FDP INSTITUTE®

Financial Data Professional Institute

FDP Institute provides world class training and education to financial professionals to meet the accelerating needs of digital transformation in the industry.

www.FDPInstitute.org

Practical Machine Learning in Asset Management: Manager Selection and Valuing the Secondary Sale of Private Assets

Lydia Ofori, CFA, CAIA
Founder & CEO
Hunter Labs Tech

Adam Duncan
Head of Investment Science Unit,
Cambridge Associates

Keith Black PhD, CFA, CAIA, FDP
Managing Director, FDP Charter
FDP Institute

Data is abundant, information is a commodity, pattern recognition is a scarce investment skill

There must be a better way to identify dislocation moments!

Human + Machine selecting the best opportunities

Qualitative

Help me to get to the point of the matter quickly

+

Quantitative

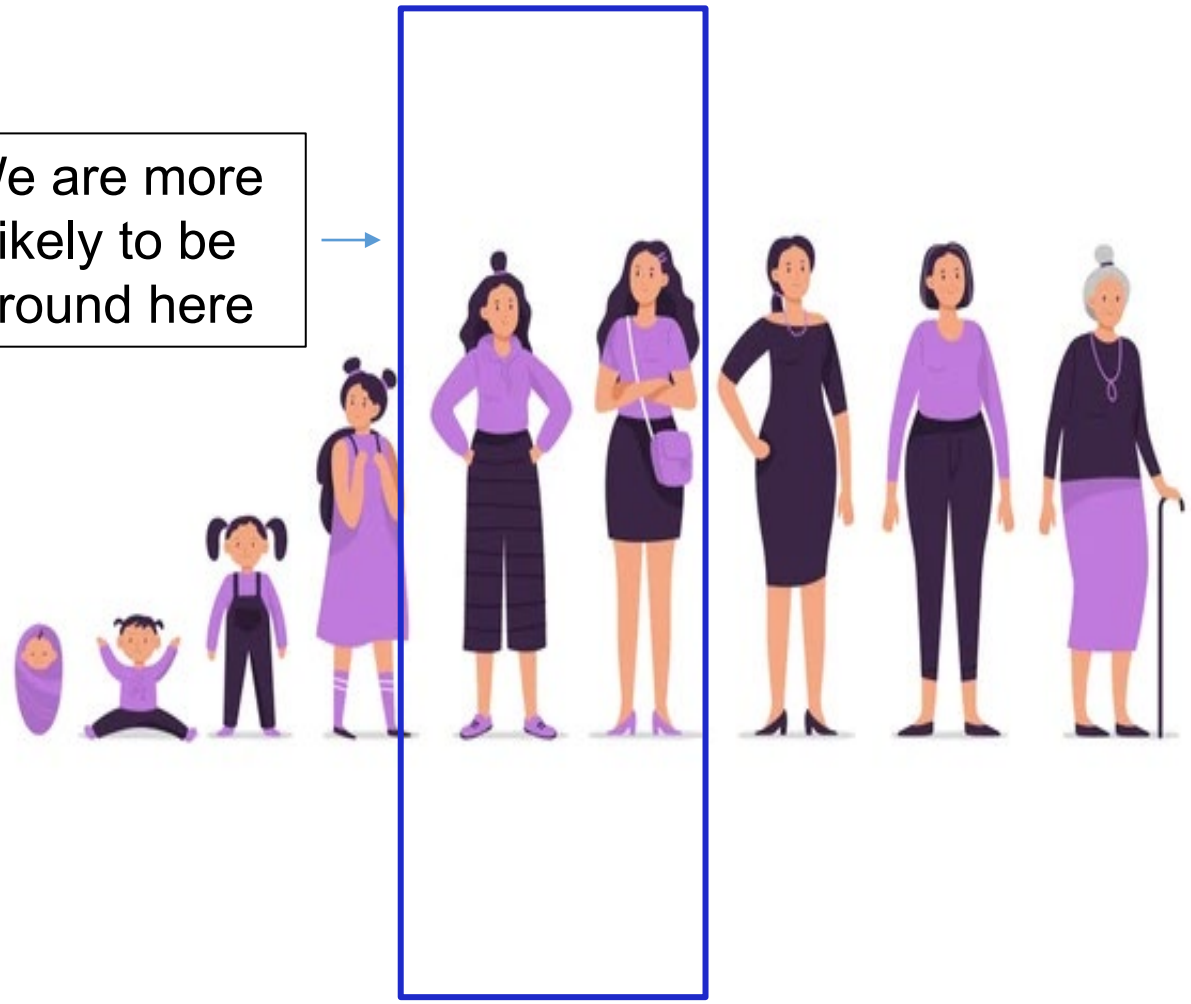
Help me to understand the pattern in the output quickly and project what could happen based on what is happening and trends

Sentiment analysis

Nowcasting

Machine - Sentiment analysis

We are more likely to be around here



What is it?

- Using a machine to read text as a human would and providing the aggregate view of a position, an idea etc.

- For example, find the aggregate view on an industry, sector, country etc based on analyst reports, newsflow, tweets, videos etc.

Requirements

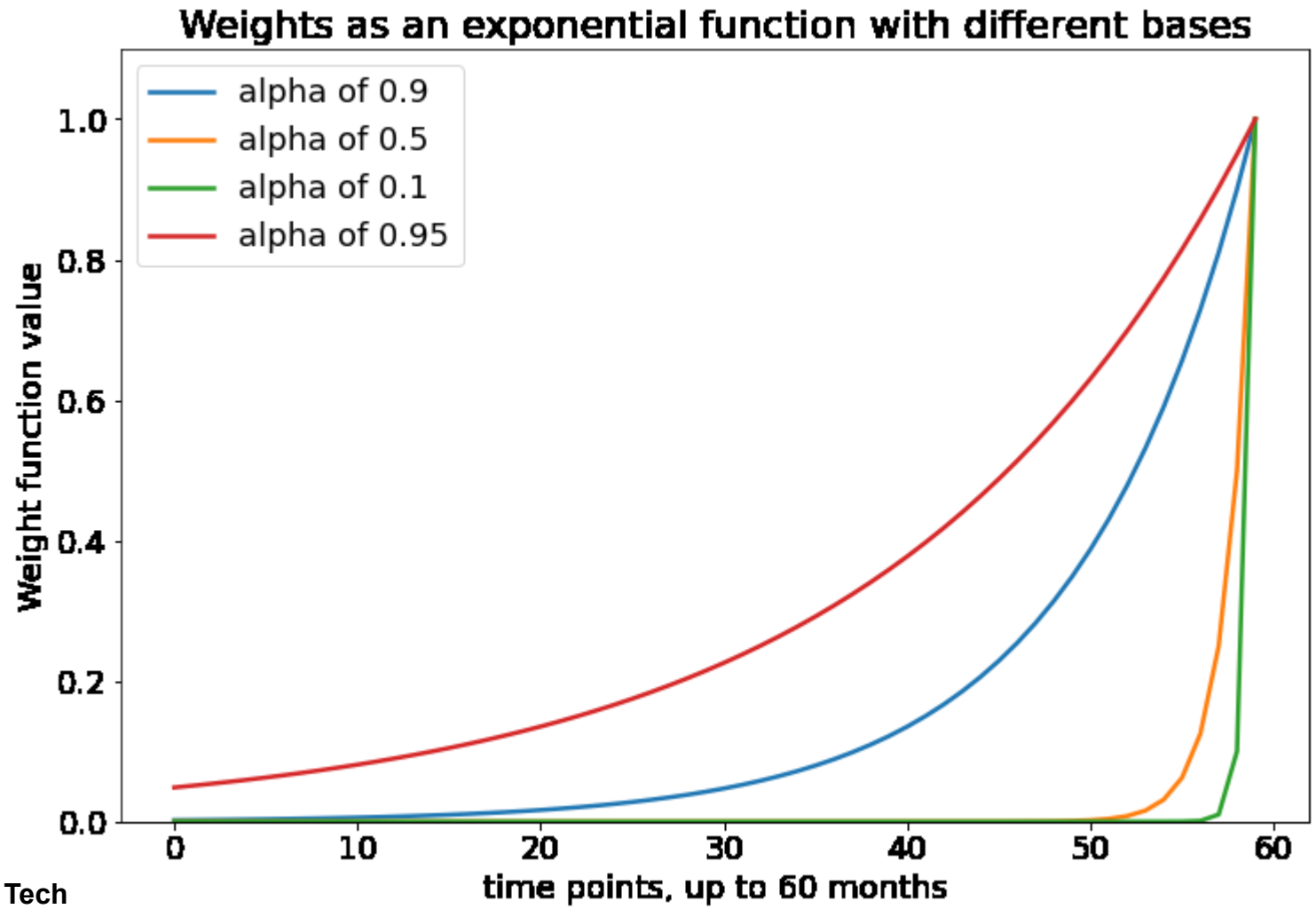
- **Library of relatable** words
- **Dictionary**
- Accurate labelling and processing a.k.a interpretation of the context of the work (i.e. natural language processing programme to use)

Considerations that could skew outcome

- Data needs to be cleaned before ingestion
- Natural Language Processing programme used
- Training engine needs to be specific to the industry for it to make sense

Machine - Sentiment analysis

Aggregation methodology and rate of news decay



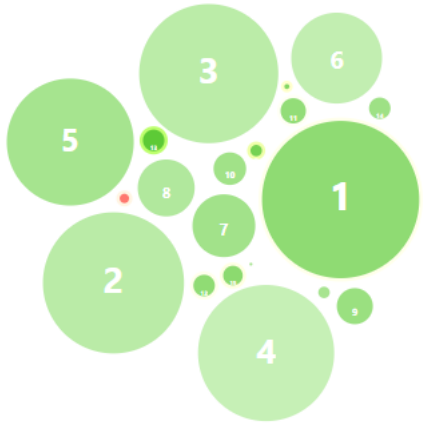
Source: Hunter Labs Tech

Machine - Sentiment analysis application example

Determine and then aggregate the sentiment of industries over time using analyst reports and newsflow

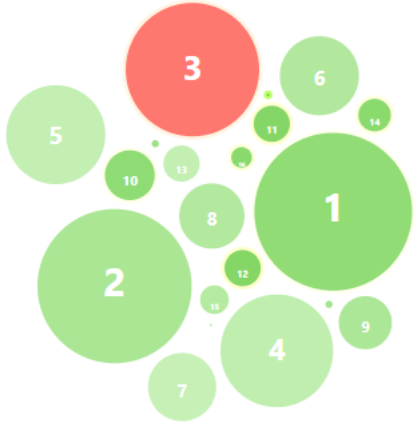
Key
Green – Positive Sentiment
Red – Negative Sentiment
US market only

Sept 2021



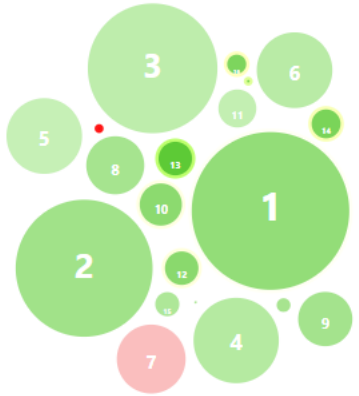
1 - USA BANKS 2 - USA MEDIA 3 - USA INFRASTRUCTURE
4 - USA ENERGY 5 - USA RETAIL 6 - USA REAL ESTATE
7 - USA INTERNET 8 - USA FOOD 9 - USA INSURANCE
10 - USA LENDING 11 - USA ADVERTISING
12 - USA FINANCIALS 13 - USA MARKETING
14 - USA MATERIALS 15 - USA FINANCIAL SERVICES
16 - USA CONSTRUCTION 17 - USA INDUSTRIALS
18 - USA SOFTWARE 19 - USA CAPITAL MARKETS
20 - USA UTILITIES

June 2021



1 - USA BANKS 2 - USA ENERGY 3 - USA RETAIL
4 - USA INFRASTRUCTURE 5 - USA MEDIA
6 - USA REAL ESTATE 7 - USA INSURANCE
8 - USA FINANCIAL SERVICES 9 - USA CAPITAL MARKETS
10 - USA FINANCIALS 11 - USA INDUSTRIALS
12 - USA MATERIALS 13 - USA WEALTH MANAGEMENT
14 - USA LENDING 15 - USA CONSTRUCTION
16 - USA RENEWABLE ENERGY 17 - USA PHARMA
18 - USA UTILITIES 19 - USA CONSUMER DISCRETIONARY
20 - USA OIL GAS

March 2021



1 - USA BANKS 2 - USA ENERGY 3 - USA RETAIL
4 - USA INFRASTRUCTURE 5 - USA INSURANCE
6 - USA REAL ESTATE 7 - USA MEDIA
8 - USA CAPITAL MARKETS 9 - USA FINANCIAL SERVICES
10 - USA LENDING 11 - USA WEALTH MANAGEMENT
12 - USA FINANCIALS 13 - USA MATERIALS
14 - USA INDUSTRIALS 15 - USA CONSTRUCTION
16 - USA RENEWABLE ENERGY 17 - USA PHARMA
18 - USA AIRLINES 19 - USA CONSUMER DISCRETIONARY
20 - USA MINING

Source: Hunter Labs Tech, Sentiment view on US only

Machine - Sentiment analysis application example

Neutral positions could suddenly become positive or negative – delta modelling

Sept 2021

1 - USA INVESTMENT BANKING 2 - USA BUILDING PRODUCTS
3 - USA DISTRIBUTORS 4 - USA HOTELS RESTAURANTS LEISURE
5 - USA CONSTRUCTION ENGINEERING
6 - USA ENERGY EQUIPMENT SERVICES
7 - USA LEISURE PRODUCTS 8 - USA MULTILINE RETAIL
9 - USA TEXTILES APPAREL LUXURY GOODS
10 - USA CONSTRUCTION MATERIALS 11 - USA GLASS
12 - USA HEALTHTECH 13 - USA LIFE INSURANCE
14 - USA METALS MINING 15 - USA MORTGAGE REITS
16 - USA AEROSPACE 17 - USA BIOTECHNOLOGY
18 - USA CONSUMER SERVICES
19 - USA HEALTH CARE EQUIPMENT
20 - USA OPERATING SYSTEMS

June 2021

1 - USA EMAIL 2 - USA INVESTMENT BANKING
3 - USA BUILDING PRODUCTS
4 - USA CONSTRUCTION ENGINEERING
5 - USA ENERGY EQUIPMENT SERVICES 6 - USA ENGINEERING
7 - USA FINTECH 8 - USA HOTELS RESTAURANTS LEISURE
9 - USA LEISURE PRODUCTS 10 - USA MULTILINE RETAIL
11 - USA TEXTILES APPAREL LUXURY GOODS

March 2021

1 - USA INTERNET 2 - USA BUILDING PRODUCTS
3 - USA CONSTRUCTION ENGINEERING 4 - USA EMAIL
5 - USA ENERGY EQUIPMENT SERVICES 6 - USA ENGINEERING
7 - USA HOTELS RESTAURANTS LEISURE
8 - USA INVESTMENT BANKING 9 - USA LEISURE PRODUCTS
10 - USA MULTILINE RETAIL
11 - USA TEXTILES APPAREL LUXURY GOODS 12 - USA FINTECH
13 - USA MORTGAGE REITS 14 - USA BEVERAGES
15 - USA BLOCKCHAIN 16 - USA LIFE INSURANCE
17 - USA OPERATING SYSTEMS 18 - USA RETAIL REITS
19 - USA ROBO ADVISORS 20 - USA SOCIAL MEDIA

Source: Hunter Labs Tech, Sentiment view on US only

Machine - Nowcasting

Useful methodology to help determine manager's potential to outperform peers based on what we know - even more useful when data have some sort of persistence – GDP, Private Equity manager selection

Machine learns behaviour for nowcasting

Our goal was to estimate overperformance/underperformance of private equity / venture funds. Why?

- ☑ Overperformers are most likely to repeat future performance
- ☑ Under and over performers can help capital managers understand which **industries and regions (feedback loop)** are performing well
- ☑ Helps to reduce familiarity bias when selecting investment teams / managers - see the forest for the trees

Machine - Nowcasting

Our approach

- ☑ We used machine learning for nowcasting – we are interested in <u>pattern recognition</u>
- ☑ Predict log(DPI) (one factor)
  - ☑ over >256k data points
  - ☑ factors considered include vintage, DPI momentum, sector and region

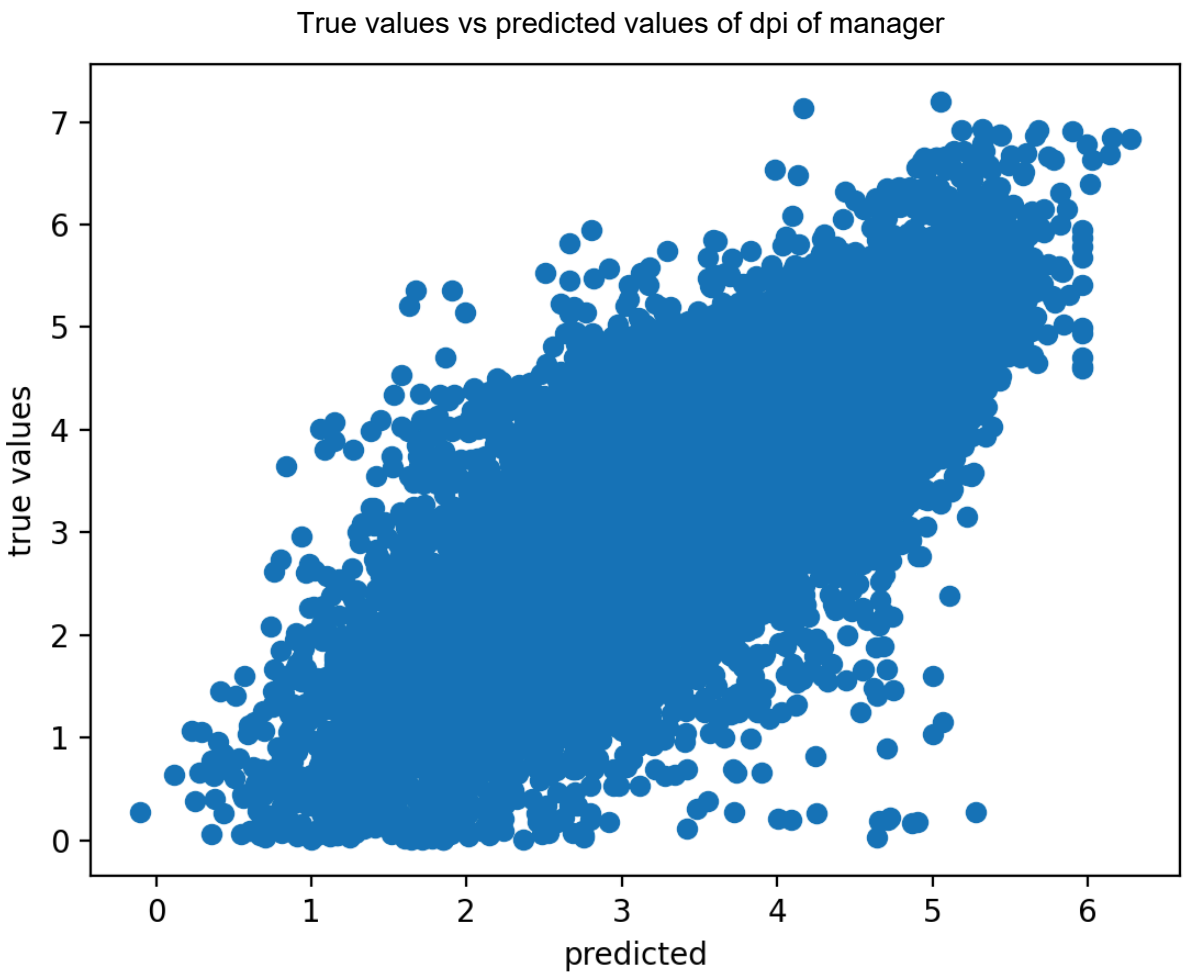
An outcome of our proprietary model using the traditional data sets

We identify a manager's true value* which can lead to better investment selection decisions.

The aim is to predict the **strength of the manager** knowing their ability to distribute back original invested capital + additional return on the invested capital.

Our results:

An output with an adjusted R2 ~70% meaning there is a strong explanatory power of our proprietary models.



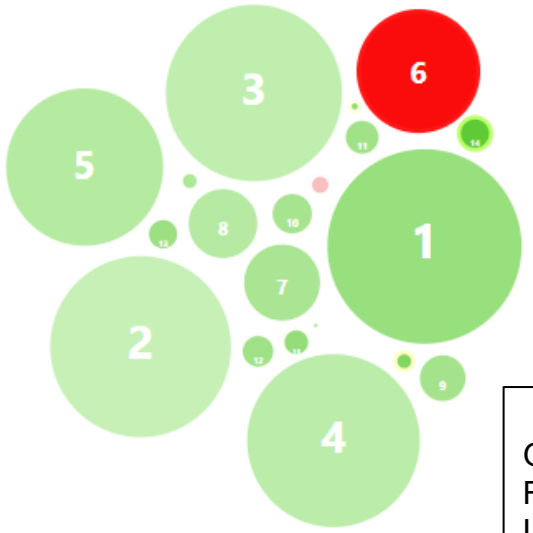
*We are predicting the immediate behavior of a manager based on their recent cadence value to determine their ability to outperform or underperform

Applications – Manager Selection; Human + Machine

In the immediate sense, for **private equity selection,** I may prefer to find value in a neutral sentiment industry / sector

For long and hedge, I may prefer to tilt toward positive sentiment ideas OR short negative sentiment ideas

Find me the best managers in applicable sectors / industries to execute



Key
Green – Positive Sentiment
Red – Negative Sentiment
US market only

1 - USA BANKS 2 - USA ENERGY 3 - USA MEDIA
4 - USA INFRASTRUCTURE 5 - USA RETAIL
6 - USA REAL ESTATE 7 - USA INTERNET 8 - USA FOOD
9 - USA INSURANCE 10 - USA LENDING
11 - USA ADVERTISING 12 - USA MATERIALS
13 - USA FINANCIALS 14 - USA MARKETING
15 - USA FINANCIAL SERVICES 16 - USA SOFTWARE
17 - USA CONSTRUCTION 18 - USA INDUSTRIALS
19 - USA CAPITAL MARKETS 20 - USA UTILITIES

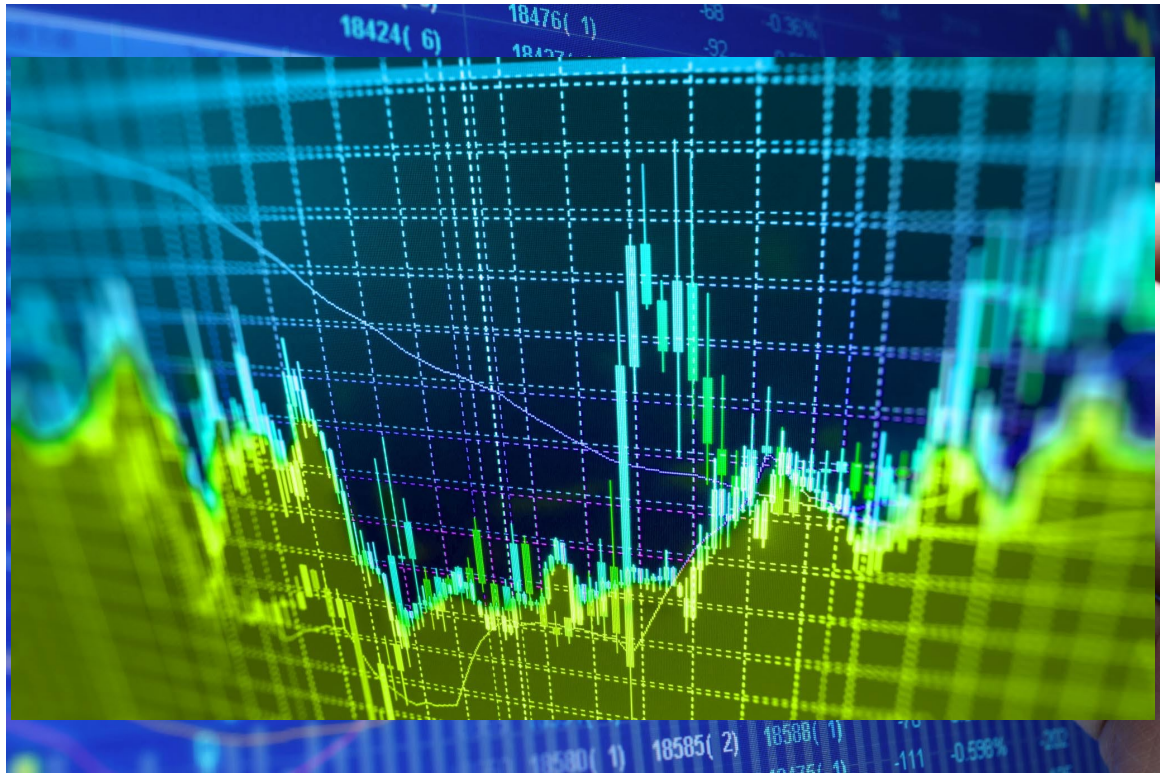
Source: Hunter Labs Technologies, as at Oct 25th

Applications – Secondary transactions

Introduce more science into portfolio secondary transactions given what we know on industry sentiments

Which of the portfolio constituents are stars and which are dogs at sale date and can dogs resurrect to become stars?

Nowcasting extended to underlying portfolio companies provides more colour to future behaviour of residual portfolio to buyer (and seller)



Applications – Secondary transactions

Fund	**Fund ABC**				
Region	**USA**				
Stratey	**Diversified**				
Sale Date	**Jan-21**				
Portfolio	**Industry**	**Sentiment**	**Company Fundamentals**	**Star / Dog**	**Impact on Price**
Portfolio Co 1	Software	Neutral	Negative Trend	Dog	NAV
Portfolio Co 2	Retail	Negative	Negative Trend	Dog	Discount
Portfolio Co 3	Healthcare	Positive	Normal Trend	Star	Premium
Portfolio Co 4	Entertainment	Negative	Slight Negative Trend	Dog	Discount
Portfolio Co 5	E-Commerce	Positive	Normal Trend	Star	Premium
Portfolio Co 6	Pet Care	Neutral	Normal Trend	Star	NAV
Portfolio Co 7	Edtech	Positive	Positive Trend	Star	Premium
Portfolio Co 8	Pharma	Positive	Negative Trend	Dog	NAV
Portfolio Co 9	Alternative Medicine	Positive	Positive Trend	Star	Premium
Portfolio Co 10	Household Durables	Neutral	Negative Trend	Dog	NAV
Overall Impact on Price					**NAV**

Source: Hunter Labs Tech, For illustration only

Applications – Secondary transactions

Fund	**Fund ABC**				
Region	**USA**				
Stratey	**Diversified**				
Sale Date	**Sep-21**				
Portfolio	**Industry**	**Sentiment**	**Company Fundamentals**	**Star / Dog**	**Impact on Price**
Portfolio Co 1	Software	**Positive**	Negative Trend	Dog	Premium
Portfolio Co 2	Retail	Negative	Negative Trend	Dog	Discount
Portfolio Co 3	Healthcare	Positive	Normal Trend	Star	Premium
Portfolio Co 4	Entertainment	**Positive**	Slight Negative Trend	Dog	NAV
Portfolio Co 5	E-Commerce	Positive	Normal Trend	Star	Premium
Portfolio Co 6	Pet Care	**Positive**	Normal Trend	Star	Premium
Portfolio Co 7	Edtech	Positive	Positive Trend	Star	Premium
Portfolio Co 8	Pharma	Positive	Negative Trend	Dog	NAV
Portfolio Co 9	Alternative Medicine	Positive	Positive Trend	Star	Premium
Portfolio Co 10	Household Durables	Neutral	Negative Trend	Dog	NAV
Overall Impact on Price					**Premium**

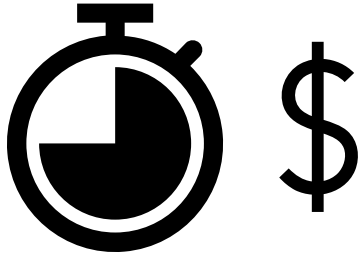
Source: Hunter Labs Tech, For illustration only

Why does this matter

Democratisation is hard to achieve if individuals have no means to sell private equity positions easily

Further applications to make idea generation easier for long only and hedge

We can all save time (and money) in analysing managers and investment opportunities



Practical Machine Learning in Asset Management

A Short Talk About Skill Detection and Other Field Notes

Adam Duncan

Head of Cambridge Associates Investment Science Unit

Fall 2021

CAMBRIDGE ASSOCIATES

A typical conversation…

Boss: Use some of this AI/ML stuff and tell me which of these 100 funds is going to outperform their benchmark over the next 5 years.

You: Um, that seems like it might be a bit ambitious.

Boss: I can order things on the Internet with my voice.

You: …

The Jesus in the Toast*. This is a massive problem for us.

*Full credit to Campbell Harvey for finding this picture and using it in a presentation. Brilliant.

Lots of literature on skill detection…

Luck versus Skill in the Cross-Section of Mutual Fund Returns

EUGENE F. FAMA and KENNETH R. FRENCH*

Talent vs Luck:
the role of randomness in success and failure

A. Pluchino*, A. E. Biondo†, A. Rapisarda‡

MUTUAL FUND PERFORMANCE: SKILL OR LUCK?

**Keith Cuthbertson*, Dirk Nitzsche*
and
Niall O'Sullivan****

FLAW IN THE FUND SKILL/LUCK TEST METHOD OF CUTHBERTSON ET AL (SSRN Abstract 665744)

John Nuttall

Lucky Factors

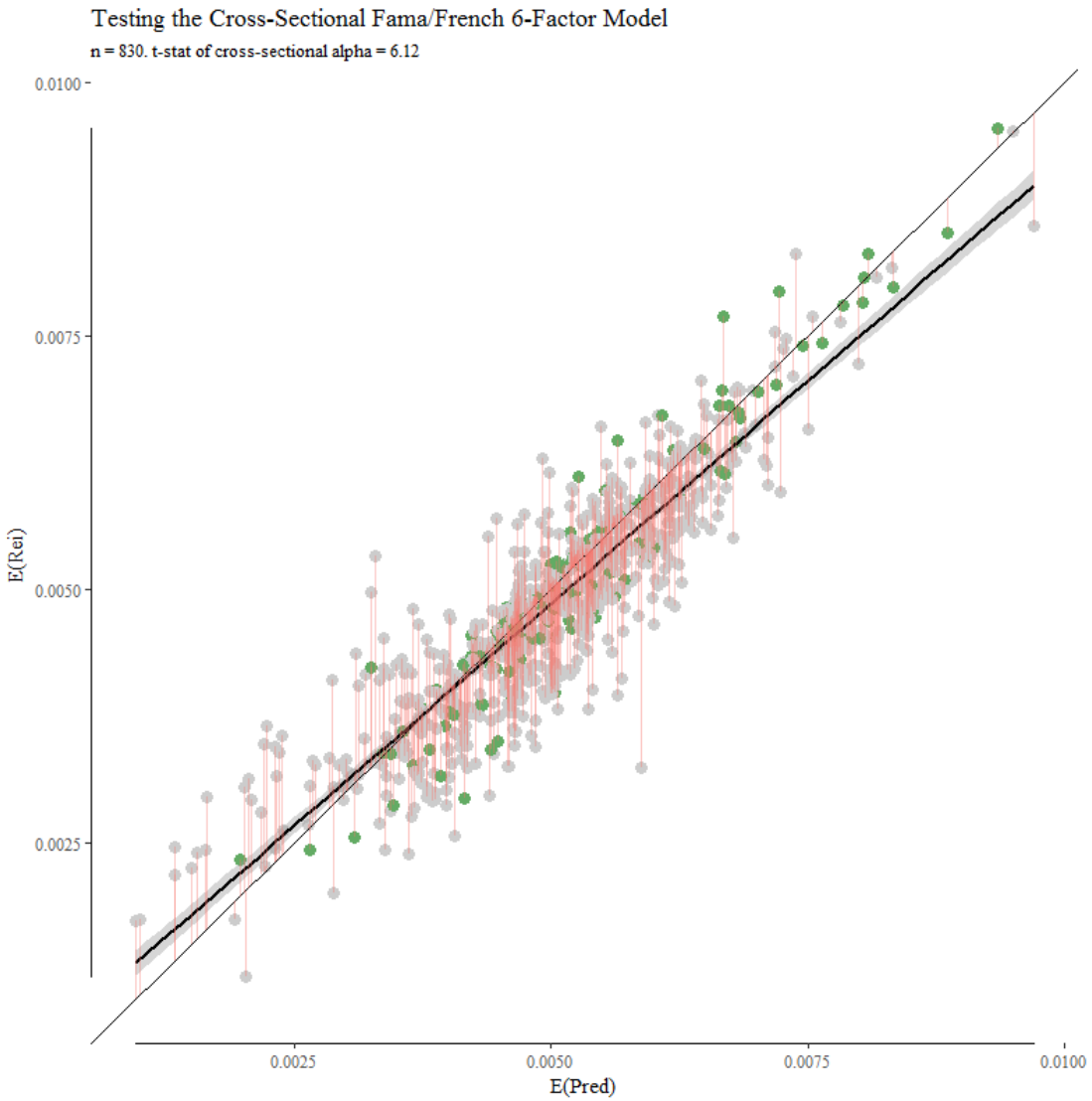
Campbell R. Harvey

Duke University, Durham, NC 27708 USA
National Bureau of Economic Research, Cambridge, MA 02138 USA

Yan Liu*

Texas A&M University, College Station, TX 77843 USA

1126 Long only equity managers. Minimum 10yr track record. Pricing error ~57 bps per annum.



Soure: Data: Cambridge Associates and AQR Data Library (https://www.aqr.com/Insights/Datasets); Analysis : A. Duncan

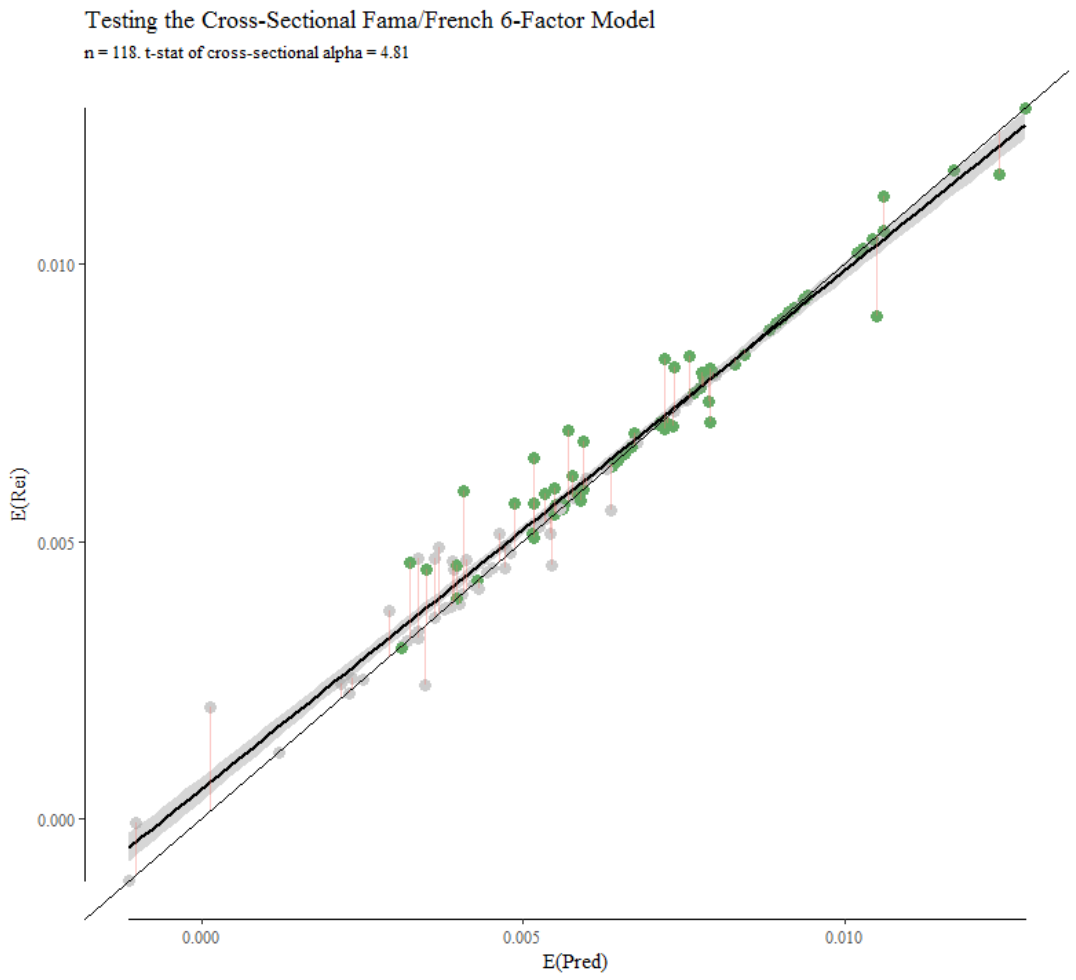
Our empirical tests of cross-sectional asset pricing models are very consistent with the empirical tests in the literature. Where we see high betas, we also see high average returns.

But the line is “too flat”.

This suggests non-zero “alpha” (or missing factors) might be present.

Regardless of your views on asset pricing models, many people compute and actively hunt for, alpha.

160 Equity Long/Short Managers.
CBOE Put Selling Index + FF Factors. Pricing error = ~ 64 bps per year.

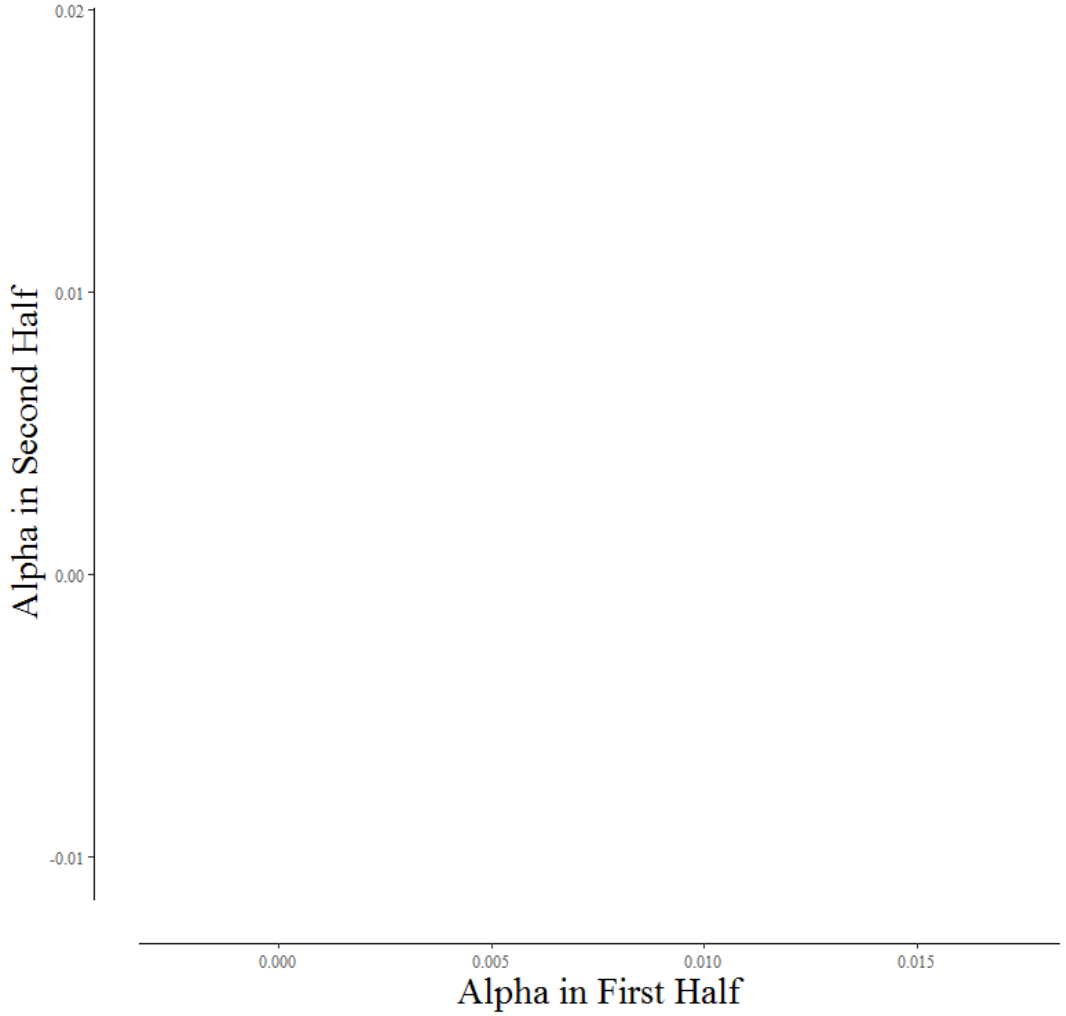


Your mileage may vary depending on the complexity of your model and which factors you include, but the alpha results are robust to changes in strategy type.

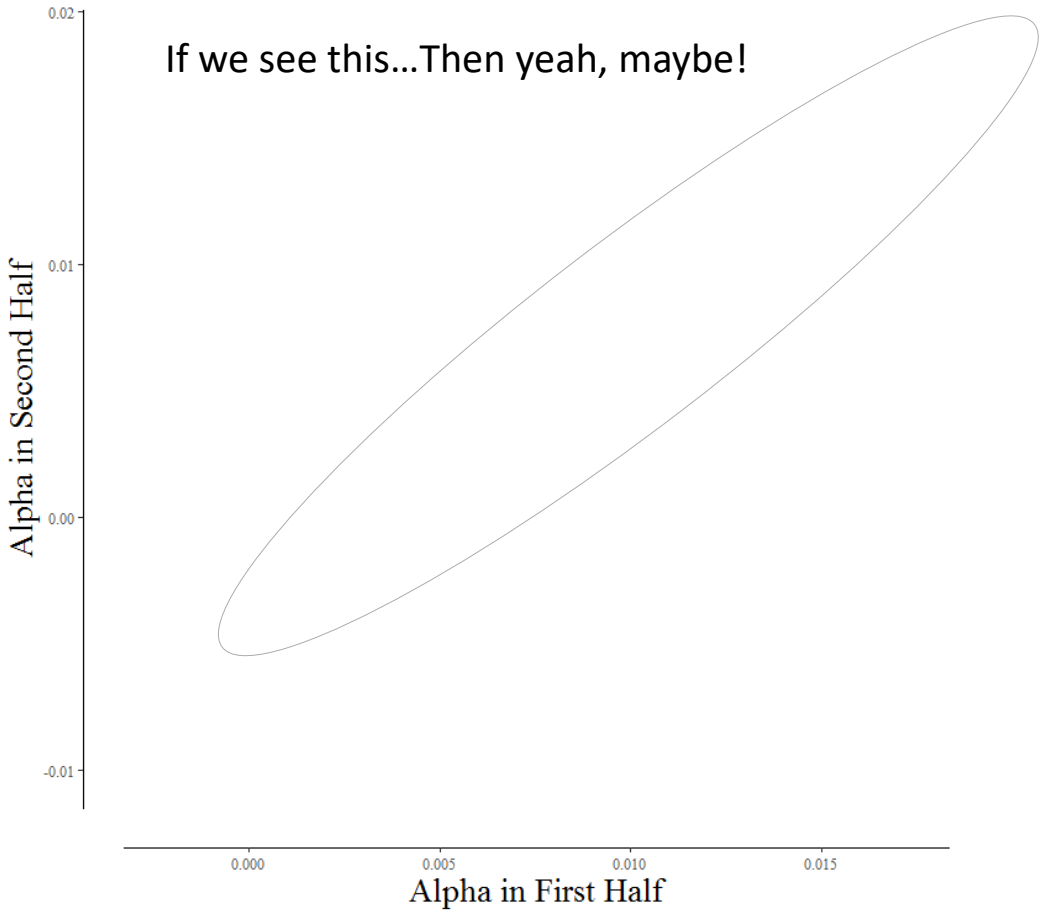
Here is a collection of long/short equity hedge funds.

Soure: Data: Cambridge Associates and AQR Data Library (https://www.aqr.com/Insights/Datasets); Analysis : A. Duncan

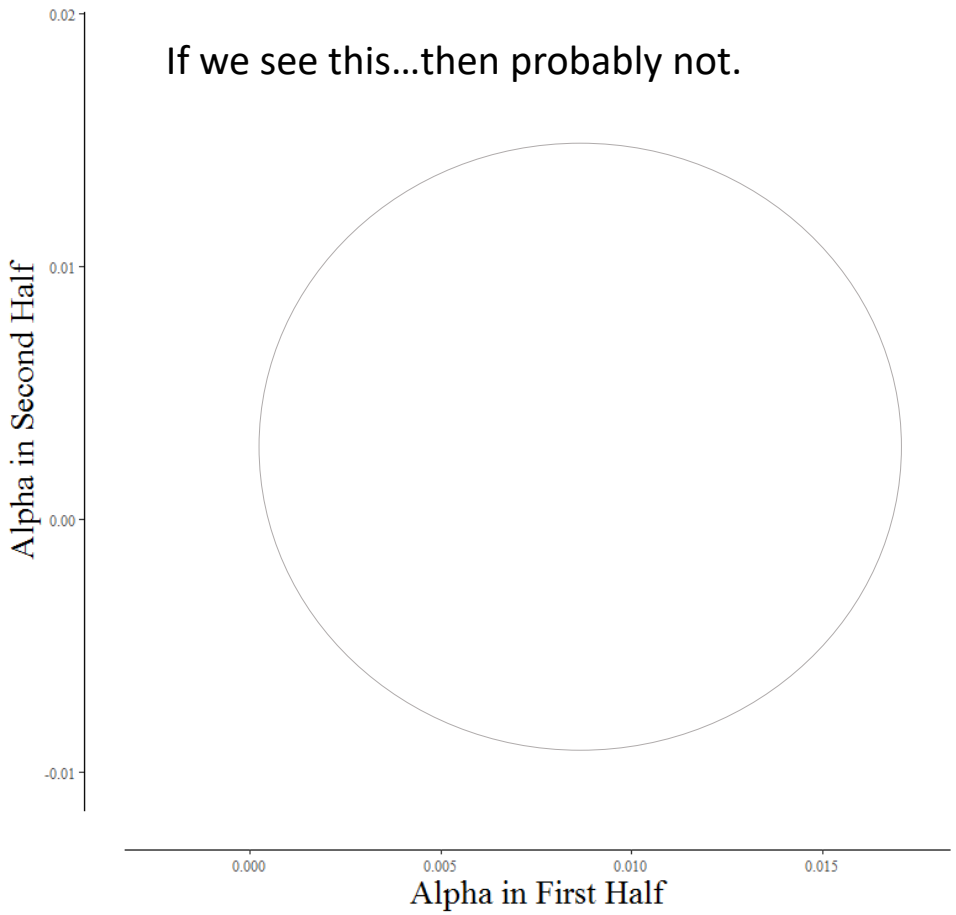
Can we pick funds based on observed alpha? Does alpha have information content?



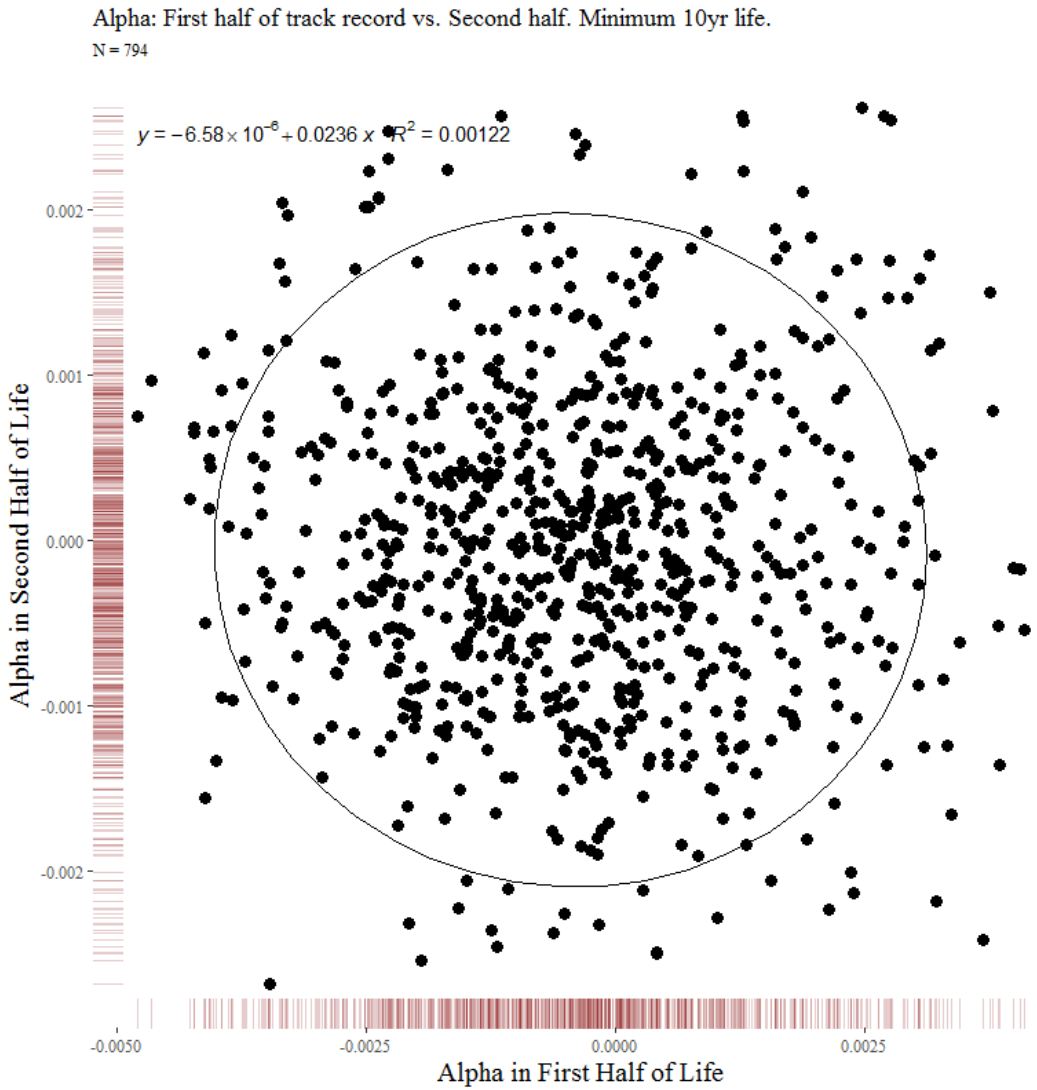
Does alpha have information content?



Does alpha have information content?

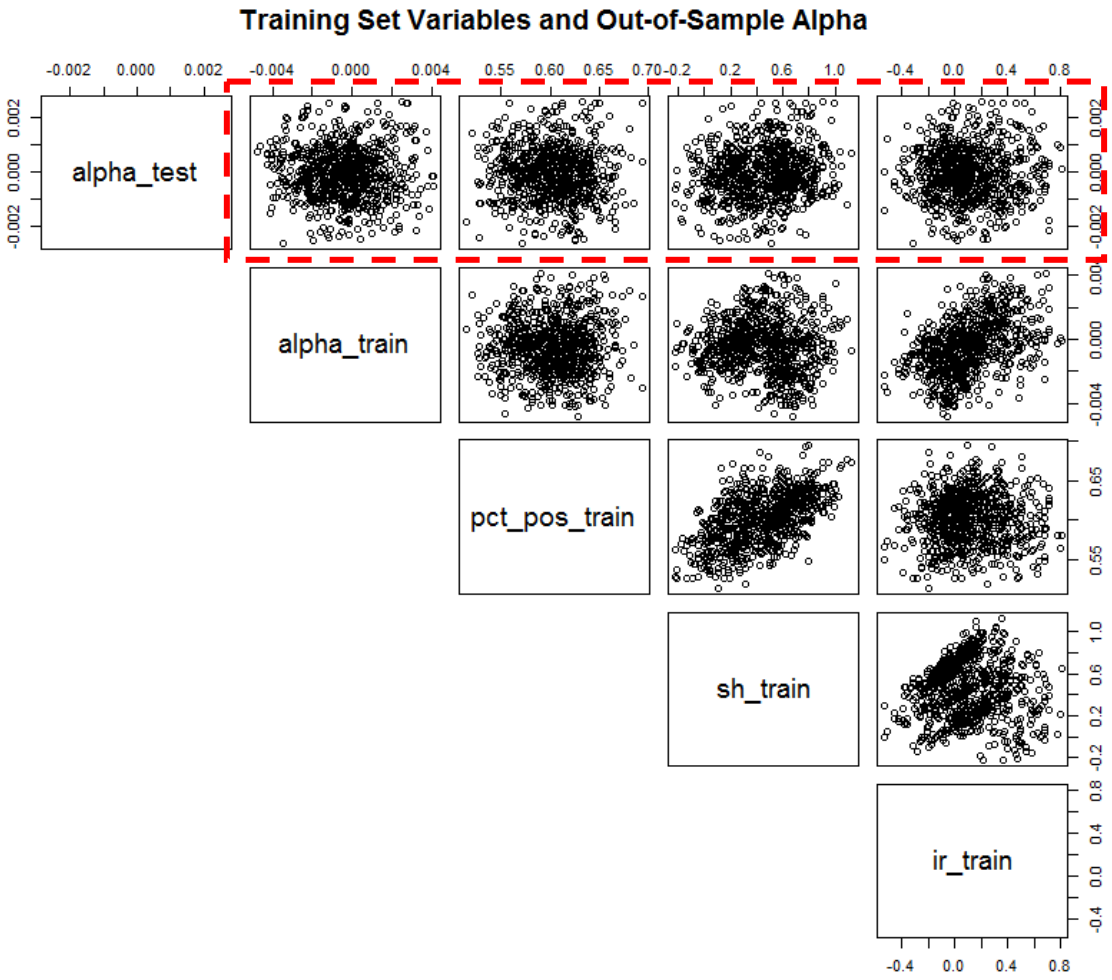


Well…



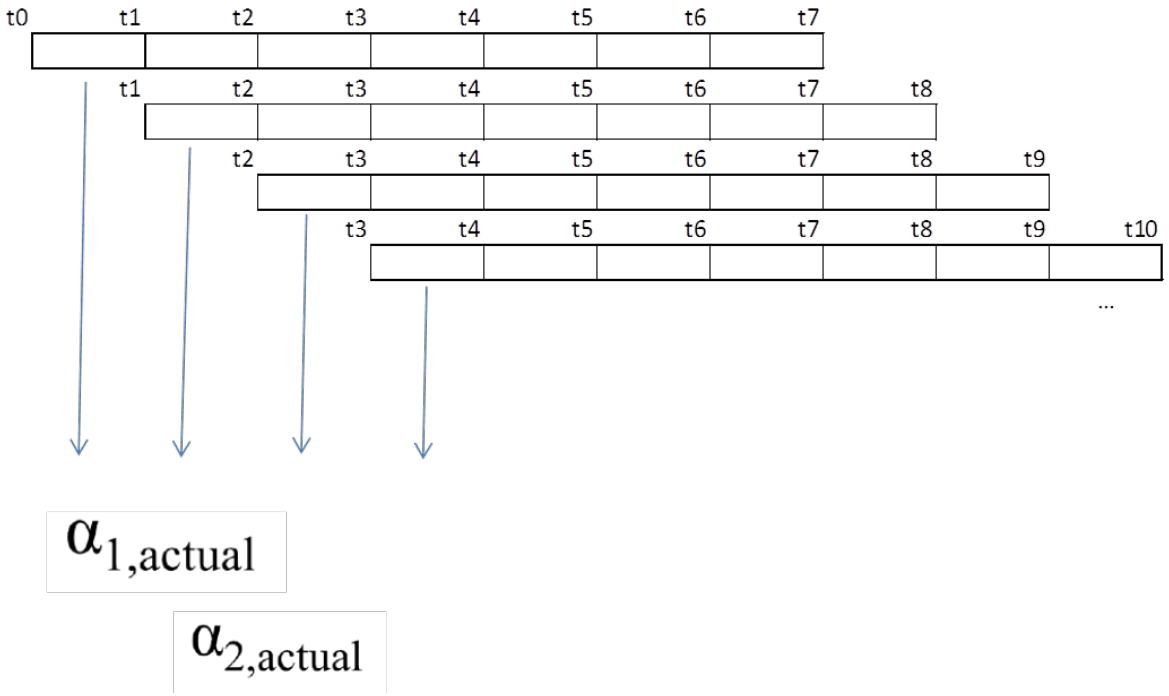
Source: Cambridge Associates. Analysis: A. Duncan

What about other metrics?



Data: Cambridge Associates. Analysis: A. Duncan Comment: Returns are memory-less and so maybe it's not surprising that we don't find any signal in them. See work by Marcos Lopez de Prado on over-differentiation.

Estimate rolling regressions and record the alphas.



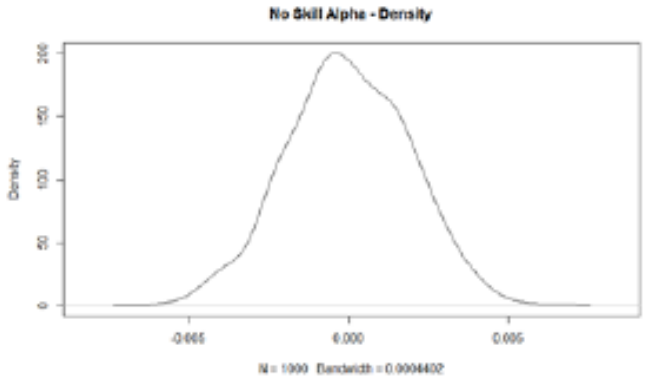
Source: A. Duncan

Bootstrap to uncover the no-skill distribution. (or make some parametric assumptions, if you like)

Bootstrap to get this.*

$$\alpha_{1,\text{actual}}$$

vs.

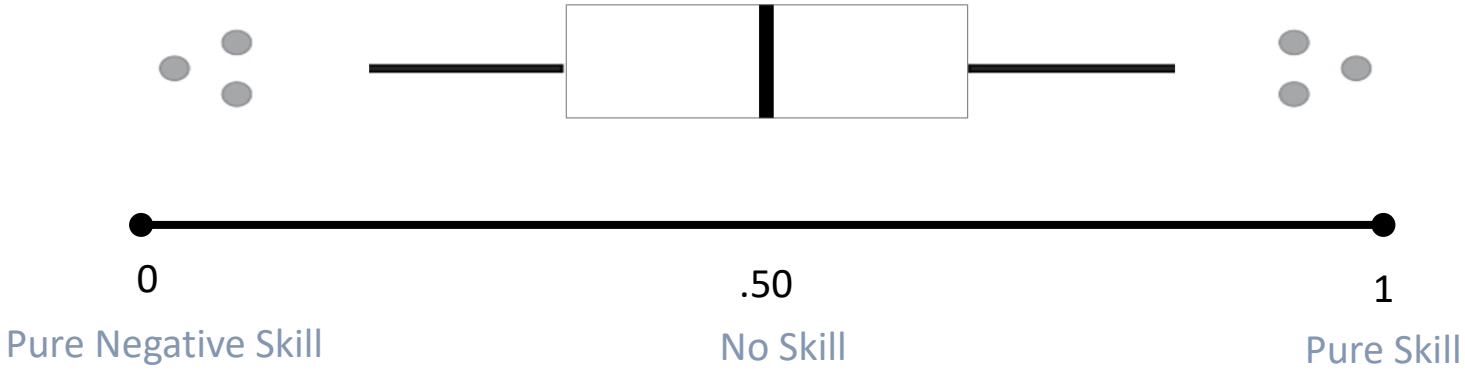


$$\sim\alpha_{1|\text{no\_skill}}$$

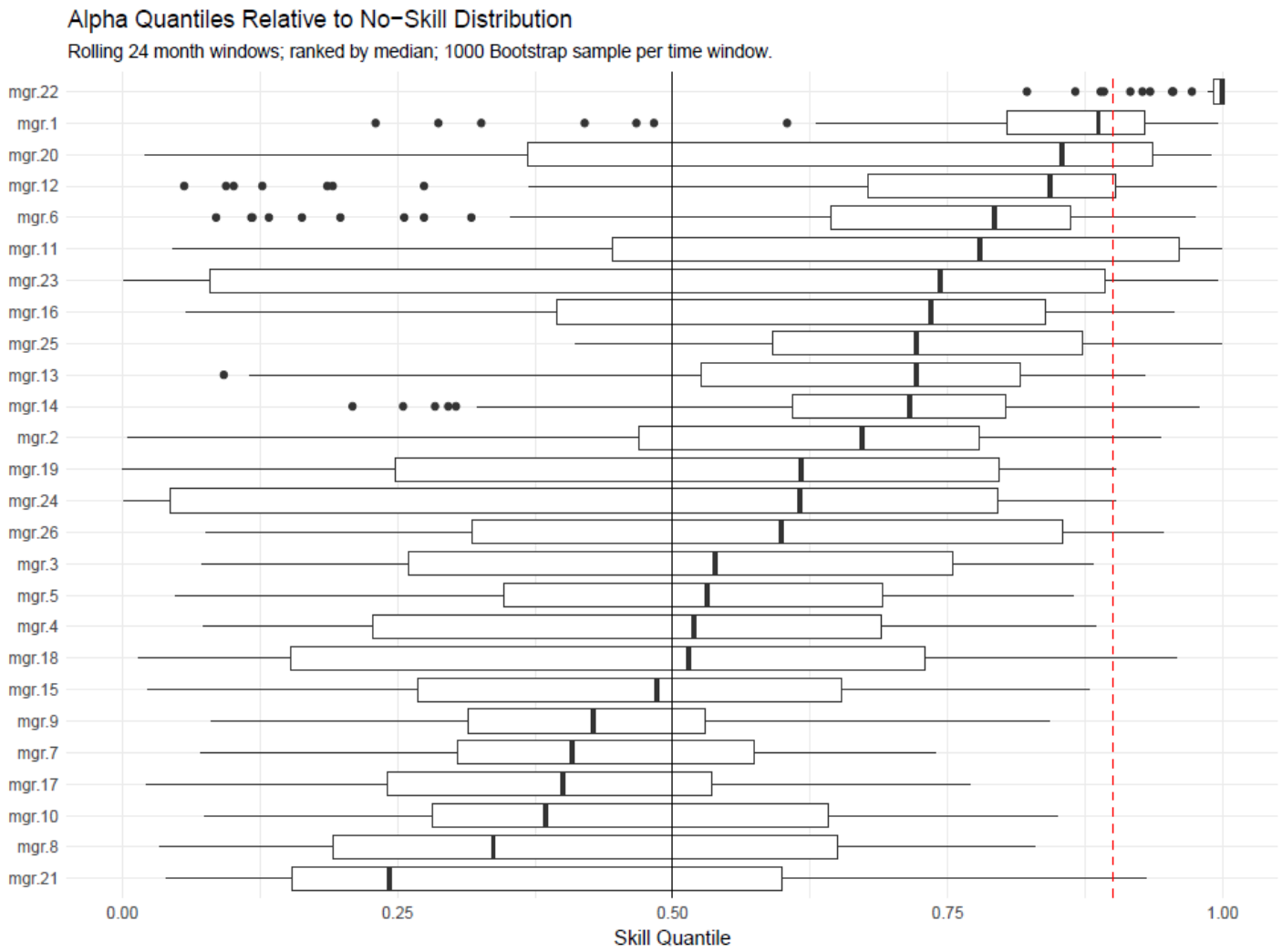
* See “Luck vs. Skill in the Cross-Section of Mutual Fund Returns “, and Gene Fama and Ken French, 2010 for descriptions of methods that are quite like what we are doing.

Skill Quantiles

"Where a manager's alpha fell relative to the no-skill distribution."



Skill Quantiles for a Group of 130/30 Equity Managers



Source: Cambridge Associates. Analysis: A. Duncan

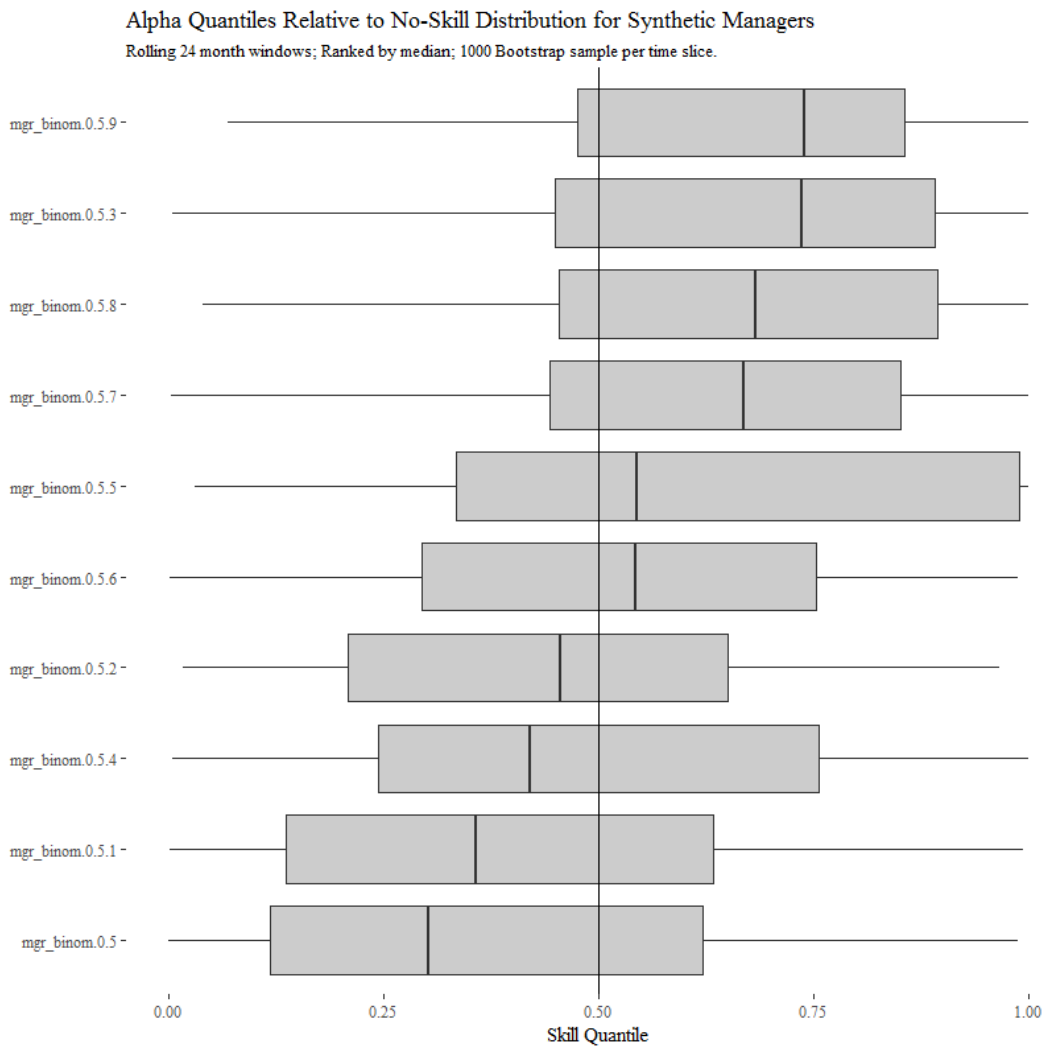
Based on this skill measure, a good portion of these managers have true skill. This should be true given our highly survivor-biased database, but nonetheless.

Look at the highest ranked fund!

Could we still be at risk of false positives here?

Yes.

These are ALL true no-skill, simulated managers. Skill is super noisy!



Analysis: A. Duncan

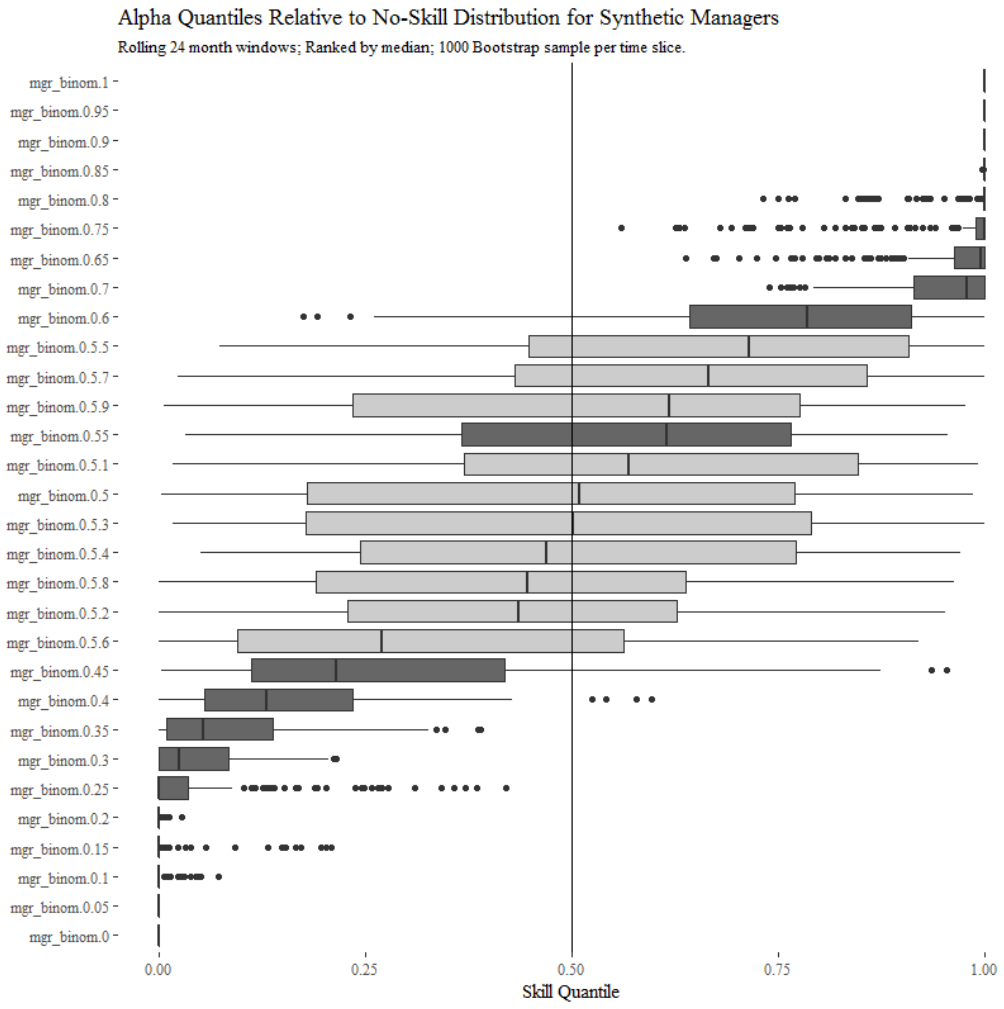
Here is a collection of *simulated* managers who follow a binomial probability process of out-performing their benchmark. We parameterize the managers to out/under-perform the benchmark in an amount consistent with the average manager in the training data.

The top 2 simulated managers had median skill-quantiles of .75!

And similar results were recorded for the bottom two managers.

This is all just randomness.

Where do you think most of the world's *true* skill lies?



Probably here!

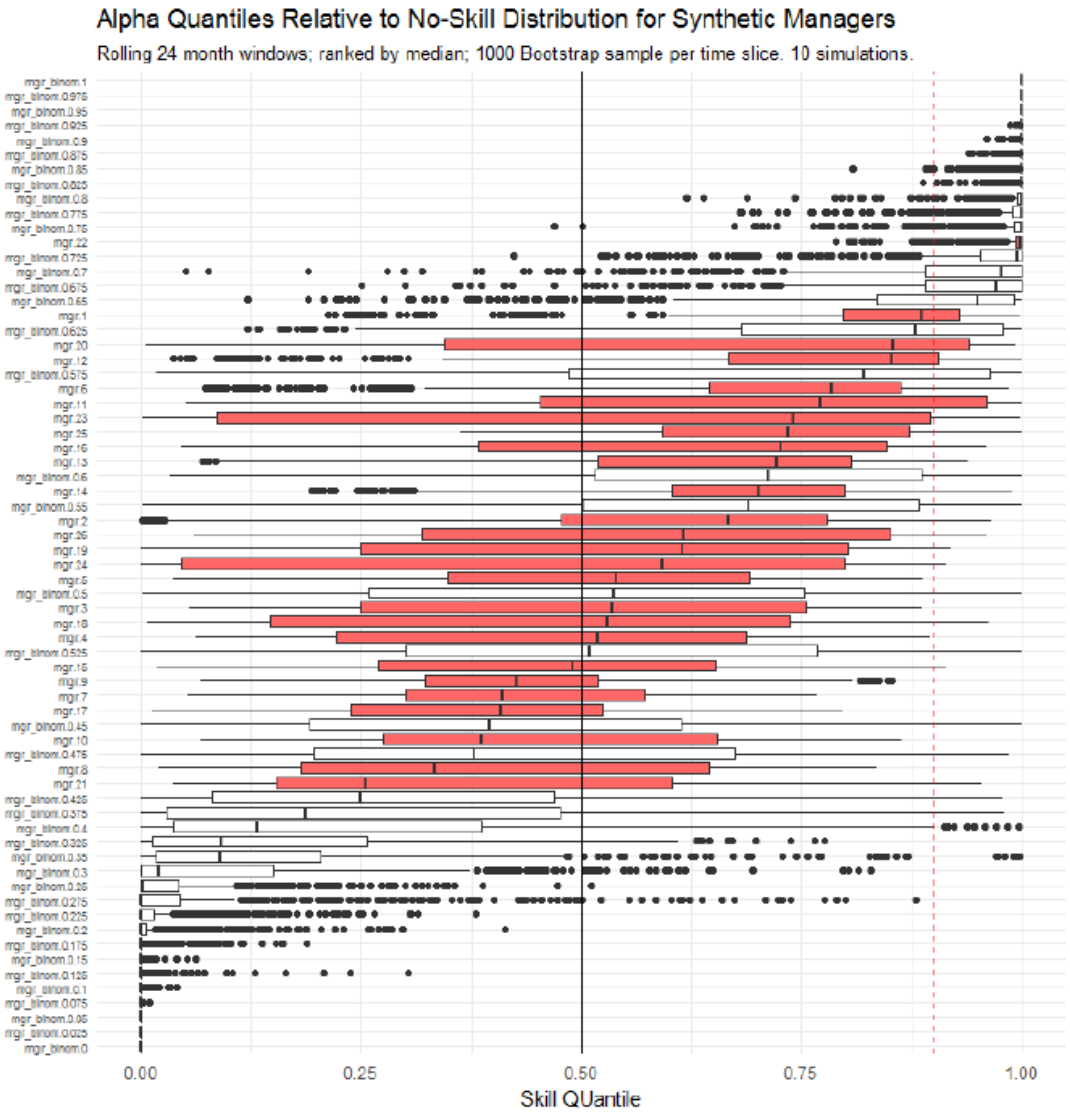
Here is what the entire space of simulated managers looks like when you run it over the interval:
true skill = [0, 1].

The light grey managers have true skill = .50.

The dark managers have non-.50 skill parameters.

Analysis: A. Duncan

Here is the same plot, but now with real managers injected. Real managers appear in red.



This process continues from here in a complicated way where we try to infer each manager's true skill parameter by clustering each manager with its nearest no-skill neighbors.

After ignoring a whole bevy of multiple testing issues, we final arrive…

…at a system that *underperforms* a simple ranking of the in-sample alphas.

This system simply *does not work.*

Source: Cambridge Associates. Analysis by A Duncan.

Some Alpha Analysis Take-Aways

Return based measures are nearly information-less.
Beware using them as your target/response variable.

But(!) financial performance (and other meta data) can be good for:

1. Predicting *what actions your colleagues or clients are likely to take.*
2. Building a model that picks funds *like your colleagues pick funds.*
3. Serving as inputs for things like anomaly detection and governance systems.

Some Machine-Learning Take-aways

Here are some practical pieces of advice for you machine learning adventures in manager selection:

- If you believe that the humans you work with are doing a good job picking managers, then start by modeling *their* process. Interview them extensively and write down what they say.
- Try to solve the problem with simple heuristics *first.* Then progress.
- Fitting a deep neural net should not be your first attempt at solving the problem.
- We spend less than 1% of our time tuning model hyper-parameters.
- We'll choose better data over a better model *every* time.
- Beware of "label entropy". You might have less useable data than you think.

General Problem Identification Take-Aways

In our view, good problems have the following characteristics:

1. There is pain point being widely felt in your organization.
2. There is proximate, relevant data that can be obtained.
3. There is clear first step that is near-term helpful.
4. There is an augmentation component to the solution:
    a. Our model reduces the total time spent on a task (Efficiency)
    b. Our model increases the value of the time spent on a task (Enrichment)

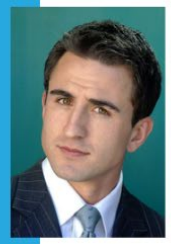
Q&A

FDP INSTITUTE | CAIA ASSOCIATION | Webinar

Turn Text to Alpha α
NLP Sentiment Thought Leadership

December 9, 2021

9am PDT | 12 noon EST | 4pm GMT | 5pm CET | 6pm SAST | 8pm GST | 9.30pm IST |



Dan Joldzic, CFA, FRM
CEO,
Alexandria Technology

Keith Black, Ph.D., CAIA, CFA, FDP
Managing Director, FDP Charter,
FDP Institute

Thank You

- Contact Us:

fdpinstitute.org

info@fdpinstitute.org

@fdpinstitute

linkedin.com/company/fdpinstitute