Lecture 5
Factors, More Factors and More Factors
&
August 2007: 6 Days that Shook Our World

Matthew Rothman

MIT Sloan School of Management

Spring 2023

Law of Active Management

Fundamental Law of Active Management

- Information Ratio: The ratio of annualized residual return to residual risk. The central measurement of active management.
- Information Coefficient: The correlation of forecasted returns with their subsequent realizations. A measure of skill.
- Breadth: The number of independent forecasts available.

Information Ratio = Information Coefficient x Breadth^{½}

This is the Fundamental Law of Active Management

So how do you increase your IR?

1. Get better at predicting returns….
2. Increase your breadth …

   Note you can do this in terms of number assets and in “time space”

Outline

- Law of Active Management
- Factor Creation
- What You Need To Be Worrying About
- Quant Crisis 2007

Law of Active Management

Fundamental Law of Active Management

- Information Ratio: The ratio of annualized residual return to residual risk. The central measurement of active management.
- Information Coefficient: The correlation of forecasted returns with their subsequent realizations. A measure of skill.
- Breadth: The number of independent forecasts available.

Information Ratio = Information Coefficient x Breadth^{½}

This is the Fundamental Law of Active Management

So how do you increase your IR?

1. Get better at predicting returns….
2. Increase your breadth …

   Note you can do this in terms of number assets and in “time space”

Multivariate (Model) Factor Creation

How Do We Do It? Multivariate Factor Selection

- Multivariate (joint) Factor Selection
  - We seek Nirvana: two or more factors that each work on an individual basis but have small correlation with each other.
  - Combing low correlation signals increases returns and reduces return volatility
- Process:

1. Calculate factor score for each factor separately
2. Calculate combined score for each stock: sum of each factor score x factor weight
3. Rank stocks by combined factor score
4. Create long/short factor portfolios (each month)
   - Long the highest ranked quintile of stocks
   - Short the lowest ranked quintile of stocks
5. Calculate forward returns for the quintiles
6. Determine efficacy

Forecasting Returns – Quantitative Research

Performance Statistics for Univariate and Multivariate Examples

	Book/ Price (B/P)	Momentum (MOM)	ROIC	B/P + MOM	**B/P + MOM + ROIC**
Return	3.5%	6.1%	2.5%	7.5%	**9.2%**
Volatility	12.2%	15.6%	8.4%	11.1%	**9.9%**
Information Ratio	0.29	0.39	0.30	0.68	**0.92**

– Combining independent signals enhances performance

- Higher expected returns
- Lower return volatility
- Higher information ratios

How Do We Do It? Multivariate Factor Selection

How Do You Test a Multivariate Factor Model?

1. Do all the same tests you did for a univariate factor.
2. When you run cross-sectional regressions, watch out for multi-collinearity!
   - My favorite (and should be easiest) interview question is:

     "When you are running a regression and your two independent variables, X1 and X2 are correlated, how does this impact the coefficients and t-stats?"

   - Why do I ask this? This is multivariate Fama-McBeth regressions when your factors are not independent. Happens all the time. What should you do?
   - Econometrics 101 and most people don't know it. It is shameful!
3. When does your model fail?
4. What are the biases in your model performance? Break it down and understand it!

We will talk more about this next week. Again. This the difference between being an investor and computer geek betting in the market.

Multivariate Factor Testing

Things to consider in testing the efficacy of a factor in multivariate setting

1. What is the correlation of signals?
2. What is the correlation of your factor returns?
3. If your signals are highly correlated but there is incremental return should you have a ‘composite’ factor?
4. What is the “Jensen’s Alpha” of your factors? Test it both ways!

$$R_{Y,t} = \alpha + R_{X,t} + e_t$$

5. What is the incremental R^2 of your model?
6. Is your factor diversifying your positioning or is exacerbating your biases?
7. Are you getting “exposure” to your new alpha source? Is the exposure appropriate?

Univariate Factor Testing

Things to consider in testing the efficacy of a univariate factor

1. Do you want to have any out-of-sample period?
2. What is the absolute return that the factor is able to produce?
3. What the Sharpe Ratio, and the Information Ratio (IR) of the factor? What is its Alpha-to-Margin ratio? Sortino Ratio?
4. What is its risk adjusted return? Is it long (or short) market beta? Is it loading up on factor exposures? Does it have it systematic sector or country exposures?
5. What are its hit-rates, both for stock selection and in terms of its performance rate?
6. What is turnover? The autocorrelation of the signal itself? What are the transaction costs associated with trading it?
7. What are its drawdowns? Given its Sharpe Ratio and IR what should you expect its drawdown and underperformance to look like? (see pp. 33-34 of Pedersen!)
8. What does it payoff profile look like? How quickly does its payoff efficacy decay?
9. How robust is the signal's performance to minor variation in its definition?
10. Are there certain universes (large cap, small cap, EM, DM, U.S., Japan) that it performs better or worse? Certain sectors? Do these make sense to you?
11. What is the time series behavior of the factor? Does the time series behavior align with macro regimes? Market regimes?

Open Questions in Factor Construction

What do you need to consider in creating your factors? What do you want to make sure you get right?

1. Make sure your sample is free of survivorship bias. How might you test this?
2. Is your dataset backfilled? How might you look for this?
3. Is your dataset point-in-time? Does it have the revision history?
4. Make sure your identifier “mapping table” is functioning and free of survivorship issues and consistent across vendors
5. Make sure you are only using information known at time t-1 (time lag is appropriate)
6. Make sure your returns are from time t to $t+\tau$
7. Make sure your LHS is demeaned (by what?).
8. Make darn sure you UNDERSTAND your data. Plot it!
9. What universe do you want to use? And how do you want “standardize” your variable (industry neutral, sector neutral, country neutral)? How does this compare to your answer in (7)?
10. How do you want to handle outliers? Do you “normalize” your variables? Is there information, or noise, in the shape of the distribution and outliers?
11. Do you understand your data?!? Did you look at it??

Questions to Ask Yourself As YouTest A Factor

How Do You Create and Test a Factor?

The entire biggest message of this class:

There are no simple individual tests or procedures to do that will give you the answer.

You Need to be a Critical Consumer of Quantitative Research. No matter who is the producer: yourself, Gene Fama, Andy Lo, the latest Journal of Finance article, or the Managing Director on your Desk.

Does it make sense? What are the biases in their result? What didn't you or they understand? What data and real world issues are biasing their (your) results?

If you don't understand this, if you don't get it right, you will blow up.

Or at best look stupid.

Univariate Factor Testing

Things to consider in testing the efficacy of a univariate factor

Remember Fama-French take the Book-to-Price Ratio of every stock and rank them according to the value.

Do they demean by sector?

Do they z-score?

- Today the S&P 500 average P/B ratio is 3.34 and P/E ratio is 19.98

	P/B	P/E
Technology	3.59	10.29
Financials	1.72	13.05
Transports	2.51	13.45
Energy	1.61	14.39
Basic Materials	1.76	17.72
Capital Goods	4.19	18.45
Utilities	1.88	24.55
Consumer Cyclical	3.31	24.80
Services	10.13	25.71
Health Care	4.43	27.59
Consumer Non-Cyclical	5.10	29.86
Retail	8.09	34.09

Univariate Factor Testing

	P/B	P/E
Energy		
Coal	0.78	7.02
Well Services & Equip	1.31	8.32
Integrateds	1.57	12.44
Oil & Gas Production	1.98	19.59
Renewable	5.22	41.97
Health Care		
Health Care Facilities	2.48	15.33
Laboratories	2.58	16.71
Equipment & Supplies	3.78	17.42
Biotech	3.65	22.47
Major Pharma	5.28	29.18
Instruments	5.59	69.39
Technology		
Peripherals	0.83	6.10
Semis	1.45	10.85
Software	5.62	18.32
Networks	5.88	19.02
Social Media	6.31	34.24
Hardware	7.72	47.26

How Do You Test a Factor? Big Picture

Every researcher better know:

- What causes your factor to produce an erroneous signal?
- When does your model underperform?

Examples:

- When does Book / Price fail? What might cause a stock to look very cheap when it really isn't?
- When might your Beta estimates be really inaccurate?
- When might Momentum fail?
- What can your NLP algorithm not handle? When will it be get something wrong? Remember Anne Hathaway vs. Berkshire Hathaway

If you don't understand this backwards and forwards, you should not be using that investment strategy. At best, you are just gambling. You are being reckless.

How Do You Test a Factor? Let's Get Specific

What Else Should You Look At?

1. Hit Rates: Separating Big Returns from Consistency
   - How often do your quintiles outperform?
   - What percent of stocks in your quintiles outperform?
2. Look for monotonicity across all your quintiles
   - Does Q1 outperform Q2 which outperforms Q3 which outperforms Q4 which outperforms Q5? Or are you U or L or "reverse" J shaped?
   - Are you monotonic in hit-rates too?
3. Does your factor perform better in certain market environments? Or macro environments?
   - High Vol vs Low Vol? High inflation vs low inflation? High interest rate regimes vs low interest regimes?
   - How time-sensitive is the performance? Did it do very poorly or well in recessions? In the Financial Crisis? Or the rebound from the crisis?
4. Does it do better (or worse) in certain sectors / industries or in certain countries (if a global model)? Do you want to bet CEF, BDC, SPACs, biopharma and medical devices? Do you want exposure to Mortgage REITS today?

How Do You Test a Factor? Let's Get Specific

What Else Should You Look At?

1. Drawdowns
   - How much of a loss can you withstand before you capital is pulled?
   - How long can you stand it?
2. Win Loss Ratio for a Day and Per Stock
   - How much do you make on winning days and much do you lose on losing days?
   - How much do you win on your average winning trade and how much do you lose or you average losing days
3. If you remove the +/- 4 Standard Deviation stocks from your portfolio how much does your Sharpe Change
4. If I remove the top 1% of winning days from your PnL stream, how much does your Sharpe change?
5. How many stocks make up 50% of your $PNL?
6. What are the characteristics of stocks where you make money and lose money? High Short Interest names? High idiosyncratic vol names? High Mo vs low MO names?

How Do You Test a Factor? Let's Get Specific

What Else Should You Look At?

7. What is the skew of your strategy? Are you picking up pennies in front of a steam roller? Look at your average return on up days / average returns on down days

8. How much of your returns comes from the Long Side and how much from the short side?

9. What is the natural turnover of your factor?

10. If you lagged your factor by a day (minute, 5 minutes, hour) how much of your return do you lose?

11. How much of your factor return comes from intra-day returns and how much comes overnight returns?

12. Have you accounted for t-costs? How much is your slippage from arrival?

13. Have you accounted for financing costs? Commissions? Dividend withholding rates?

More Factor Testing Questions

Things to consider in testing the efficacy of a univariate factor

1. Do you want to have any out-of-sample period?
2. What is the absolute return that the factor is able to produce?
3. What the Sharpe Ratio, and the Information Ratio (IR) of the factor? What is its Alpha-to-Margin ratio? Sortino Ratio?
4. What is its risk adjusted return? Is it long (or short) market beta? Is it loading up on factor exposures? Does it have it systematic sector or country exposures?
5. What are its hit-rates, both for stock selection and in terms of its performance rate?
6. What is turnover? The autocorrelation of the signal itself? What are the transaction costs associated with trading it?
7. What are its drawdowns? Given its Sharpe Ratio and IR what should you expect its drawdown and underperformance to look like? (see pp. 33-34 of Pedersen!)
8. What does it payoff profile look like? How quickly does its payoff efficacy decay?
9. How robust is the signal's performance to minor variation in its definition?
10. Are there certain universes (large cap, small cap, EM, DM, U.S., Japan) that it performs better or worse? Certain sectors? Do these make sense to you?
11. What is the time series behavior of the factor? Does the time series behavior align with macro regimes? Market regimes?

Open Questions in Factor Construction

What do you need to consider in creating your factors? What do you want to make sure you get right?

1. Make sure your sample is free of survivorship bias. How might you test this?
2. Is your dataset backfilled? How might you look for this?
3. Is your dataset point-in-time? Does it have the revision history?
4. Make sure your identifier “mapping table” is functioning and free of survivorship issues and consistent across vendors. Know the difference between company id’s and security id’s
5. Make sure you are only using information known at time t-1 (time lag is appropriate)
6. Make sure your returns are from time t to $t+\tau$
7. Make sure your LHS is demeaned (by what?).
8. Make darn sure you UNDERSTAND your data. Did you even look at it and plot it?!?
9. What universe do you want to use? And how do you want “standardize” your variable (industry neutral, sector neutral, country neutral)? How does this compare to your answer in (7)?
10. How do you want to handle outliers? Do you “normalize” your variables? Is there information, or noise, in the shape of the distribution and outliers?

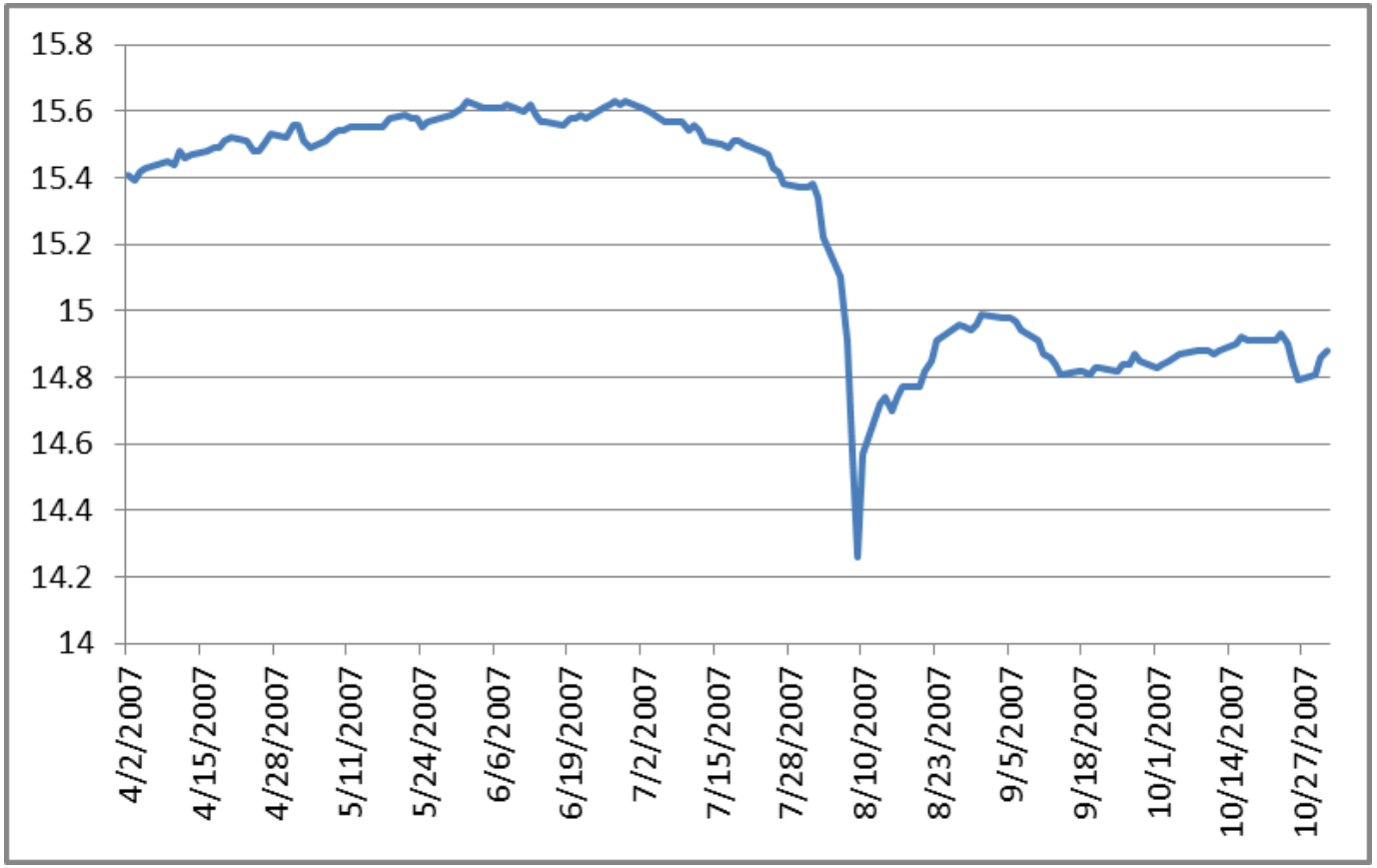
Quant Meltdown

Outline

- Quant Meltdown: Anatomy of The Crisis
- How Correlated are the Quants?

Fund Returns

HSKAX



Highbridge Statistical Market Neutral mutual fund (Quant Equity Russell 1000 fund)

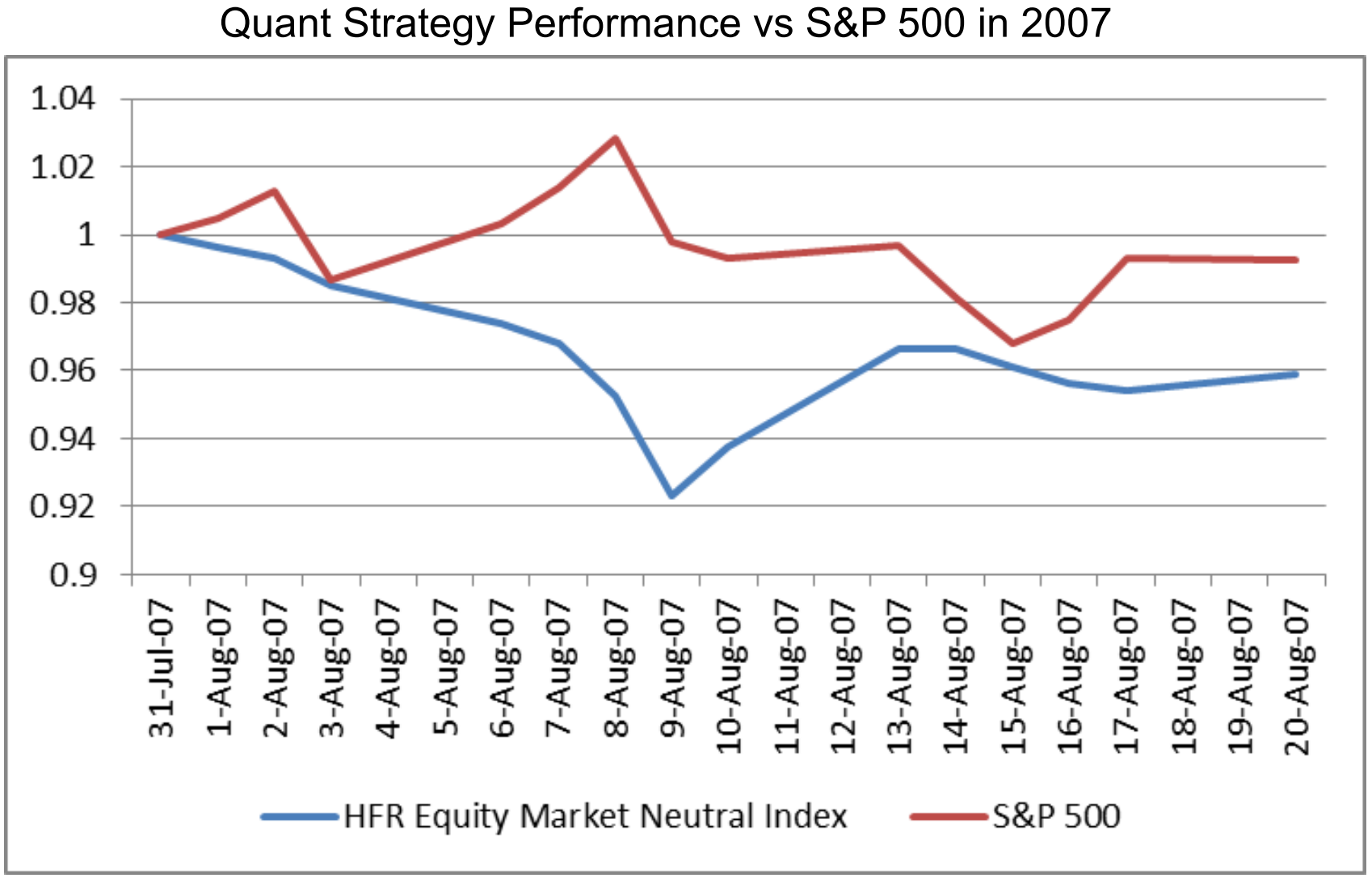
Fund Returns

HSKAX Performance
(mean=0.025%, stdev =0.0016045)*

Date	Price	Return	Z-score
8/1/2007	15.38	0.07%	0.25
8/2/2007	15.34	-0.26%	-1.78
8/3/2007	15.22	-0.78%	-5.03
8/6/2007	15.1	-0.79%	-5.07
8/7/2007	14.91	-1.26%	-8.00
8/8/2007	14.57	-2.28%	-14.37
8/9/2007	14.26	-2.13%	-13.42
8/10/2007	14.57	2.17%	13.40
8/13/2007	14.72	1.03%	6.26
8/14/2007	14.74	0.14%	0.69
8/15/2007	14.7	-0.27%	-1.85
8/16/2007	14.74	0.27%	1.54
8/17/2007	14.77	0.20%	1.11
8/20/2007	14.77	0.00%	-0.16

*calculated 12/5/2005 – 6/30/2007

Fund Returns



Longer Term Perspective

Cumulative Performance of Lehman Model Portfolio (January 2006 – September 30, 2007)

Source: Lehman Brothers Quantitative Equity Strategies; CRSP; Compustat; Russell; Thompson Financial Services
Past results are no guarantee of future performance.

Longer Term Perspective

Company Name / Hedge Fund	Early August 2007 Results	August 2007 Monthly Results
Alpha Equity Leveraged Market Neutral Fund		-6.4%
AQR	-13.0%	
GMN Capital	-19.0%	
ACI Market Neutral Fund	-19.0%	
Man Group PLC AHL Fund	-7.0%	
Ascend US Market Fund		-19.2%
Ascend US Market Fund Mid-Large Cap Fund		-12.7%
Ayrie Goden Eagle Fund		-16.2%
Barclays Global Investors 32 Capital Fund	-8.0%	0.2%
BattenKill Market Neutral Fund		-2.3%
Black Mesa Capital	-7.5%	-10.7%
DE Shaw Composite Fund	-15.0%	
Goldman Sachs -- Global Alpha Fund	-28.2%	-22.5%
Goldman Sachs -- Global Equity Opportunities Fund	-30.0%	
Hamton US Hedge 2 Limited		-4.1%
JP Morgan -- HighBridge Capital Fund	-18.0%	
Jasper Partners LP		-1.7%
Morgan Stanley Process Driven Trading (PDT)	-$500 mil*	
Renaissance Institutional Equity Fund (REIF)	-9.0%	0.4%
Renaissance Medallion Fund		3.9%
Temujin Fund Management	-4.0%	-0.4%
Tewksbury Investment Fund		-8.0%
Thales Fund Management		-3.7%
Tykhe Capital	-31.0%	-20.0%
Whitebox Statistical Arbitrage Fund		-4.0%

Longer Term Perspective

Factor Return and Historical Context for the Returns (January 1973 – July 2007)

Factor	August 3 - August 9	% of Observed Times Factors Returns More Extreme
EBITDA to EV	-5.9%	0.0%
Book to Price	-2.2%	2.3%
Gross Free Cash Flow to Price	-5.0%	0.1%
Sales to Price	-3.5%	0.8%
Trailing Earnings to Price	-4.5%	0.1%
ROIC	-1.2%	4.1%
Change in Debt to Assets	-1.8%	1.4%
Sales Growth	1.6%	4.5%
Change in Shares Outstanding	-0.9%	8.5%
Asset Turnover	-3.1%	0.8%
Earnings Revision Ratio	-1.4%	3.7%
Price Momentum	-4.5%	2.3%

Source: Lehman Brothers Quantitative Equity Strategies; CRSP; Compustat; Russell; Thompson Financial Services
Past results are no guarantee of future performance.

Causes of the August 2007 Quant Crisis

- Crisis in the sub-prime market – contagion
  - Few large multi-strategy hedge funds sustained large losses in fixed income portfolios
    - Reluctance to realize losses / reduce risk by selling (or marking to market) fixed income portfolios
    - Raise equity (or de-lever) by quickly selling most liquid assets in the world – U.S. equities
  - Majority of these multi-strategy hedge funds were quant funds
    - Thus, positions unwound were series of quantitative factor based portfolios
    - Dramatic de-levering spurred other quantitative managers' portfolios to suffer losses
    - Lack of transparency for losses induced panic among managers

      → "Run on the Bank" mentality
  - Panic averted, capital deployed, short covering cause market to rebound.
    - Was it the "Goldman injection"? Transparency into the crisis? Action by the Fed?

Model and Theme Portfolios: Monthly Returns

Monthly Long / Short Model and Theme Portfolio Returns

	Overall Model	Valuation	Quality	Market Dynamics
December, 2006	2.7%	2.0%	0.5%	2.0%
January, 2007	-0.4%	0.2%	-1.1%	0.2%
February, 2007	-0.5%	0.2%	-0.9%	0.4%
March, 2007	0.1%	0.4%	-1.9%	2.7%
April, 2007	-1.0%	0.3%	-0.3%	-1.4%
May, 2007	-1.0%	0.4%	-0.9%	-1.3%
June, 2007	-0.2%	-1.2%	-0.7%	0.3%
July, 2007	-1.1%	-2.8%	-2.8%	2.1%
August, 2007	0.2%	-1.3%	-0.5%	1.1%

Source: Lehman Brothers Quantitative Equity Strategies; CRSP; Compustat; Russell; Thompson Financial Services
Past results are no guarantee of future performance.

How Alike Are Quants?
Or, Is Quant A Crowded Trade?

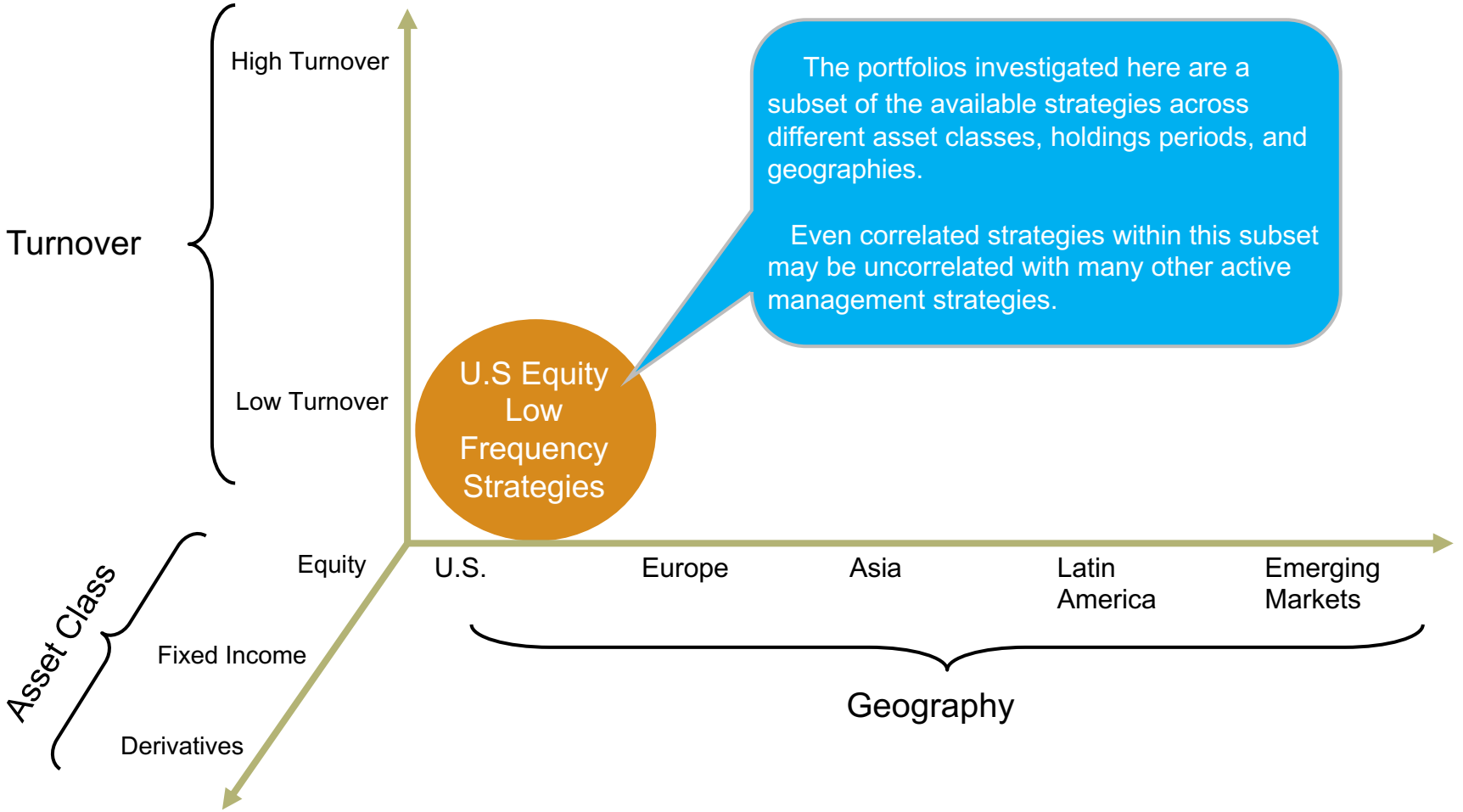
How Alike Are Quants?

- 10 out of the 15 largest quantitative managers agreed to give us the rankings coming out of their Core models for 2006 Q1, 2006 Q2 and 2006 Q3
  - Top quintile ranked stocks in S&P 500; Bottom quintile ranked stocks in S&P 500
  - Top quintile ranked stocks in R2000; Bottom quintile ranked stocks in R2000

- We picked a small sample – want homogeneity. These are pure quants. True quants

- Focus solely on top and bottom quintiles in order to maximize agreement, plus this is where the models take their largest positions

How Alike Are Quants?



Agreement and Disagreement Across Funds

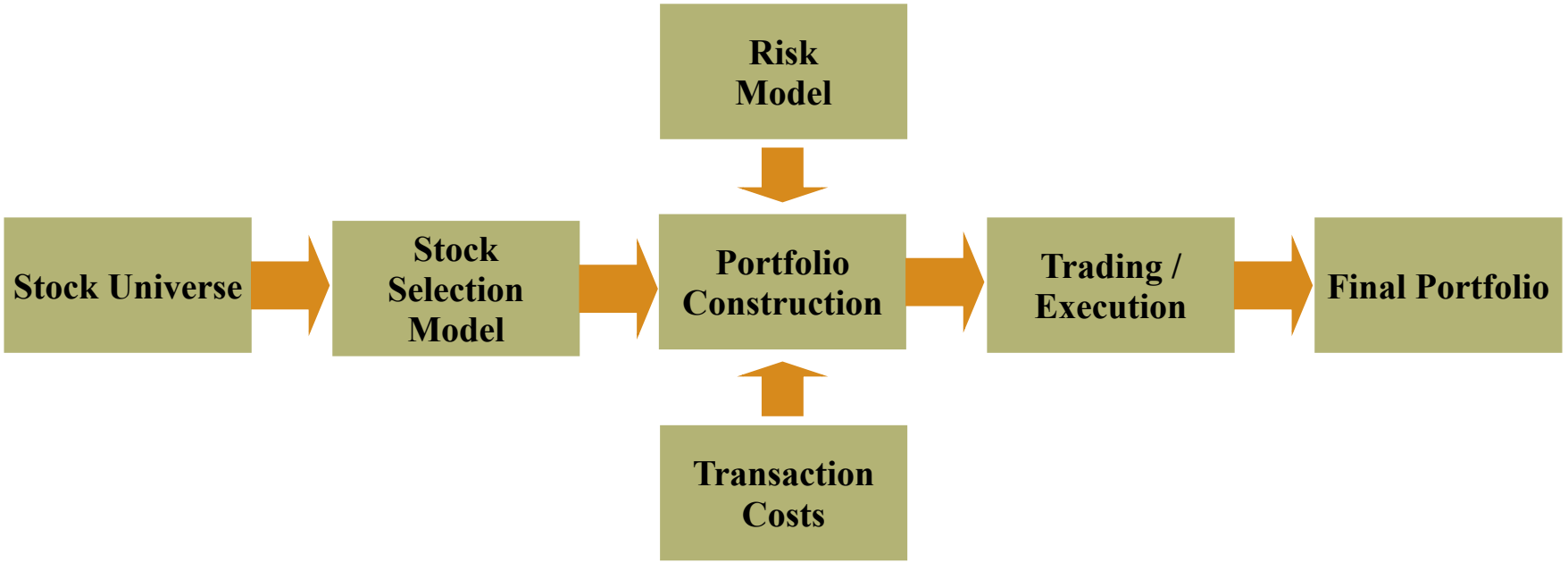
March 2006: Average Percent of Fund's Holdings in Another Fund's Portfolio

Fund	Agreement		Disagreement		Net Agreement
	Longs	Shorts	Longs	Shorts	
A	34%	35%	8%	10%	26%
B	35%	34%	10%	7%	27%
C	35%	43%	8%	8%	31%
D	37%	45%	8%	9%	32%
E	36%	36%	7%	7%	29%
F	43%	38%	7%	5%	34%
G	34%	36%	12%	7%	25%
H	44%	40%	5%	5%	37%
I	37%	37%	7%	6%	30%
J	40%	39%	8%	6%	33%
Barclays Capital Model Portfolio	41%	34%	8%	6%	30%
Fundamental Revisions Portfolio	29%	23%	17%	14%	10%
Fundamental Analyst Recommendations	16%	26%	18%	18%	3%
Average Across Quant Funds	**38%**	**38%**	**8%**	**7%**	**30%**

Source: Barclays Capital Quantitative Equity Strategies

- 20% reflects independence for agreement and disagreement.
- Net agreement is the average agreement minus the average disagreement; 0% reflects independence.

The Quantitative Investment Process



- 1,000's of stocks
- Create a Security Master Mapping Table (SEC MASTER) that maps across vendors and controls for corporate actions
- Liquidity test

- Rank stocks based on alpha or price forecast

- Model portfolio subject to risk limits and transaction costs
- Risks include sector mismatch, style, individual position size
- Limit execution costs

- Manual execution
- Advanced execution algorithms
- Utilize dealer "Pipes"
- Feedback execution to earlier processes
- Transaction cost analysis

Correlation of Funds' Daily Active Returns

Average Quant Long Portfolio Active Return Correlations: April – Dec. 2006

Fund	Equal-Weight Long Portfolios	Optimized Long Portfolios	Equal-Weight Long/Short Portfolios
A	38%	49%	39%
B	9%	45%	26%
C	41%	49%	48%
D	33%	34%	28%
E	40%	49%	53%
F	50%	49%	53%
G	23%	47%	28%
H	49%	55%	53%
I	40%	48%	51%
J	48%	53%	54%
Barclays Capital Model Portfolio	46%	55%	48%
Average	**38%**	**48%**	**44%**

Past results are no guarantee of future performance.

Source: Barclays Capital Quantitative Equity Strategies

Correlation of Funds' Daily Active Returns

Average Optimized Long Portfolio Active Return Correlations With Different Average Daily Volume (ADV) Constraints: April – Dec. 2006

Fund	15% ADV Constraint	30% ADV Constraint	50% ADV Constraint
A	66%	47%	45%
B	69%	36%	37%
C	67%	51%	45%
D	67%	32%	28%
E	72%	51%	48%
F	72%	59%	54%
G	59%	46%	45%
H	73%	58%	57%
I	74%	54%	49%
J	72%	57%	53%
Barclays Capital Model Portfolio	70%	57%	54%
Average	**67%**	**49%**	**46%**

Past results are no guarantee of future performance.

Source: Barclays Capital Quantitative Equity Strategies

Myths Emerging From the Crises

- Myth #1 – Crises was caused because all quantitative fund managers are running highly correlated strategies
- Myth #2 – All quant managers are the same – all look at the same data, all went to the same schools, all go to the same conferences and read the same papers. What do you expect?
- Myth #3 – Crises was caused because quantitative risk models are poorly designed
- Myth #4 – The money being run by quantitative fund managers has clearly reached its capacity. Don't allocate any more to these folks
- Myth #5 – Don't give money to a "genius" because he/she will inevitably "blow you up"

Summary

Summary

- In a crisis of impacting now illiquid assets, why might we expect to see highly liquid assets' value impacted?
- What is the mechanism by which liquidity crisis might impact the highly liquid assets' value?
- What was the role of leverage in the Quant 2007 crisis? If none of the funds were leveraged would there have been a crisis?
- Why did the crisis subside on August 10th ? What does the large snap-back on August 10th tell you about the crisis, broadly speaking?
- Was the fact that most quants use valuation signals a major source of the crisis?
- What was the role of portfolio construction in the quant crisis? Discuss both a risk and liquidity perspective.
- How unique were the signals that most managers were using? How unique were their returns streams? What is your evidence for both? And clearly explain the difference here.
- If you were running a quant fund in August 2007, what actions were you likely to have taken? What actions should you have taken? How might "agency" problems impacted what steps you would have taken? Would these have been different if you were running a long-short fund, a levered long-short fund, or a long only fund?
- How important in the aftermath of the August 2007 should it be for quantitative managers to look for proprietary signals ("unique alpha")? What might be some of the risks to doing this? Answer this question from both a business perspective (Chief Marketing Officer and Chief Executive Officer) (e.g. how your clients will react) and from the perspective of being the Chief Investment Officer.