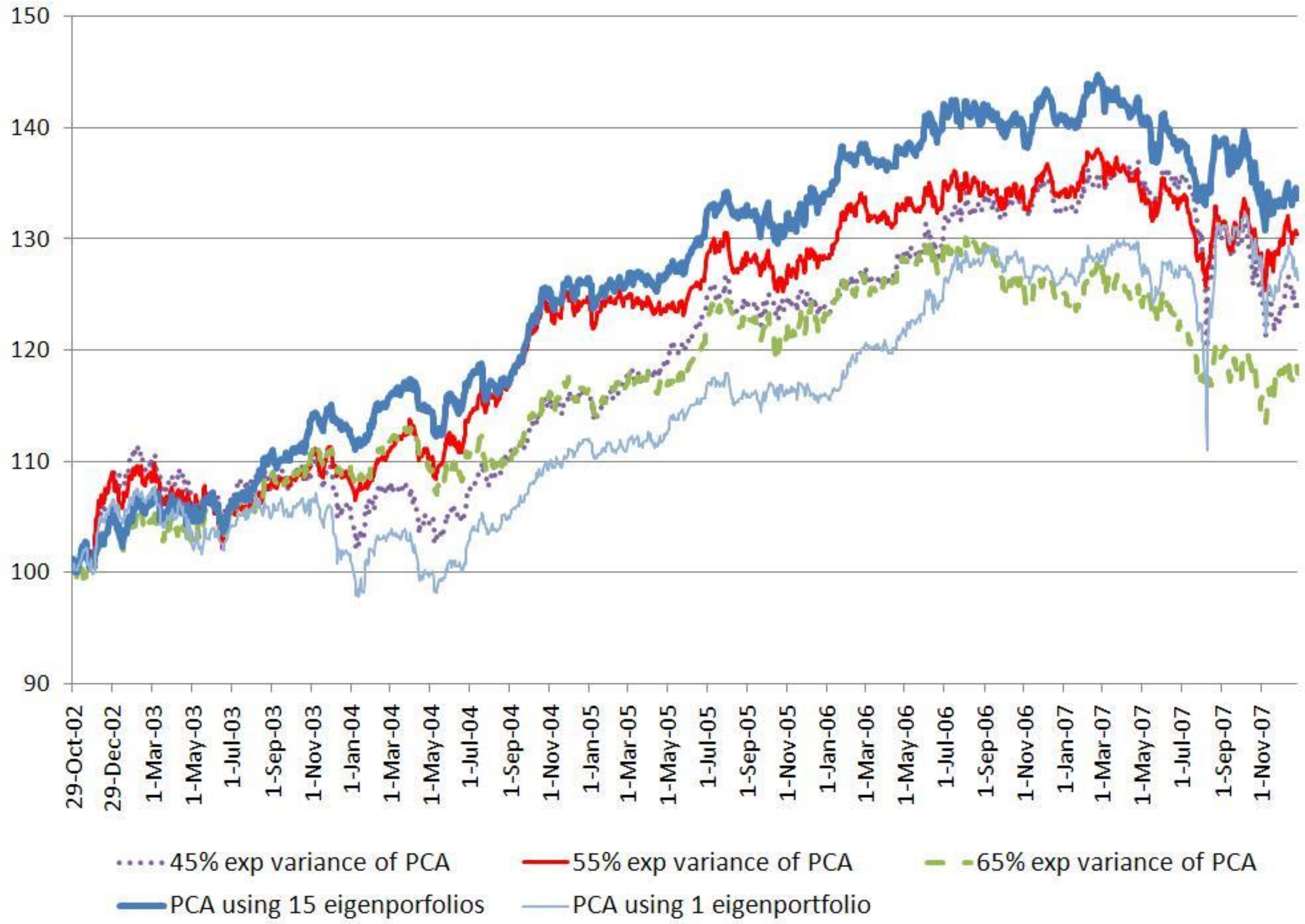
-- **Too many factors** lead to noise trading (negative P/L).

-- Interesting question: is the ``correct'' number of factors variable? (I.e. dependent on market conditions?) . Yes.

Evolution of the Number of Factors for different levels of explained variance



P/L for different truncation levels



Sharpe Ratios for Variable PCA strategies (2003-07)

	1 Eigenportfolios	15 Eigenportfolios	45 % Exp Variance	55% Exp Variance	65% Exp Variance
2003	-0.7	0.9	-0.5	-0.1	0.4
2004	1.7	2.2	1.7	2.6	1.3
2005	0.8	1.2	1.3	0.6	1.0
2006	1.8	1.0	1.3	0.9	0.3
2007	0.0	-0.7	-0.7	-0.4	-0.9
Since Inception	0.7	0.9	0.6	0.7	0.4

Trading Time vs. Actual Time, II

Using the daily trading volume, construct a residual process which measures the change in price **per share**

$$\varepsilon = \frac{\Delta S}{S} - \beta \frac{\Delta I}{I} \qquad \text{(usual residual)}$$

$$\bar{\varepsilon} = \frac{<\Delta V>}{\Delta V}\varepsilon, \qquad \Delta V = \text{daily volume} \qquad <\Delta V> = \text{average } \Delta V$$

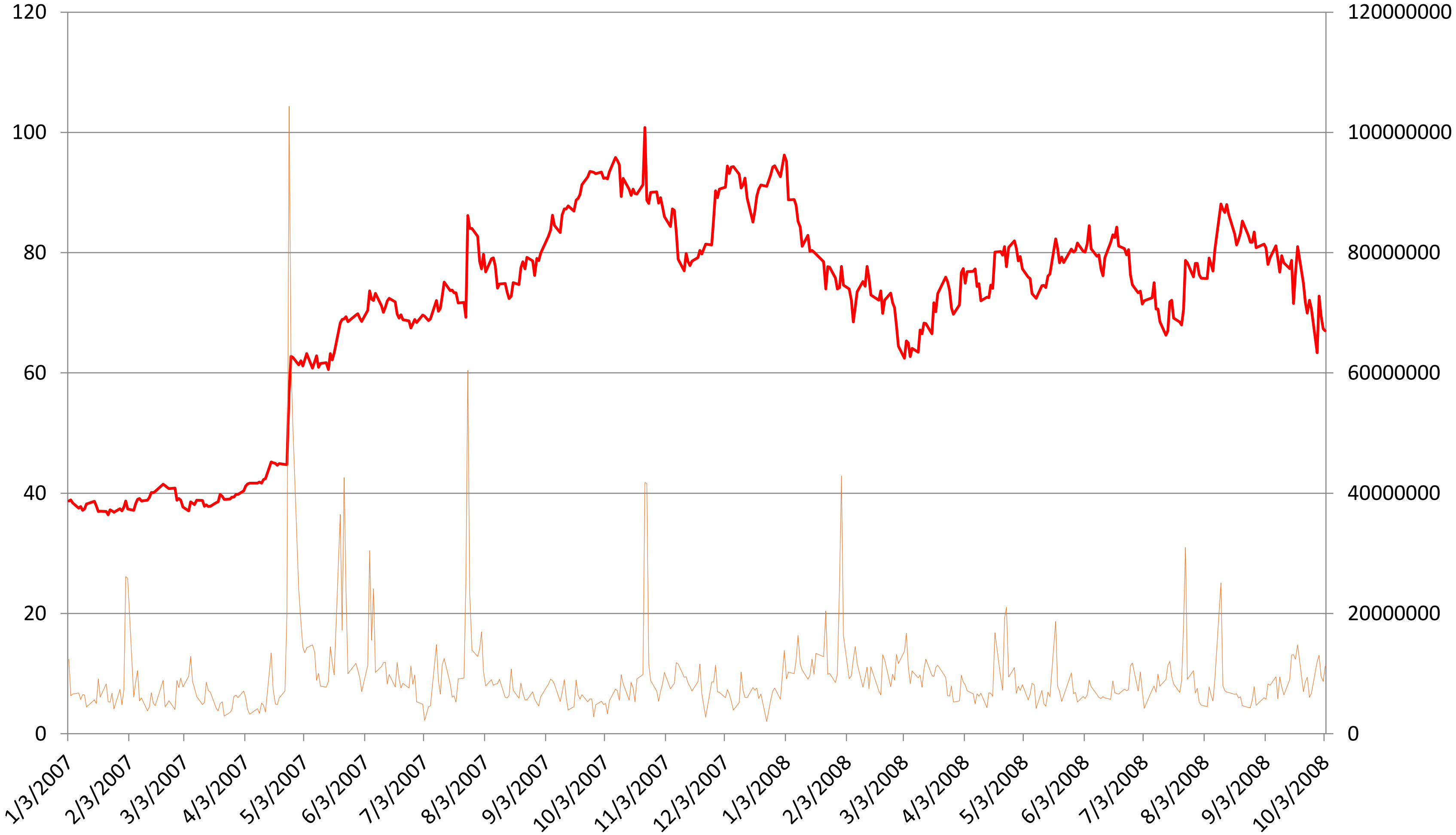
$$Y_t = \sum_{i=1}^{t} \bar{\varepsilon}_i$$

$$dY = \kappa(m - Y)dt + \sigma dW$$

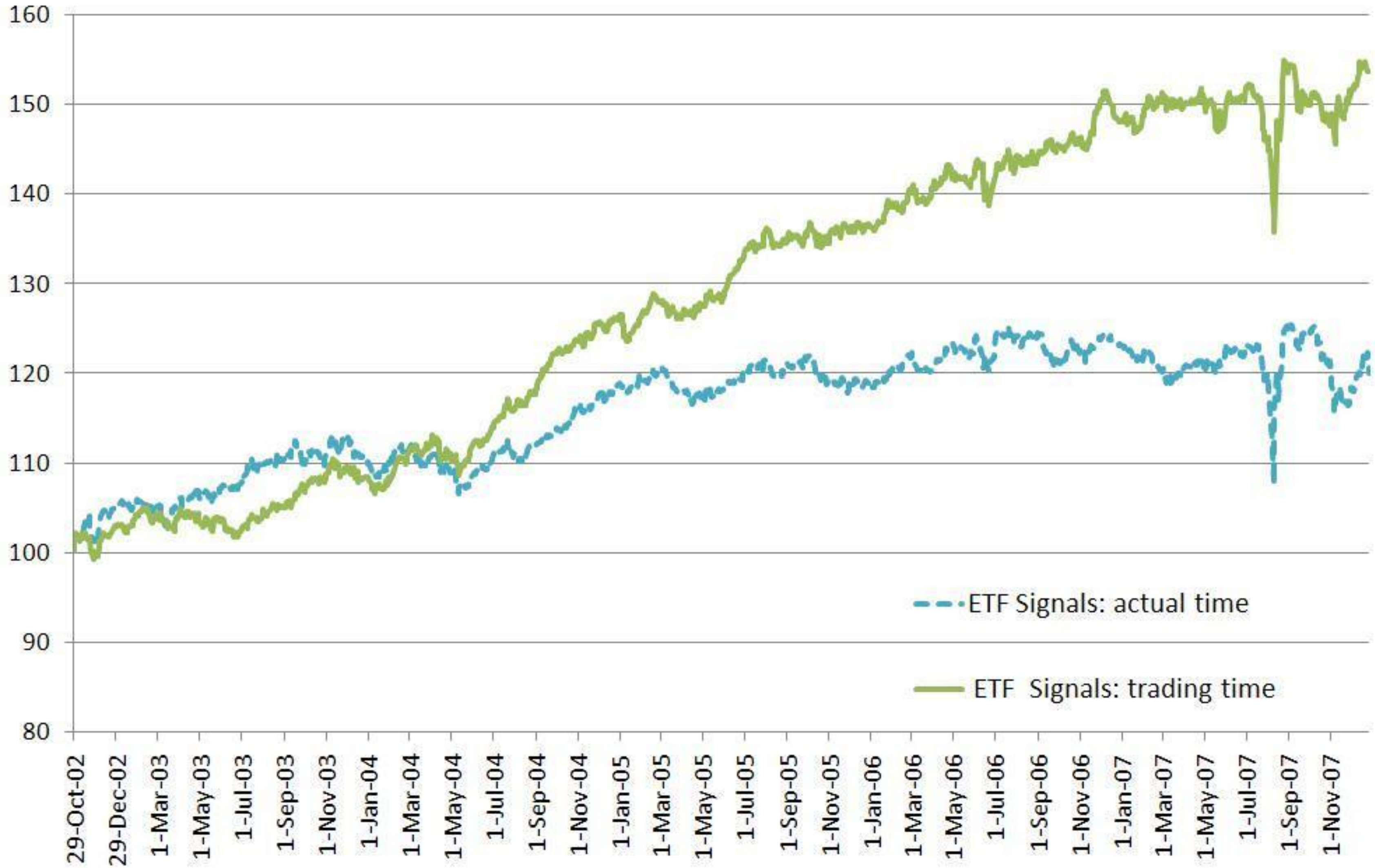
Estimate AR-1 / OU process for the new process Y(t)

This makes deviations on unusually high-volume more likely, so the signal is weaker

Amazon.com Jan 2007-Oct 2008: Avoiding short-selling on large volume



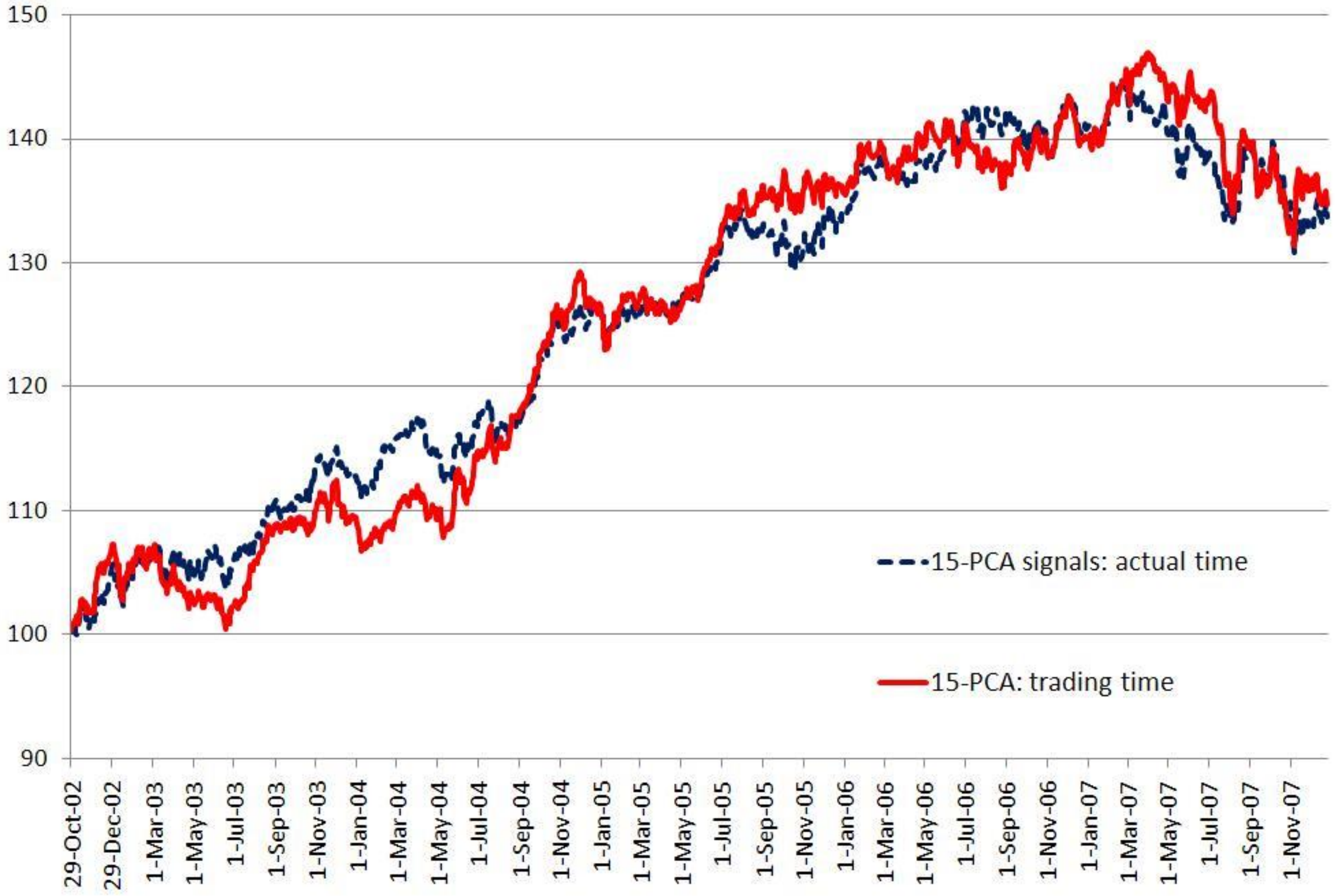
ETF signals: trading time vs. actual time



ETF signals in trading time: Sharpe Ratios

	HHH	IYR	IYT	OIH	RKH	RTH	SMH	UTH	XLE	XLF	XLI	XLK	XLP	XLV	XLY	Porfolio
2003	0.3	1.4	(1.4)	0.1	(0.1)	0.9	(1.4)	0.1	0.0	(0.8)	1.1	2.1	(0.1)	0.0	1.0	0.9
2004	0.9	2.4	0.2	0.9	1.7	1.8	0.1	0.2	1.8	1.4	1.0	1.6	0.4	0.5	0.4	3.1
2005	0.1	0.1	1.3	0.6	(0.2)	(0.2)	0.8	0.4	1.2	1.5	0.7	2.6	(1.4)	2.1	(0.2)	1.6
2006	2.4	1.6	(1.5)	0.3	1.1	(0.6)	(0.5)	1.3	0.7	0.4	1.6	1.1	(0.5)	(0.0)	1.9	1.5
2007	1.7	(2.4)	0.4	0.8	(1.0)	(0.4)	0.4	0.7	0.7	(1.2)	1.9	2.0	0.3	1.8	0.7	0.4
Since 2003	1.08	0.61	(0.21)	0.52	0.30	0.29	(0.12)	0.54	0.88	0.28	1.26	1.91	(0.28)	0.87	0.76	1.51

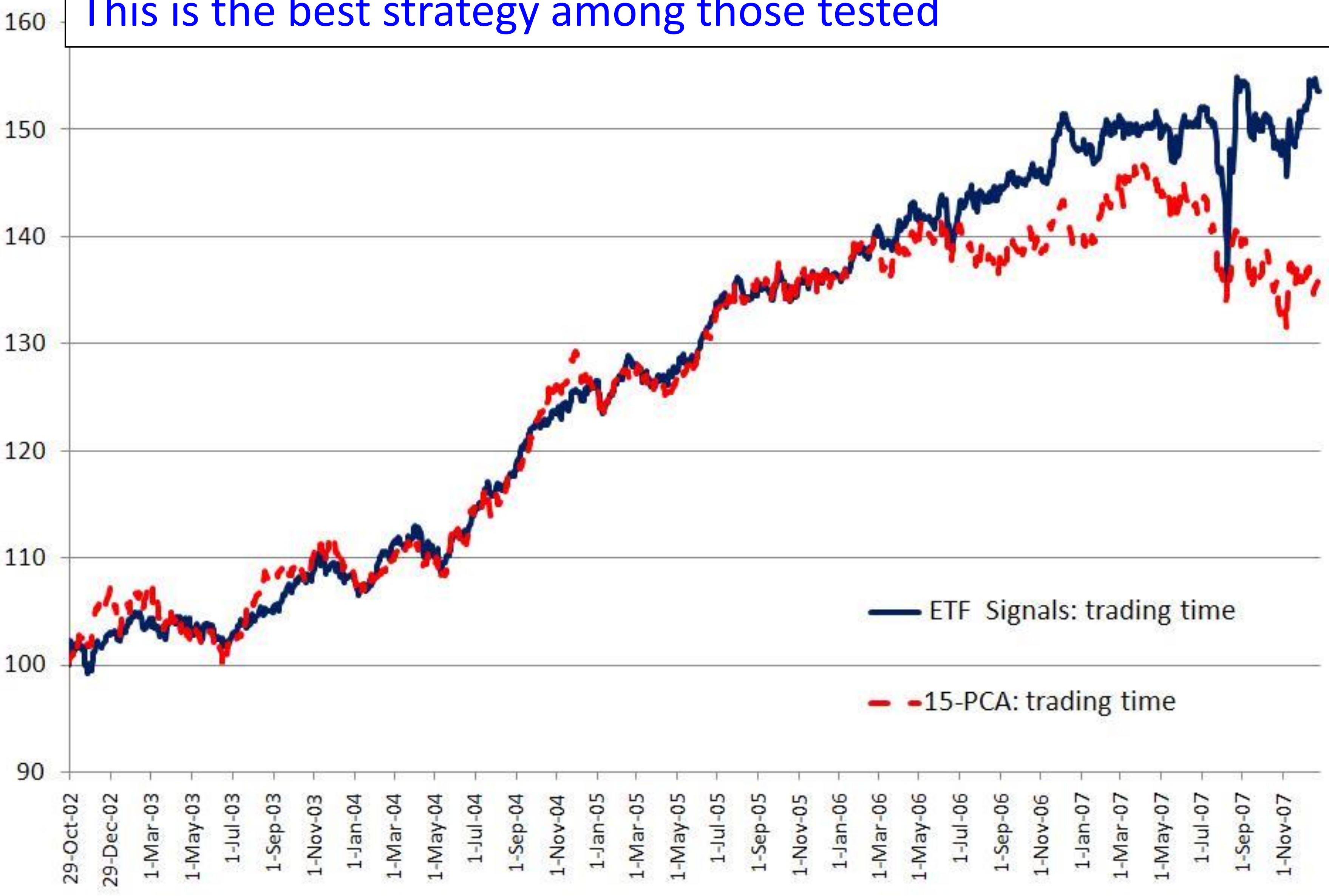
PCA: trading time vs. actual time



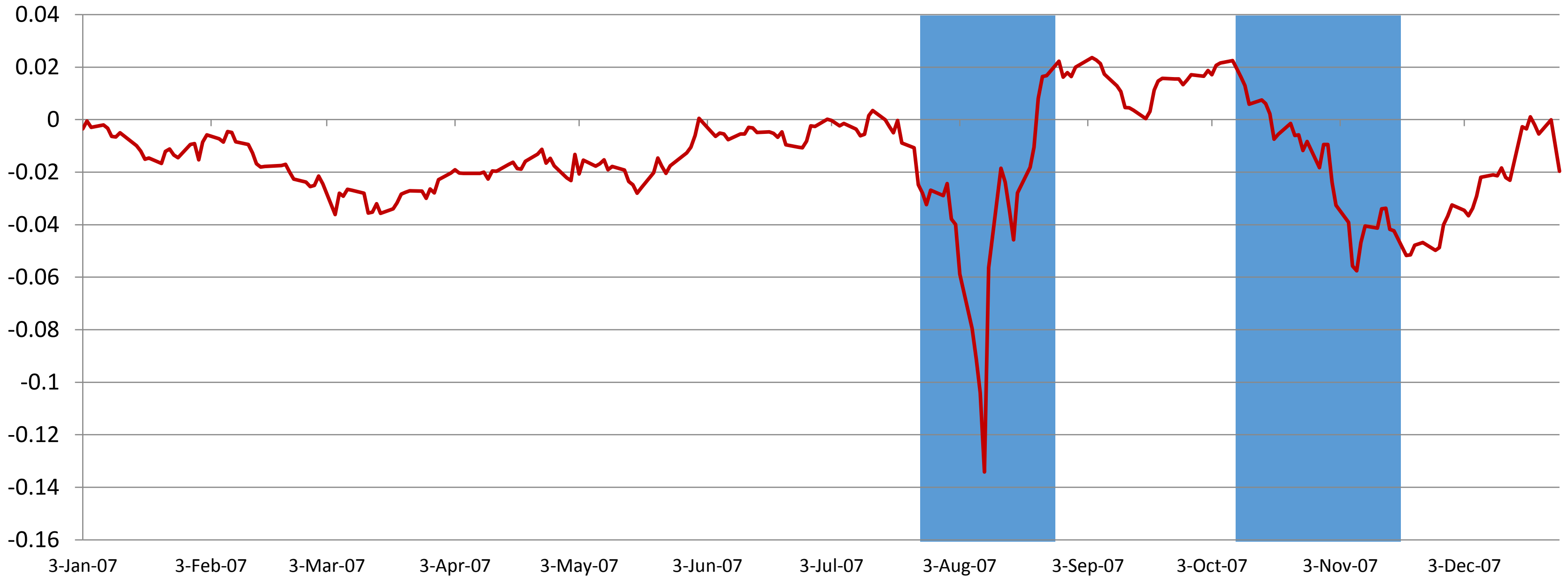
15 PCA signals in trading time: Sharpe Ratios

	HHH	IYR	IYT	OIH	RKH	RTH	SMH	UTH	XLE	XLF	XLI	XLK	XLP	XLV	XLY	Porfolio
2003	0.9	(0.1)	(0.0)	0.6	0.4	0.6	(1.7)	(1.3)	(0.0)	0.9	0.3	1.7	(0.5)	(0.4)	1.5	0.2
2004	1.5	1.8	0.6	0.9	0.9	0.8	(0.5)	0.2	1.7	1.9	1.1	1.4	0.7	0.1	1.4	2.4
2005	(1.1)	(0.8)	0.8	1.0	0.1	1.9	0.4	0.6	1.4	0.8	1.4	1.5	(1.7)	1.7	(0.6)	1.2
2006	0.3	1.2	(0.8)	1.1	0.9	(0.9)	(0.1)	0.8	(0.5)	0.2	0.4	0.0	(0.1)	0.3	0.3	0.6
2007	(0.2)	(0.7)	(0.6)	1.0	(1.1)	(1.9)	0.2	0.8	0.9	(1.7)	2.6	0.9	0.4	0.9	(1.4)	(0.5)
	0.29	0.27	(0.00)	0.93	0.22	0.13	(0.35)	0.21	0.69	0.42	1.16	1.10	(0.23)	0.50	0.24	0.80

ETFs in trading time outperformed PCAs after 2005
This is the best strategy among those tested

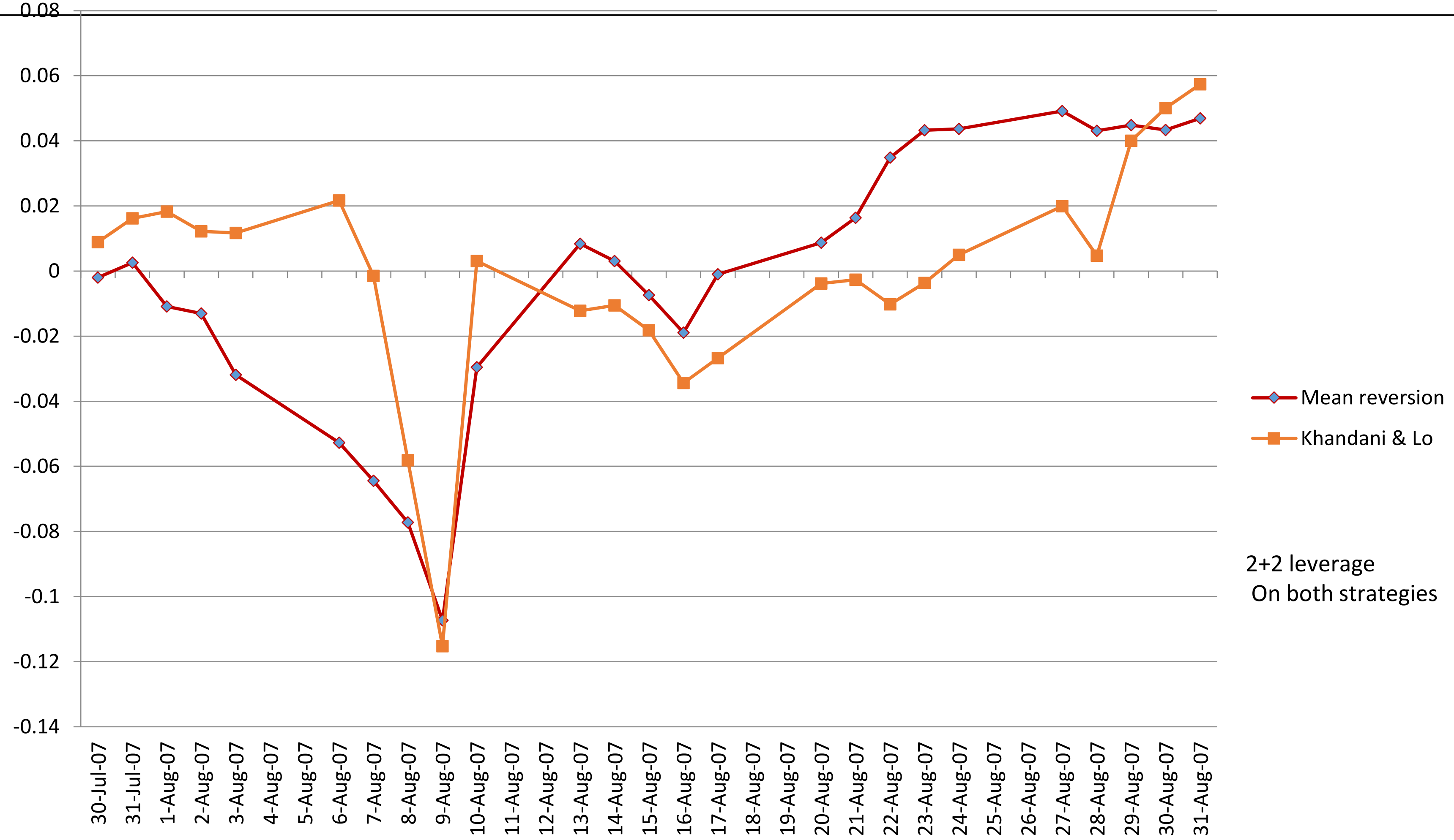


Focus on August 2007



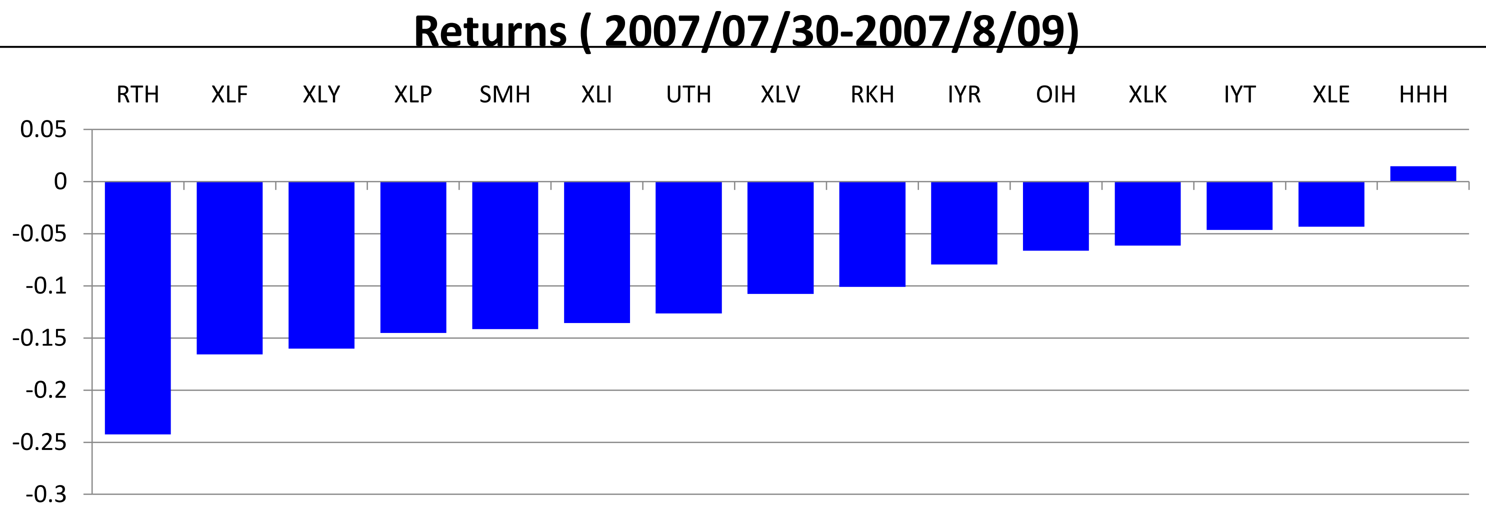
2007: Flat/small losses year with ``cataclysmic'' dips in August and November. Leverage=2+2

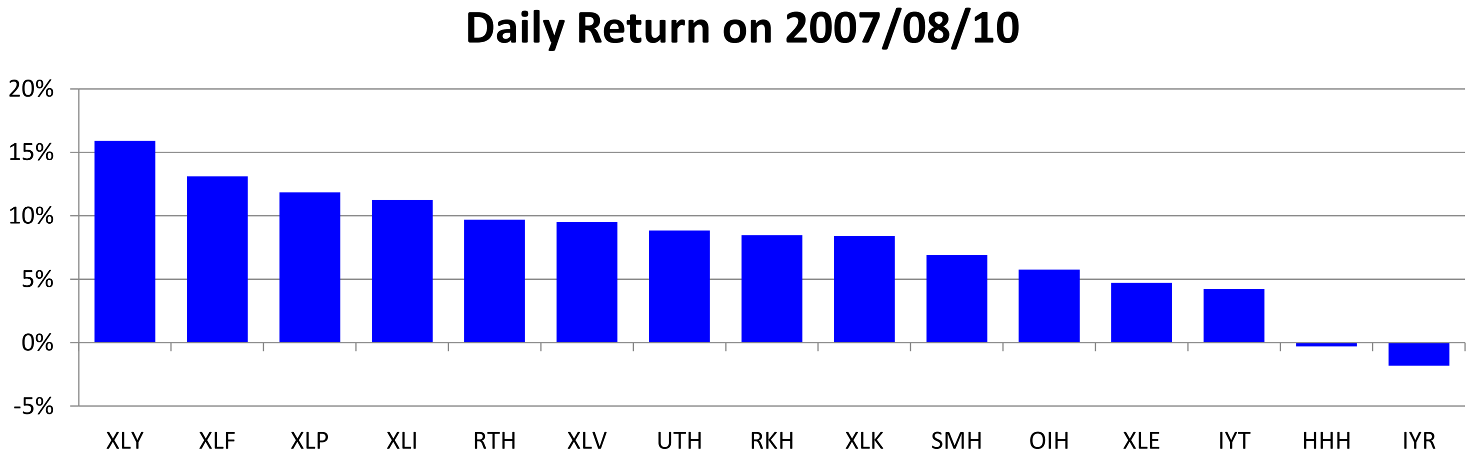
August 2007: comparison with Khandani & Lo



2+2 leverage
On both strategies

Sector view in Aug 2007





Conclusions

-- We presented a general `blueprint' for market-neutral statistical-arbitrage strategies

-- Market-neutrality is enforced using PCA factors or ETFs (synthetic or actual)

-- Signals are based on price mean-reversion around the industry sector ETFs or synthetic PCA factors

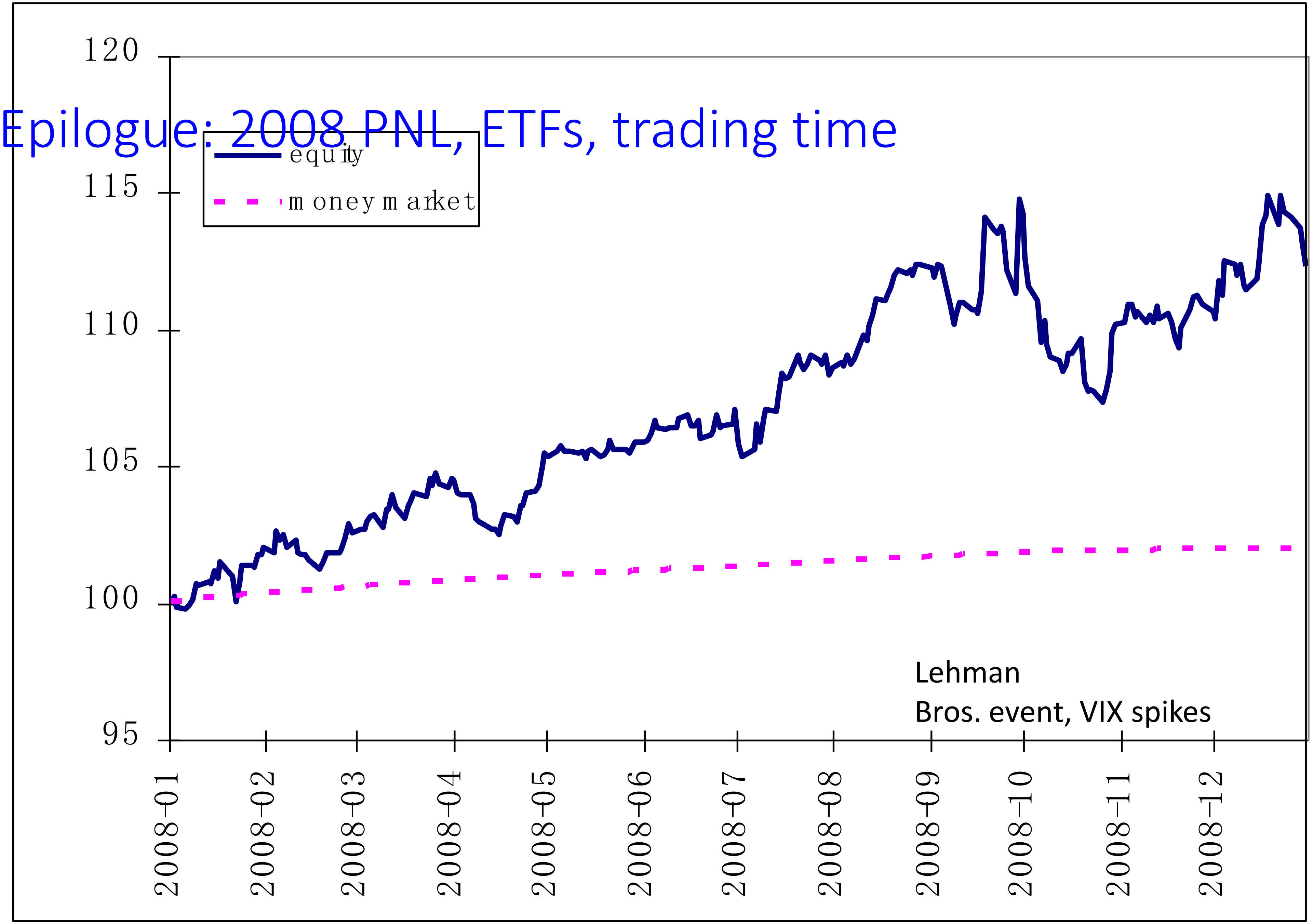
-- Performance depends significantly on which method is used to generate stock residuals (de-factoring).

-- Conjecture :quant MN works better when volatility is high, but steady.

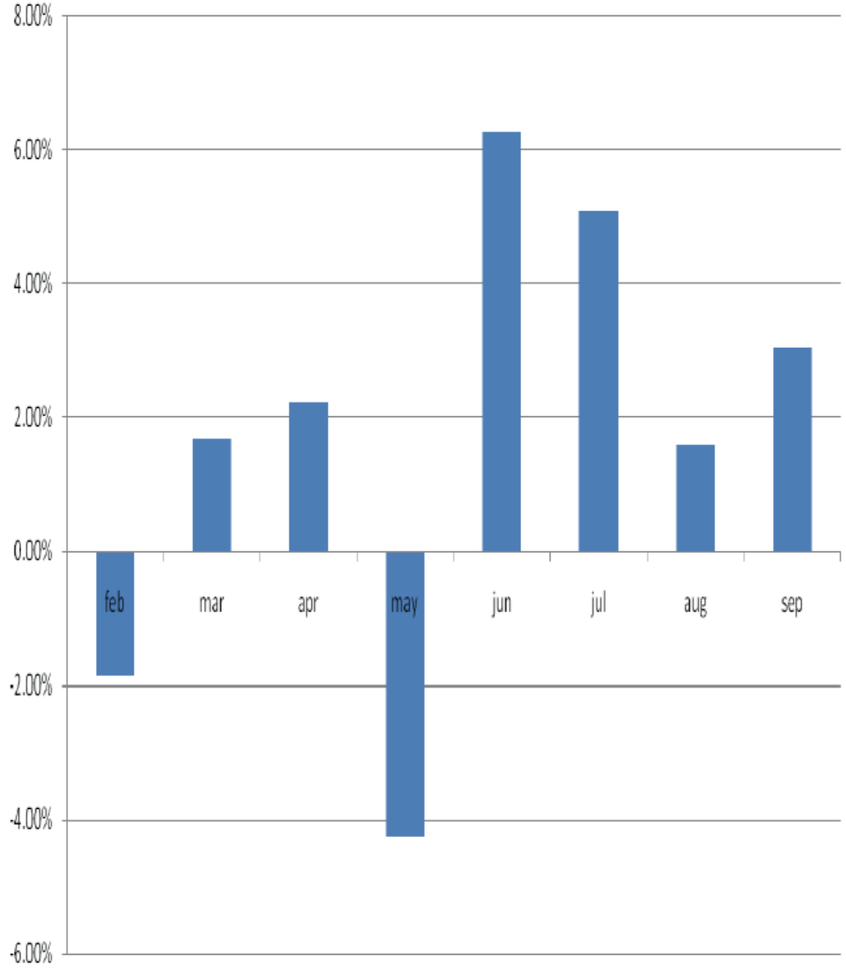
-- There are significant improvements in performance if we use ``trading time''

-- Parallels with Khandani-Lo suggest that this strategy is subject to periodic very sharp ``commodity-like'' draw-downs due to liquidity trap. Usually, the strategy recovers, but leverage can be problematic.

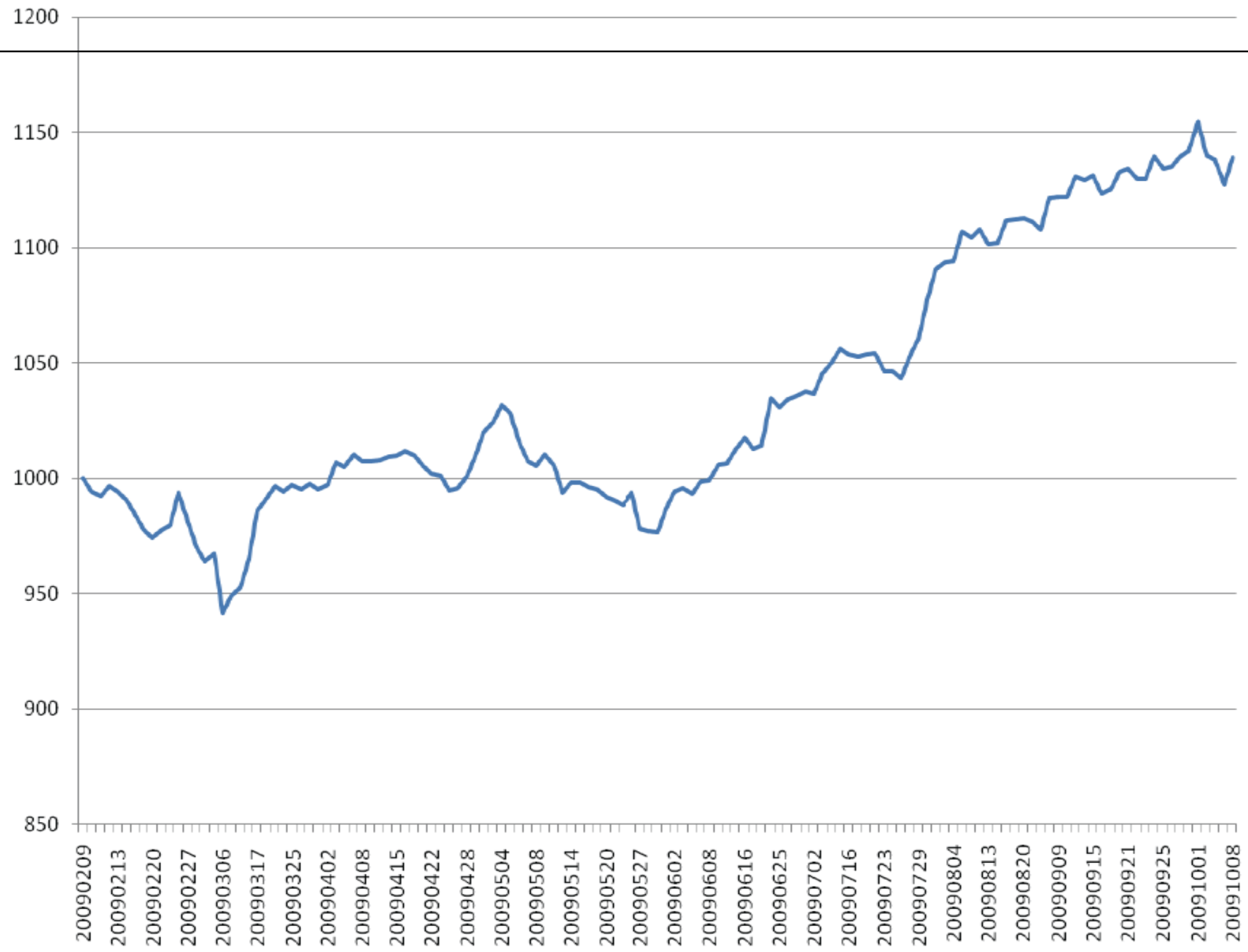
Epilogue: 2008 PNL, ETFs, trading time



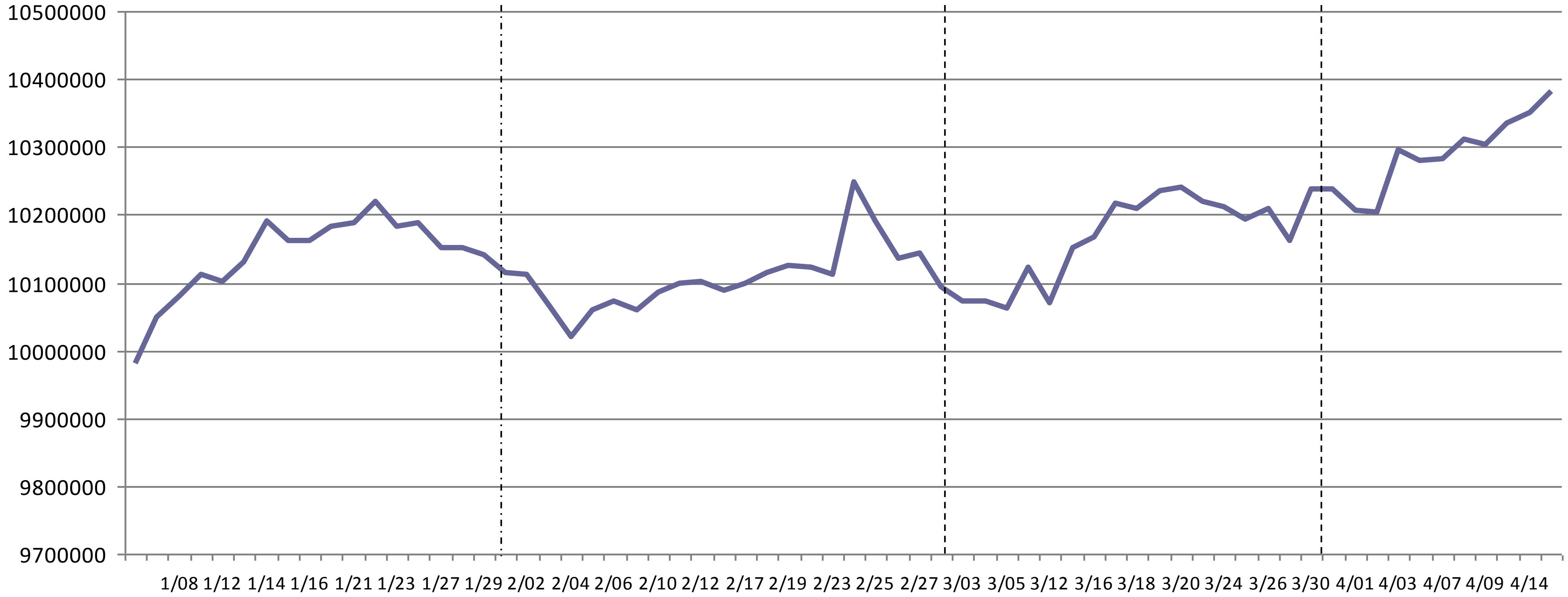
Monthly performance 2/2009-9/2009



Daily performance 2/2009-9/2009



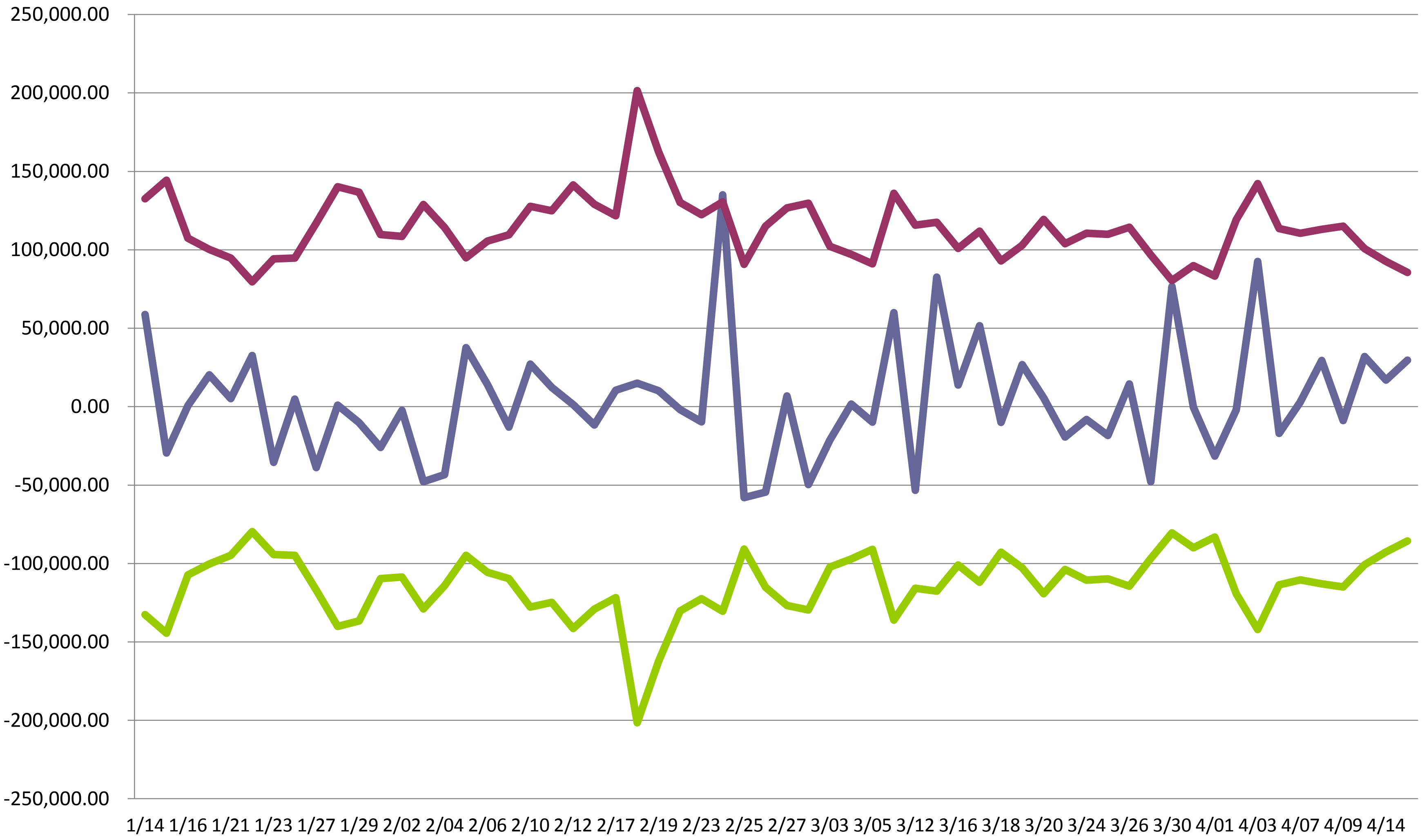
Performance: Jan 2009-April 2009



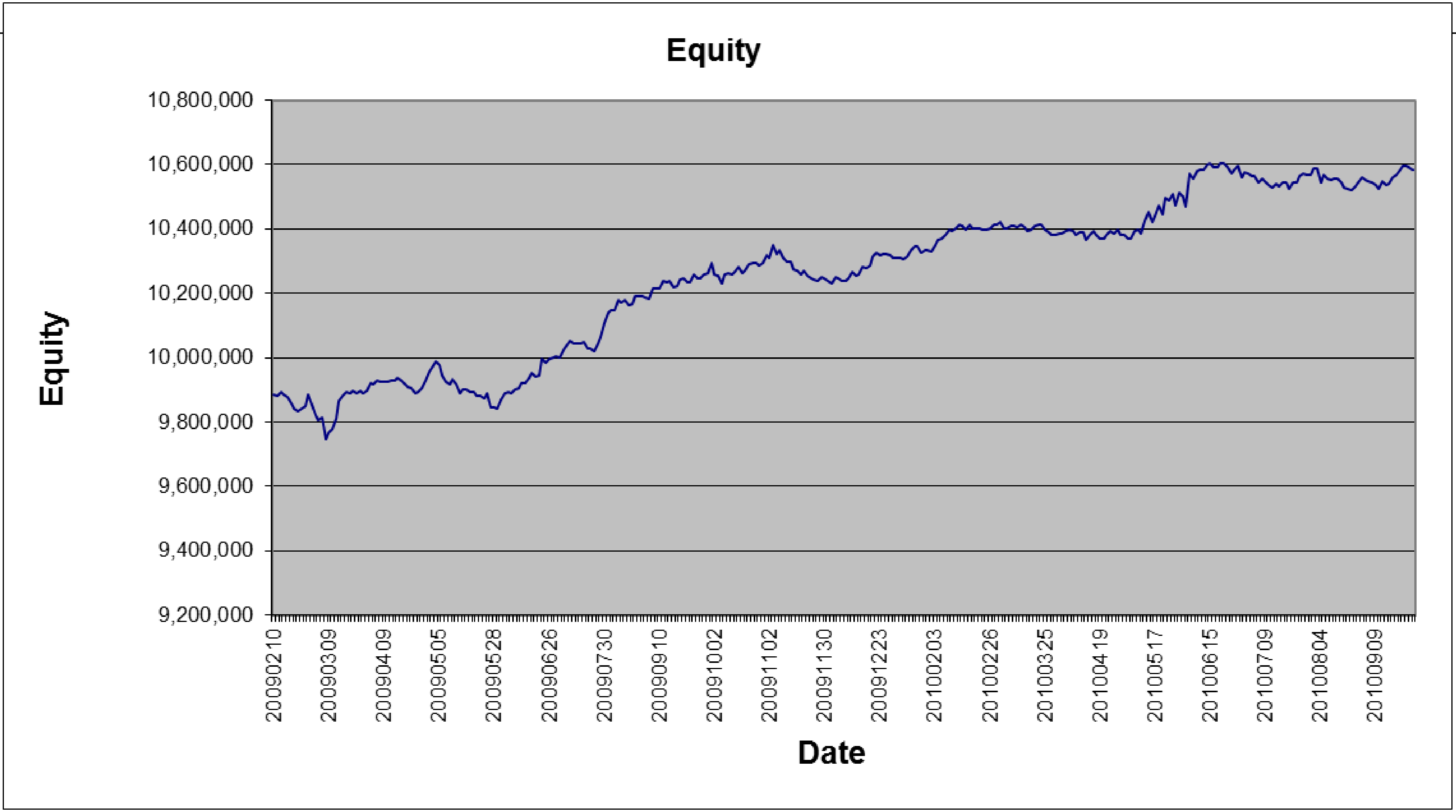
Trading days:	67
Number of trades:	142
Average P/L per trade:	3.7 bps
Average holding time:	5 days

Average P/L per day:	6.8 bps
Stdev. PL per day:	37 bps
Sharpe Ratio:	2.66
Target VaR	150 bps/day
Leverage:	0.5+0.5

2009: Daily Returns vs Target VaR



Feb 2009 to Nov 2010



Portfolio Risk-Management

$Q_i \quad i = 1,\ldots, N$ holdings in stock or ETF i.

$V(Q_1, Q_2, \ldots, Q_N) = \sum_{ij=1}^{N} C_{ij} Q_i Q_j$ daily portfolio variance

C: covariance matrix estimated with 3m window

$\mathbf{Q} = (Q_1, Q_2, \ldots, Q_N)$ position at date T

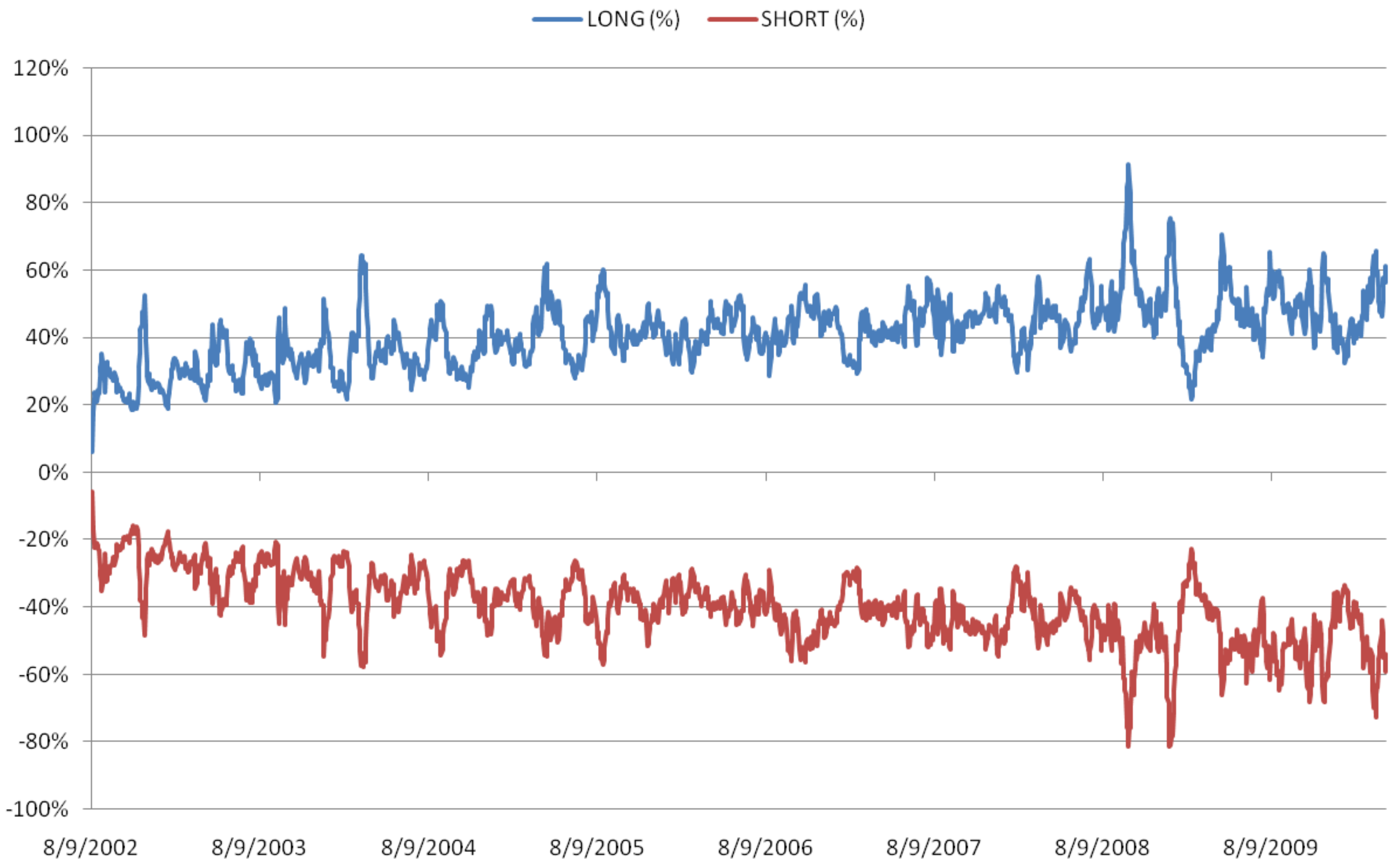
$\mathbf{Q}' = (Q_1', Q_2', \ldots, Q_N')$ target (desired) position date $T+1$

<u>Risk-control algorithm</u> : let ε denote a risk threshold (e.g. 25bps of daily stdev)

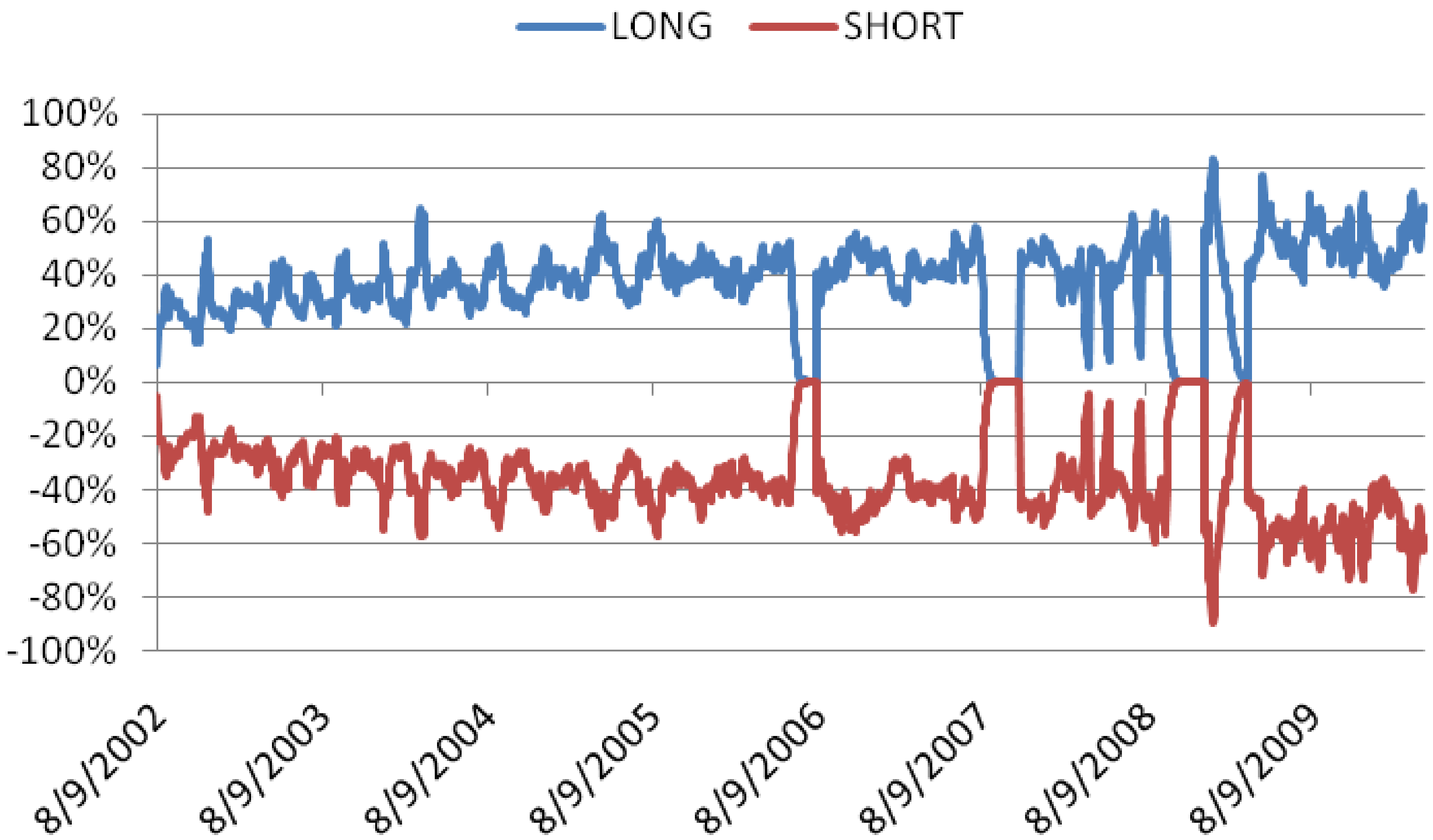
If $V(\mathbf{Q}') > \varepsilon$ then only execute `closing trades' $\mathbf{Q} \to \mathbf{Q}'$

If $V(\mathbf{Q}') \leq \varepsilon$ then execute all trades $\mathbf{Q} \to \mathbf{Q}'$

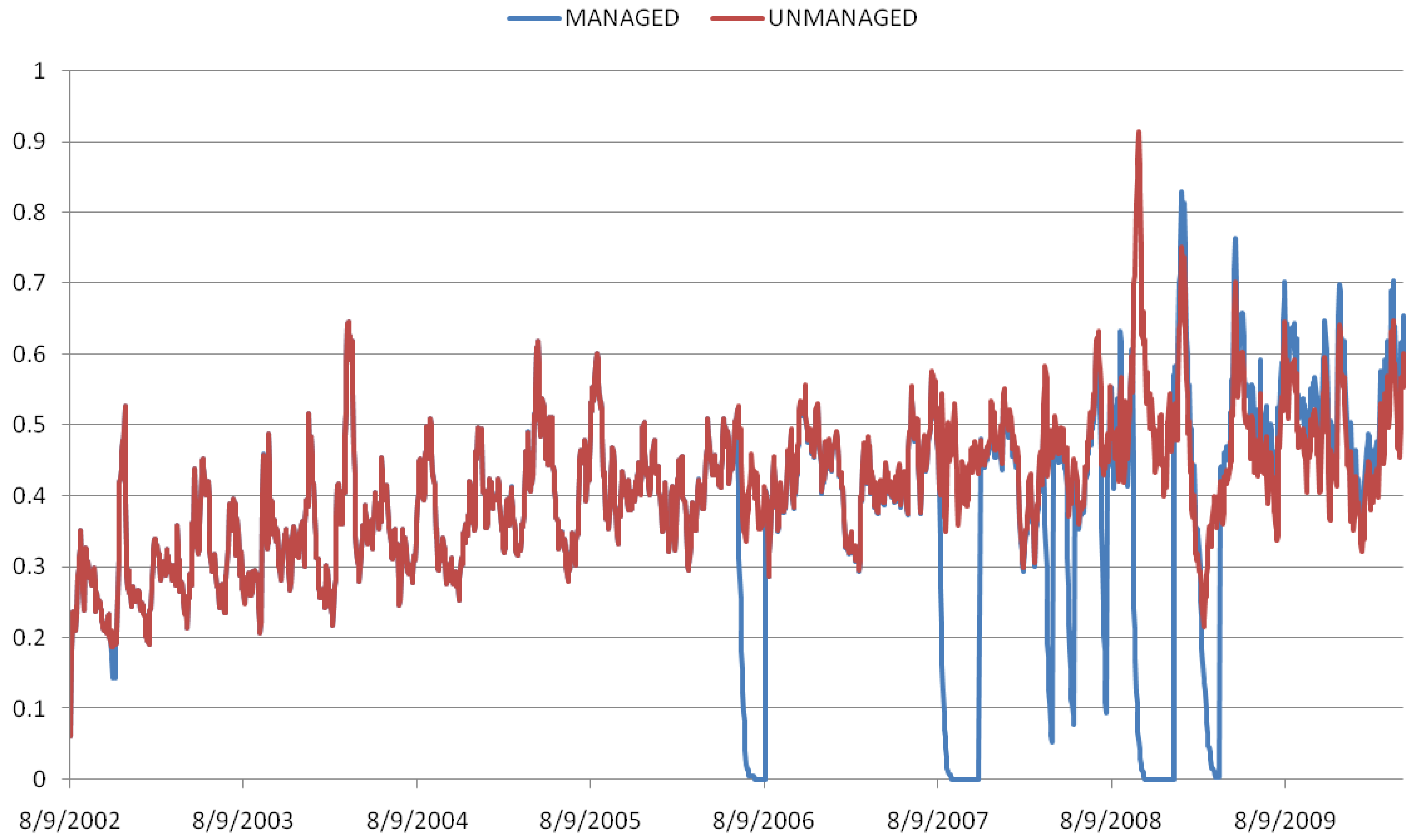
Long and Short positions for unmanaged portfolio (% of account equity)



Long/short balances with portfolio risk-management



Comparison between Long balances: managed & unmanaged simulations



Building a portfolio from ETF-based signals: the ``PLATA'' strategy

-- Large, diversified trading universe of equities (~ 500 names)

-- Select those stocks within the trading universe that have a trading signal (s-score) and open trades

-- All trades consist of stocks paired with ETFs

-- Monitor for closing trades through s-score

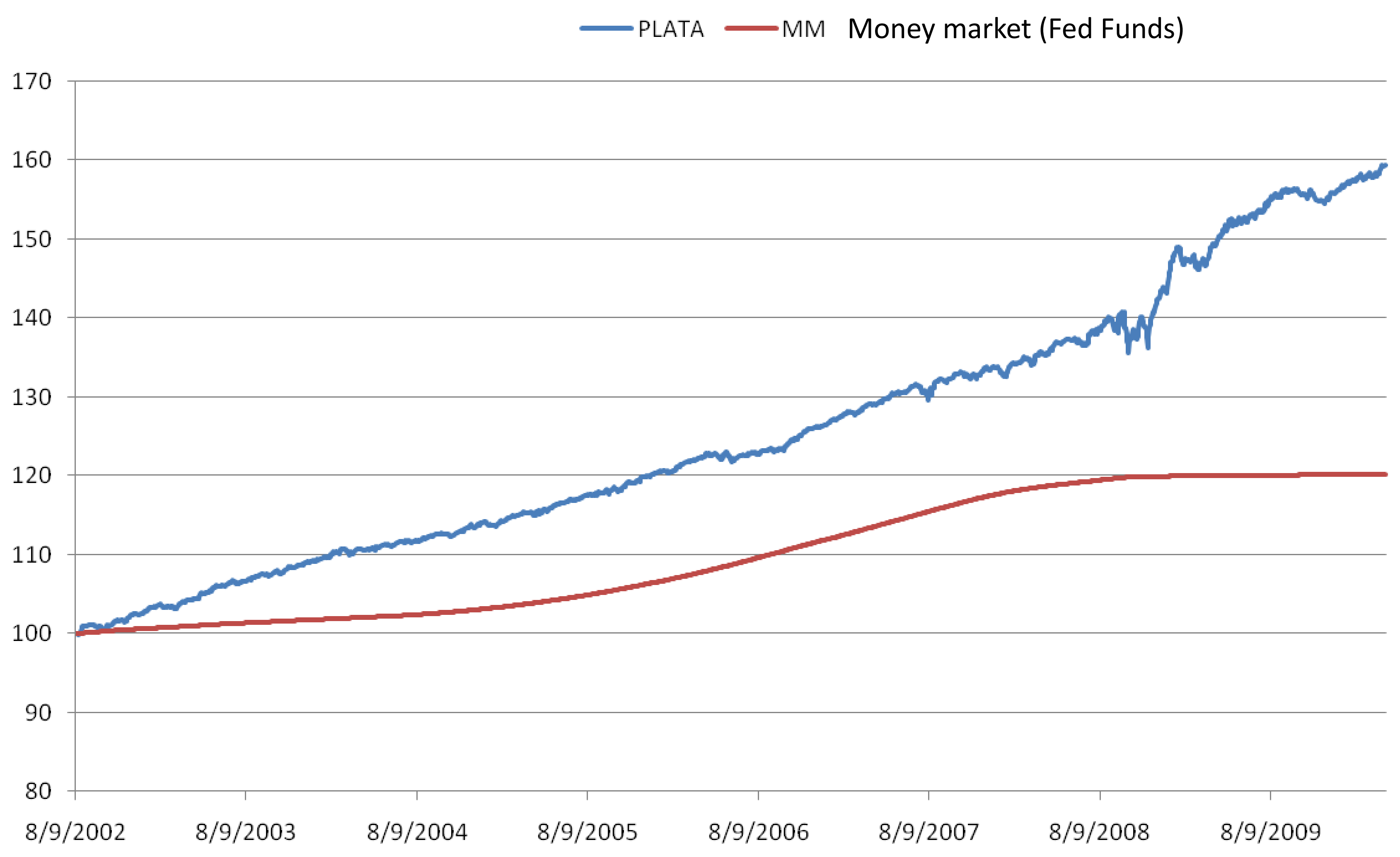
-- Monitor for degradation of statistical parameters, stop-losses, etc.

-- Investment per stock ~ 25 bps (~250K per 100MM notional capital)

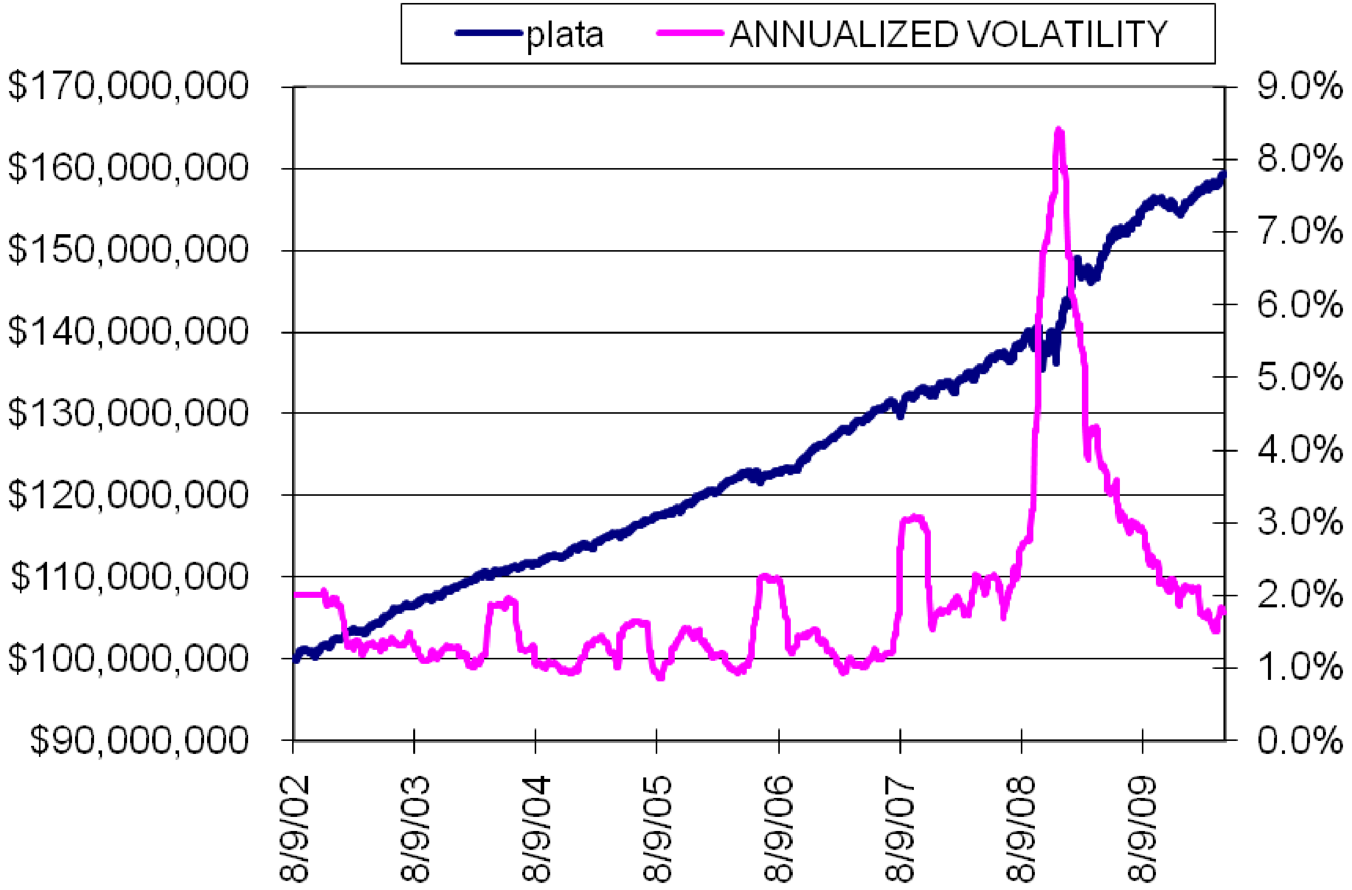
-- Typical profile 30 to 50 % long / 30 to 50 % short, dollar-neutral.

-- Portfolio-level risk management used to ``vet'' trades.

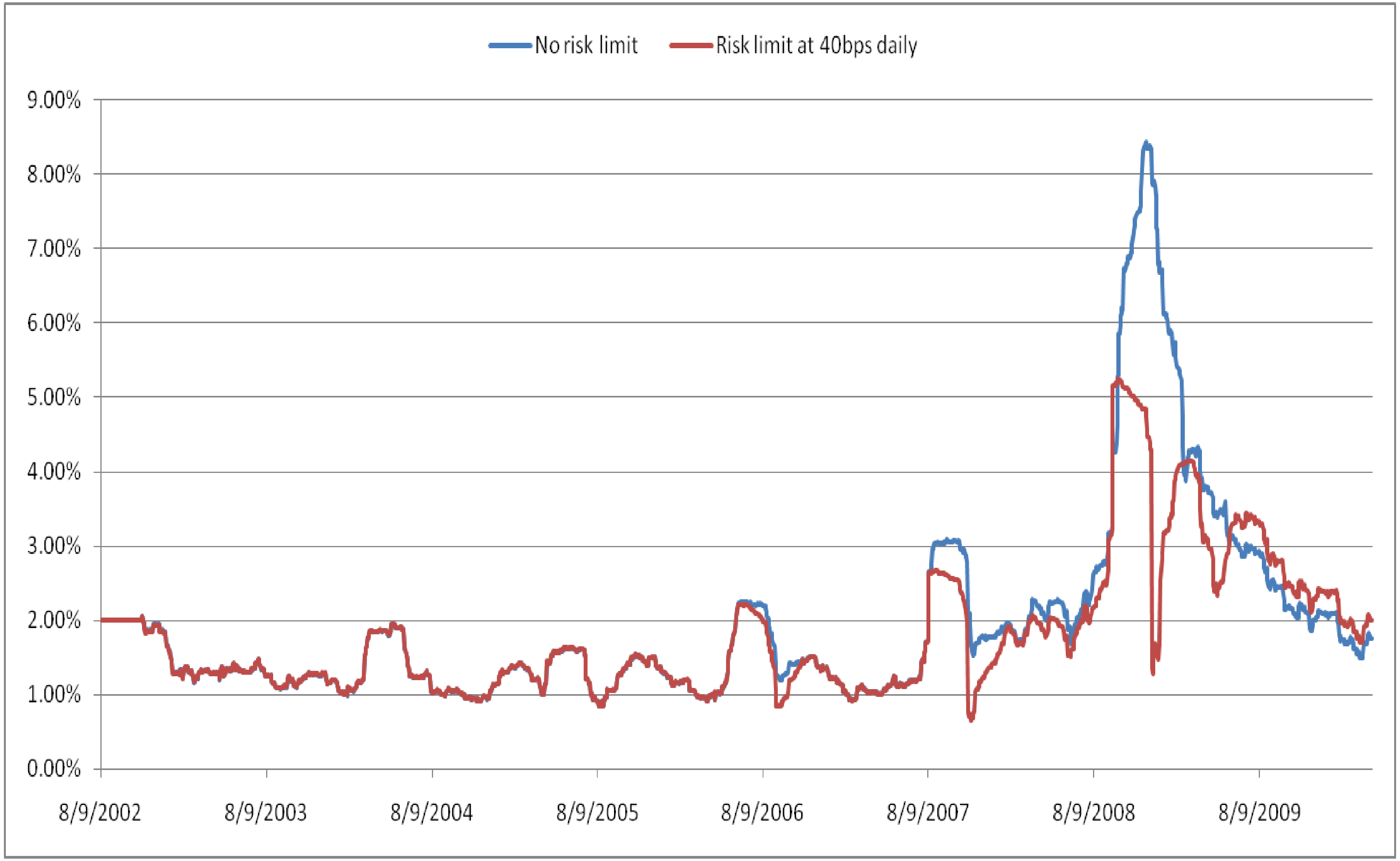
Basic PLATA : 0.25% per stock position
Leverage: 20/20 to 40/40, target daily vol=25 bps



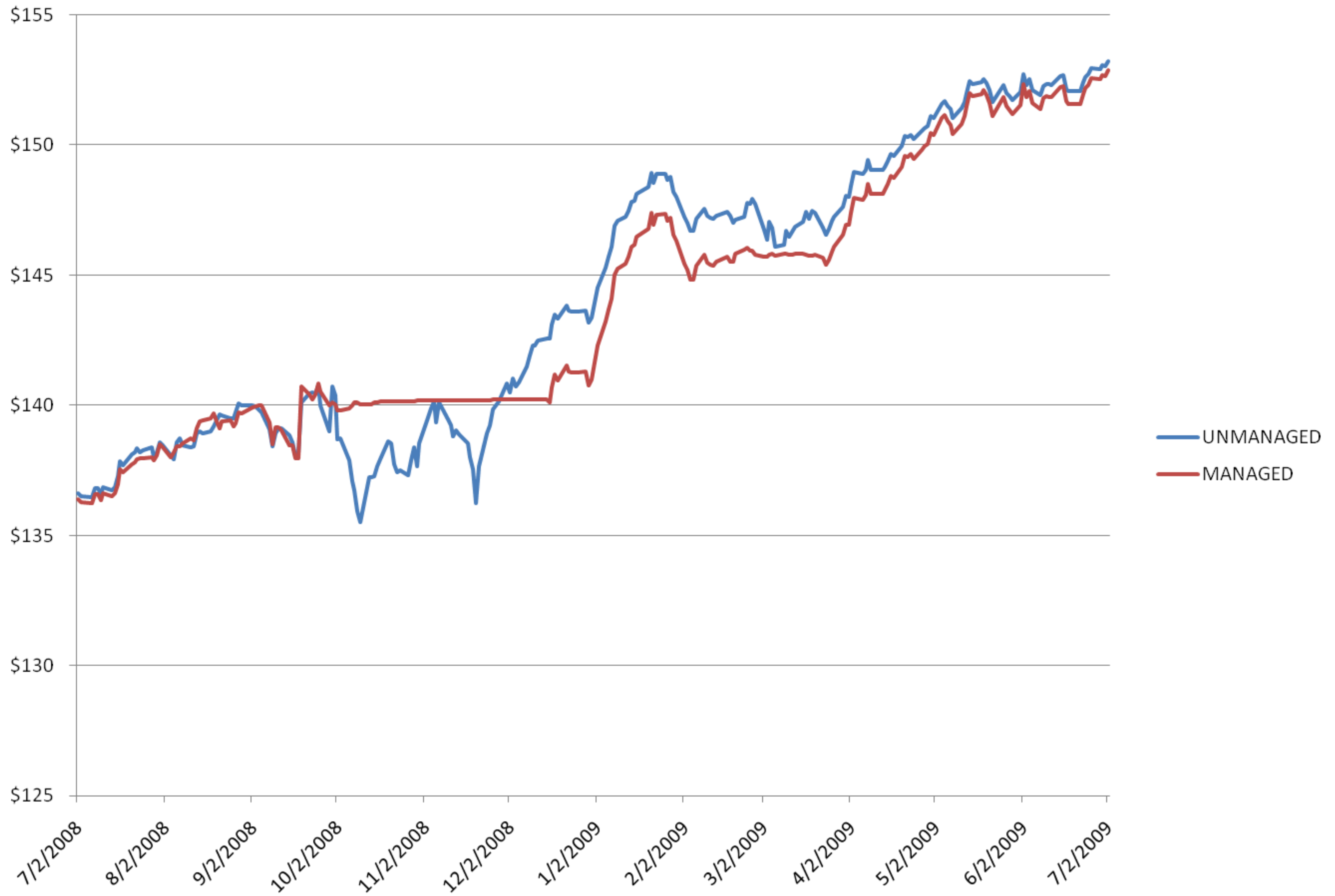
Trailing annualized volatility without portfolio risk management



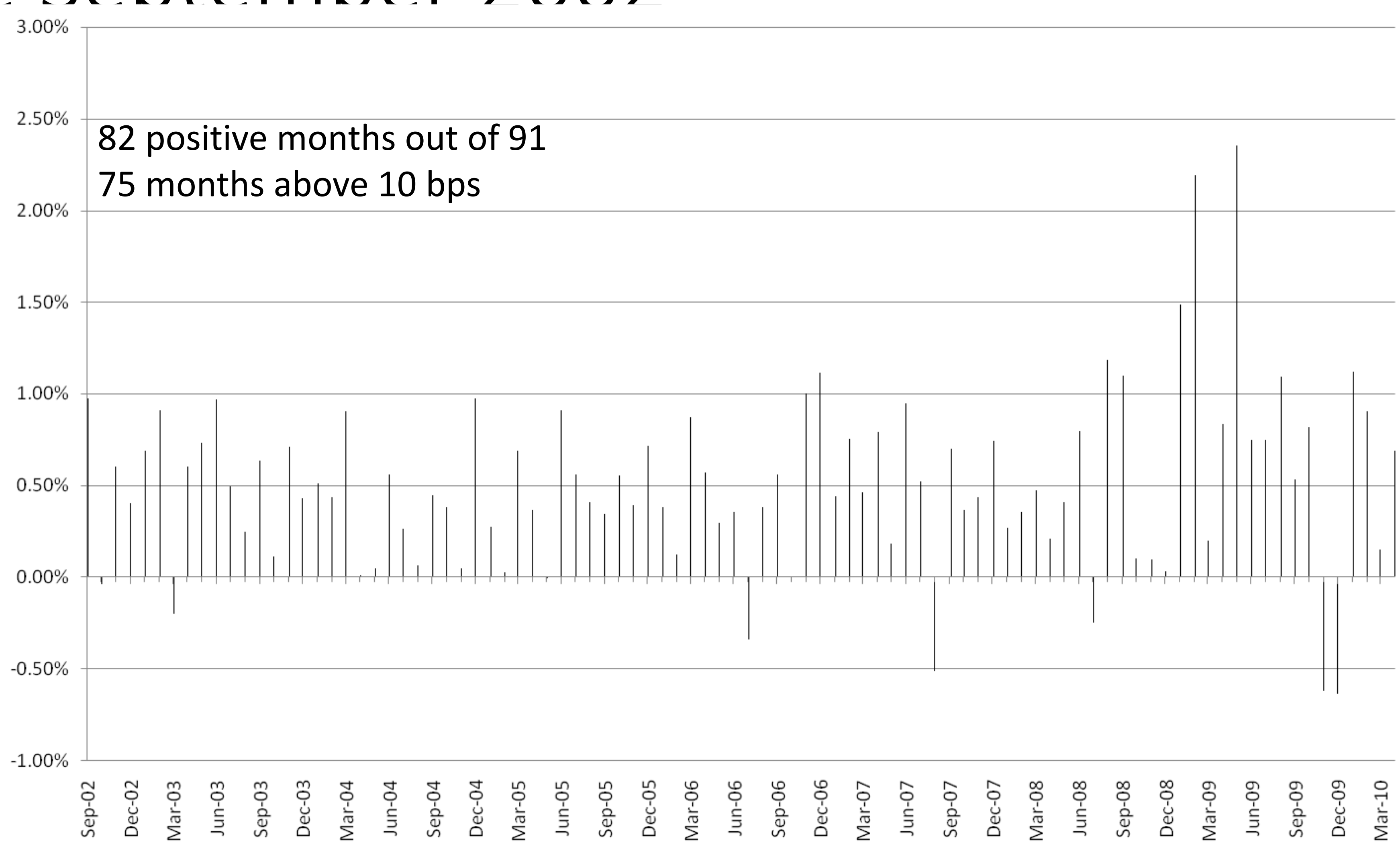
Volatility of PLATA with and without risk limit



Difference between managed risk and unmanaged risk in the Fall of 2008

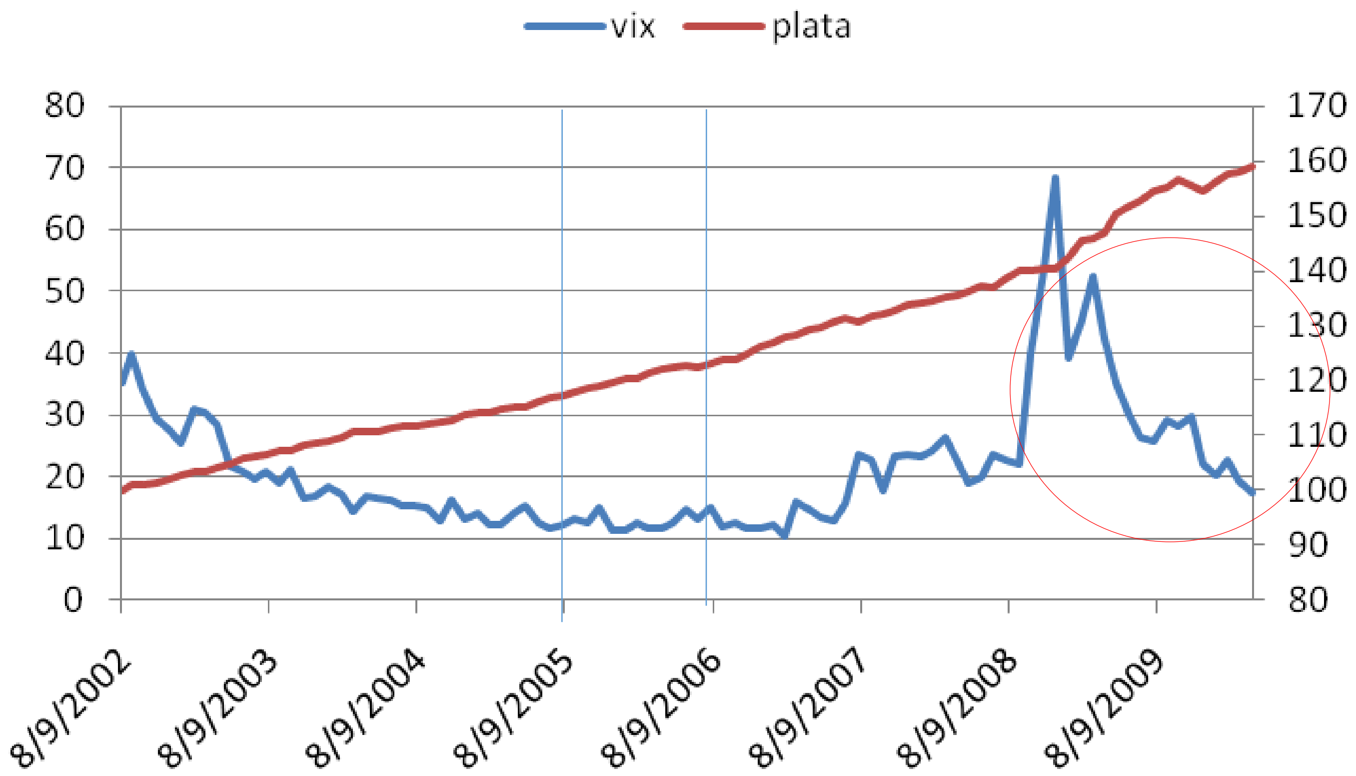


PLATA monthly performance since September 2002



PLATA Monthly Statistics	
Mean	0.50%
Standard Error	0.05%
Median	0.46%
Mode	NA
Standard Deviation	0.48%
Sample Variance	0.00%
Kurtosis	3.01
Skewness	0.75
Range	3%
Minimum	-1%
Maximum	2%
Sum	0.46
Count	91
Largest(5)	1.12%
Smallest(5)	-0.25%
Confidence Level(99.0%)	0.13%

Volatility and Stat Arb (Plata)



PLATA works better in the aftermath of volatility spikes and less well when volatility drops. It is therefore reasonable to blend it with an index strategy

Statistical Arbitrage and 130/30

Indexers: mutual fund managers and long-only managers

Objective: Track (or beat) returns of the overall market or sector

Underlying theory: CAPM, etc.

130/30 Managers: Long 130%, short 30% with periodic revisions of the portfolio

Essentially, a beta strategy with stock picking

Market-neutral managers: Seek returns that are uncorrelated with the market (alpha)

Long-Short Equity MN: Fundamental Stock Picking with shorting

Statistical Arbitrage: Quantitative long-short MN

SPY+PLATA: a synthetic 130/30 fund

Based on a notional amount of 100 MM:

-- go long 100 MM SPY and

-- enter into a PLATA strategy based on 100MM notional amount
(30 to 50 mm long/ 30 to 50 mm short)

(parameters for PLATA: big universe, 25bps per stock,
target daily stdev of portfolio=25bps)

Due to market-neutrality of PLATA, this portfolio looks essentially like
a 130/30 to a 150/50 depending on the volatility in the market and the
turnover.

-- Assume fee structure: ~ 1.20% per year (average for quant mutual funds)

Comparing SPY+PLATA with SPY

MainStay 130 30 Core C (MYCCX)

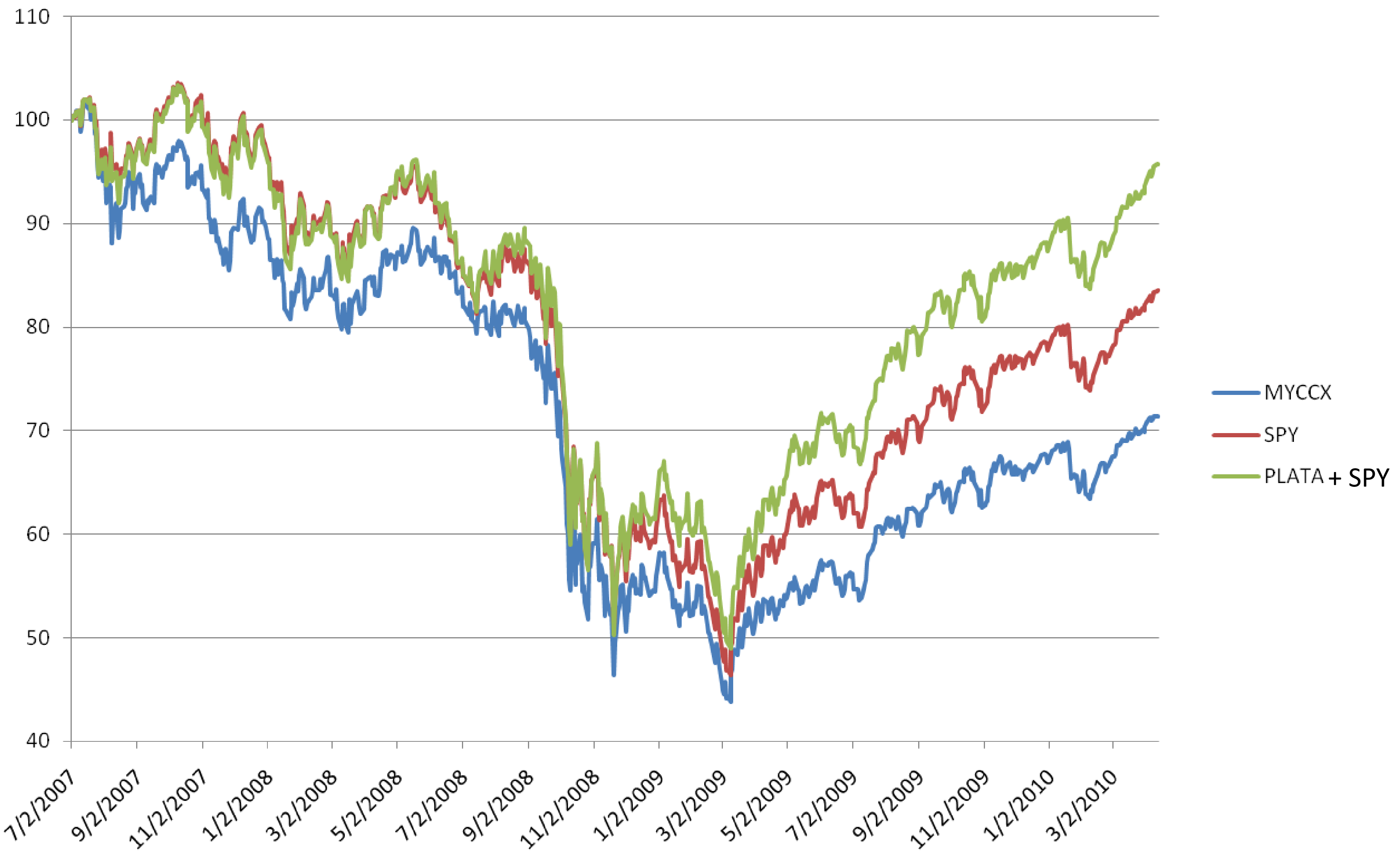
The investment seeks long-term growth of capital, with income as a secondary consideration. The fund primarily invests in common stocks of well-established U.S. companies, primarily those with large-capitalizations that are similar to companies in the Russell 1000 index. It generally holds long positions, either directly or through derivatives, equal to approximately 130%(within a range of 120% to 140%) of net assets, and short positions, either directly or through derivatives, equal to approximately 30% (within a range of 20% to 40%)of net assets.

Expense	MYCCX	Category Avg
Total Expense Ratio:	2.35%	1.12%
Max 12b1 Fee:	1.00%	N/A
Max Front End Sales Load:	N/A	5.28%
Max Deferred Sales Load:	1.00%	2.53%
3 Yr Expense Projection*:	1,027	596
5 Yr Expense Projection*:	1,741	894
10 Yr Expense Projection*:	3,631	1,787

AUM: 327 Million USD

Source: Yahoo!Finance

MainStay 130 30 Core C (MYCCX)



Fidelity Advisor Large Cap 130/30

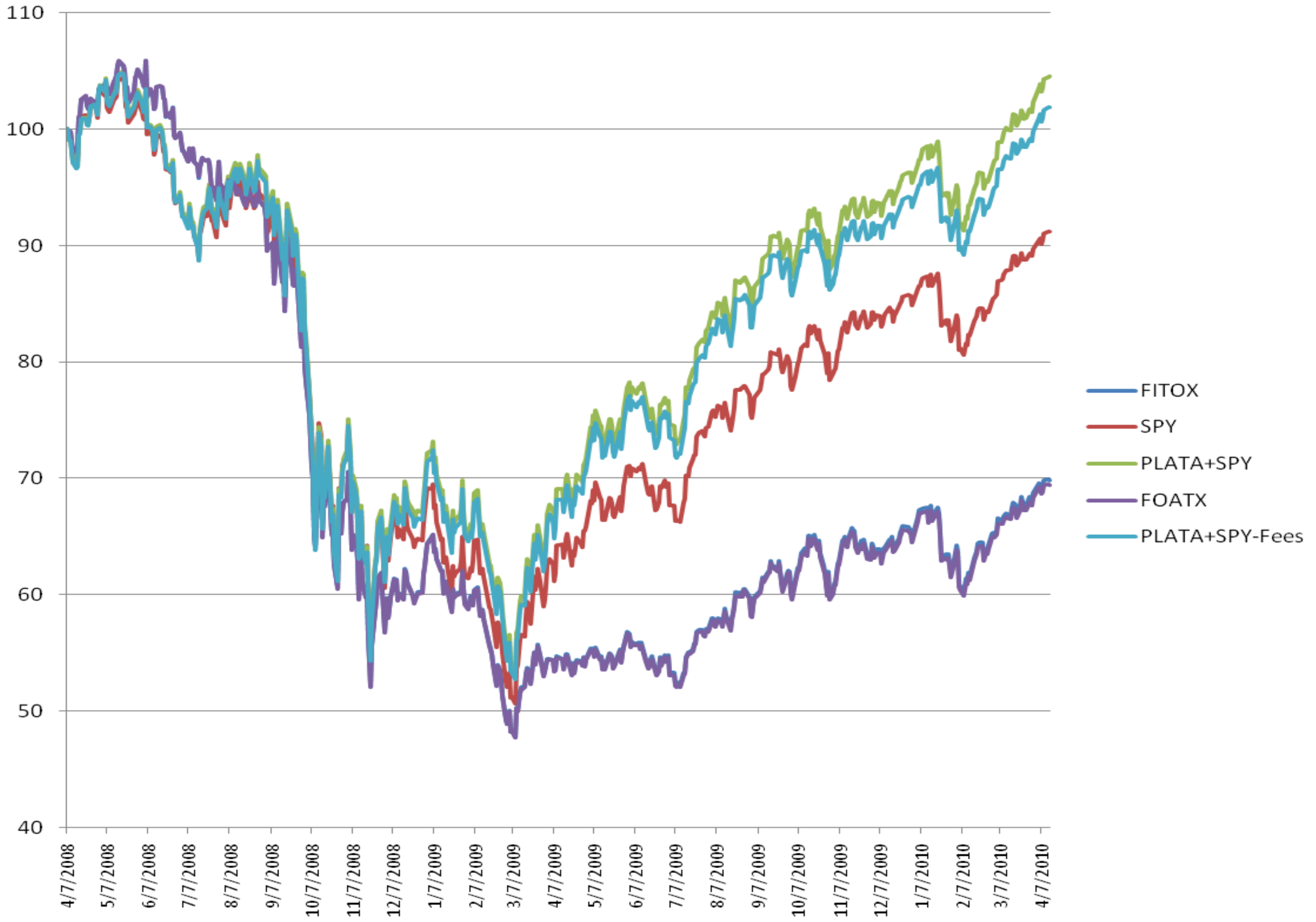
FITOX, FOATX

The investment seeks long-term growth of capital. The fund normally invests at least 80% of assets in common stocks of companies with large market capitalizations. It normally establishes long and short positions in equity securities. The fund invests either in "growth" stocks or "value" stocks or both. It may invest in domestic and foreign issuers.

Expense Ratio: 1.12 %., 1.56%
AUM: 26 MM,.

Source: Yahoo!Finance

Fidelity Advisor Large Cap 130/30 (FITOX, FOATX)



BNY Mellon US Core Equity 130/30

MUCIX

The investment seeks capital appreciation. The fund normally invests at least 80% of net assets in equity securities. It focuses on growth and value stocks of large cap companies. Although the fund typically invests in seasoned issuers, it may purchase securities of companies in initial public offerings (IPOs) or shortly thereafter.

AUM: USD 222 Million
Fee: 2%

Source: Yahoo!Finance

BNY Mellon US Core Equity 130/30 (MUCIX)

CSM: the first 130/30 ETF

-- Proshares launched the first 130/30 ETF in July 2009

-- Based on the 13030 Large Cap Index constructed by Andrew Lo (MIT) and Panjak Patel (Credit Suisse)

-- Based on ranking stocks in S&P 500 according to 10 quantitative criteria (Book to Value, Momentum, etc)

-- Monthly rebalancings

-- Data available does not include the credit crunch

The investment seeks investment results, before fees and expenses, that track the performance of the Credit Suisse 130/30 Large Cap Index. The fund invests in a combination of equity securities and derivatives that Proshares Advisors believes should track the performance of the Index. It invests typically the rest of the assets in money market instruments. The fund is non-diversified.

AUM: USD 60 Million
Fee: 0.95%

Source: Yahoo!Finance

Comparison of SPY+ PLATA with CSM since inception (7/2009)