DESAUTELS Faculty of Management
Faculté de gestion

The Cross-Sectional Distribution of Fund Skill Measures

Laurent Barras (McGill), Patrick Gagliardini (USI and SFI), and Olivier Scaillet (UNIGE and SFI)

McGill

Outline

- Motivation & Contribution
- The Skill Measures
- The Non-Parametric Approach
- Data Description
- Empirical Results

Motivation

- **Extensive literature on performance**
  - Ample evidence of negative net alphas
    - (i) Average: Jensen(68), Elton et al.(93)
    - (ii) Distribution: Barras et al. (10), Harvey & Liu(18a)

- **Far less is known about skill**
  - Usually: bad performance=unskilled…
  - But skill and performance are very different
  - Skill: viewpoint of funds
    - Do they create value via their investment/trading skills?
  - Performance: viewpoint of investors
    - Do they receive some of this value (bargaining power)?

Motivation

- **In this paper, we focus on skill**
  - Estimate the entire distribution of mutual fund skill
  - Beyond the average: capture heterogeneity in skill

- **Novelty: we use a non-parametric approach**
  - No parametric structure imposed on the distribution
  - Bias adjustment: infer *true* skill from estimated skill

- **Why is a non-parametric approach important?**

  (i) Avoid the challenge of specifying the skill distribution

  (ii) Allow to examine jointly multiple skill measures

  - In contrast with std parametric/Bayesian approaches

Contribution

- **We jointly examine four skill measures**
  - For the entire universe of US equity funds

- **The two dimensions of fund skill**

  Investment versus trading skills

  (i) First dollar(fd) alpha: ability to detect trading opportunities

  (ii) Size coefficient: ability to override capacity constraints

- **Aggregating the two skill dimensions**

  Value added (profits) from exploiting skill

  (i) Lifecycle: profits over the entire lifecycle (different size)

  (ii) Steady state: profits after reaching the average size

Contribution

- **Main results: the two skill dimensions**

  – Funds are skilled at detecting profitable trades

  …but are unskilled at resisting capacity constraints

    - For 85% of funds: (i) fd alpha>0, (ii) size coeff.>0

      Average fd alpha: 3.4% per year

      Average size coeff.: 1.4% per year for 1σ

    - Strong heterogeneity + highly non-normal

      Looking at the average not sufficient

  – Both skill measures are positively correlated

    - Importance of aggregating them

Contribution

- **Main results: aggregating the skill dimensions**
  - Funds earn large profits from their skills
    - For 70% of the funds: value added>0

      Average (lifecycle): 1.9 mio. per year

      Average (steady state(ss): 7.1 mio. per year
    - Strong heterogeneity + highly non-normal
  - Profits are not far from level predicted by theory
    - Steady state value added: 2/3 of optimal value added
  - Overall interpretation for funds
    - Skill + bargaining power over investors

Contribution

- **Literature on performance (net alpha)**
  - Average: Carhart(97), Elton et al.(93), Jensen(68), Kosowski et al.(06), Pastor and Stambaugh(02)
  - Distribution: Barras, Scaillet, and Wermers(10), Ferson and Chen(18), Harvey and Liu(18a), Jones and Shanken(05)
- **Literature on fund skill**
  - Berk and van Binsbergen(15), Grinblatt and Titman(89), Pastor, Stambaugh, and Taylor(15), Wermers(00)
- **Literature on time-varying fund returns**
  - Chen et al.(04), Harvey and Liu(18b), Kasperczyk et al.(14), Pastor, Stambaugh, and Taylor(15,17)

The Skill Measures

1) The two skill dimensions

- Gross alpha-size relation (Berk-Green (BG) model):

$$\alpha_{i,t} = a_i - b_i q_{i,t-1}$$

- The first dollar(fd) alpha
  - Ability to detect profitable trades
  - "Paper" return: no drag of real world implementation
- The size coefficient
  - Ability to override capacity constraints
  - Execution costs of trading large orders
  - Fund-specific (departs from previous work)

The Skill Measures

2) Aggregating the two skill dimensions

– Value added: profits from exploiting skill

  - Like monopolist rent: markup price x quantity

– Lifecycle value added (over the fund lifecycle):

$$va_{i,l} = E\left[\alpha_{i,t} q_{i,t-1}\right] = a_i q_{1,i} - b_i q_{2,i}$$

with $q_{1,i} = E\left[q_{i,t-1}\right]$ and $q_{2,i} = E\left[\left(q_{i,t-1}\right)^2\right]$

– Steady state value added (at average size $q_{1,i}$):

$$va_{i,ss} = E\left[\alpha_{i,t}\right] E\left[q_{i,t-1}\right] = a_i q_{1,i} - b_i \left(q_{1,i}\right)^2$$

– We have $va_{i,ss} > va_{i,l}$ as long as $b_i > 0$

The Non-Parametric Approach

- **Overview**

  – Allow for the estimation of the distribution of all skill measures $m_i = \{a_i, b_i, va_{i,l}, va_{i,ss}\}$

    - Applies to density, cumulative function, moments (e.g., mean, variance, skewness), quantiles
    - Focus for the presentation: density

  – Main challenge: inference of the *true* measures m_i based on the estimated ones $\hat{m}_i$

    - Error-in-Variable (EIV) bias

The Non-Parametric Approach

- **Benefits (vs Bayesian/parametric approaches)**
  - Immune to misspecification errors
    - No assumption on the shape of the *true* distribution
  - Joint analysis of all four skill measures
    - No inconsistencies/curse of dimensionality
  - Simple and fast (like an histogram)
    - No Gibbs sampling, EM algorithm
  - Asymptotic theory (double asymptotics)
    - Derivation of asymptotic distribution/bias of the different estimators

The Non-Parametric Approach

1) Estimation of the skill measures

– Times-series regression for each fund (gross return)

$$r_{i,t} = a_i - b_i q_{i,t-1} + \beta'_i f_t + \varepsilon_{i,t} \quad i = 1, \ldots, n$$

- Random coefficient model

– Estimated skill measures:

$$\hat{m}_i = \{\hat{a}_i, \hat{b}_i, \; \hat{a}_i \bar{q}_{1,i} - \hat{b}_i \bar{q}_{2,i}, \; \hat{a}_i \bar{q}_{1,i} - \hat{b}_i (\bar{q}_{1,i})^2\}$$

– Fund selection (I_i=1 if fund *i* is selected)

- Rule 1: minimum number of observations T (60)
- Rule 2: condition number on covariance mat. (15)

The Non-Parametric Approach

2) Standard non-parametric estimator

– Density estimator at a given point m (Silverman(86))

$$\hat{\varphi}(m) = \frac{1}{nh}\sum_{i=1}^{n} I_i K\left(\frac{\hat{m}_i - m}{h}\right)$$

where K is the Kernel function (Gaussian)

h is the bandwidth

– **Prop. 1**: properties of $\hat{\varphi}(m)$ (n,T→∞, h →0)

$$\sqrt{nh}(\hat{\varphi}(m) - \varphi(m) - bs(m)) = N\left(0, K_1\varphi(m)\right)$$

where K_1 is a constant ($K_1 = \int K(u)^2\, du$)

bs is the bias term

The Non-Parametric Approach

- **Key insights from Proposition 1**
  - $\widehat{\varphi}(m)$ is asymptotically normally distributed
  - Help with the choice of the optimal bandwidth
    - If T small wrt n: we have $h^{*} \div (n/T)^{-1/3}$
  - The bias term bs has two components

    (i) Smoothing bias bs_1 (standard)

    (ii) Error in variable (EIV) bias bs_2 (non-standard)

    - Both terms depend on the *true* density $\varphi(\alpha)$

The Non-Parametric Approach

3) Adjustment for the bias

– Use a reference model to estimate the bias bs

$m_i \sim N(\mu_m, \sigma_m)$ and $s_i \sim N(\mu_s, \sigma_s)$ with correl ρ

where $s_i = \log(S_i)$

S_i is the asymptotic variance of $\sqrt{T}\hat{m}_i$

– **Prop. 2**: bias under the ref. model (as $n, T \rightarrow \infty$, $h \rightarrow 0$)

(i) Smoothing bias: $bs^r_1(m) = f(\mu_m, \sigma_m)\phi\left(\frac{m-\mu_m}{\sigma_m}\right)$

(ii) EIV bias: $bs^r_2(m) = f(\mu_m, \sigma_m, \mu_s, \sigma_s, \rho)\ \phi\left(\frac{m-\mu_m-\rho\sigma_m\sigma_m}{\sigma_m}\right)$

The Non-Parametric Approach

- **Key insights from Proposition 2**
  - Offer several advantages:
    - Simple to compute
    - Estimation precision (5 parametres)
    - Closed form: comparative static analysis
  - One thing to check:
    - Accuracy when the reference model does not hold?
    - Checked it is the case with Monte-Carlo analysis

The Non-Parametric Approach

4) Bias-adjusted density estimator

– Estimate the five moments of the reference model

  - From the estimated skill measures
  - From the estimated variance (Newey-West)

– Compute the bias terms using Prop. 2

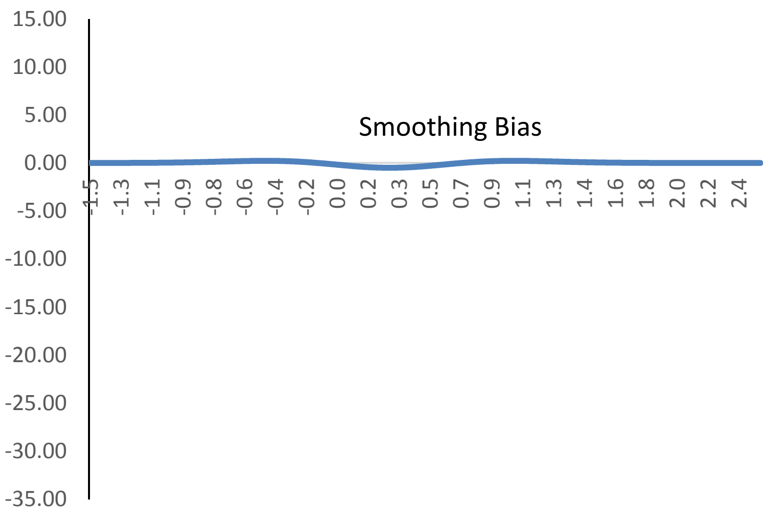
– Bias-adjusted density estimator

$$\hat{\varphi}^*(m) = \hat{\varphi}(m) - \widehat{bs}^r(m)$$

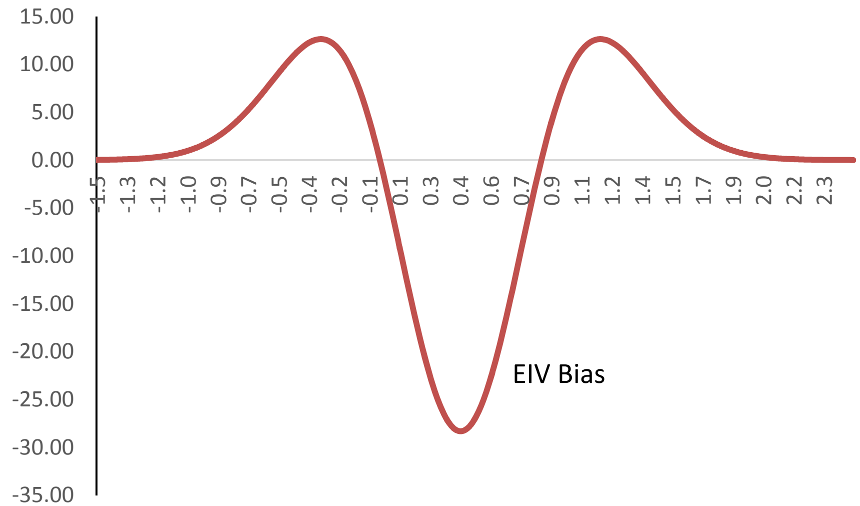
The Non-Parametric Approach

Bias components for first dollar alpha

Smoothing bias



EIV bias



Play no role given the large number of funds n (more than 2,200)

Bias is massive

Intuition: $\hat{a}$=a+estimation noise overestimate prob. of observing large a

Shape is general as long as *true* density peaks at the mean

Data Description

- **Mutual funds**
  - All open-end, actively-managed US equity funds
  - Monthly return data from CRSP
  - Gross returns computed by value-weighting the different shareclasses
  - Exclusion of funds with $15 mio.
    - Chen et al.(04), Pastor, Stambaugh, Taylor(15)
  - Sample period: January 1979 - December 2015

Data Description

- **Benchmark models**
  - Four-factor models
    - Market, size, value, momentum
  - Cremers Petajisto, Zitzewitz (CPZ) version
    - Market is the SP500
    - Size, value obtained from Russell indices
  - Rationale for CPZ model over Carhart model
    - Carhart model fails to price the Russel indices

      Ann. alpha of -2.4% for Russell 2000 (over 85-05)

Data Description

- ***Summary Statistics (value-weighted portfolio)***

Panel A: Gross Excess Return

	Average Nb. Funds	Mean (Ann.)	Volatility (Ann.)	Skewness	Kurtosis
All Funds	1103	7.68	15.01	-0.8	5.4

Panel B: Estimated Betas

	Market	Size	Value	Momentum	Adj. R2
All Funds	0.93	0.26	-0.10	0.01	0.98
Investment Styles					
Small-cap	0.98	0.81	-0.29	0.05	0.97
Large-cap	0.94	0.15	-0.06	0.01	0.99
Growth	0.95	0.33	-0.27	0.03	0.97
Value	0.92	0.13	0.19	-0.01	0.98
Fund Characteristics					
Low Expense	0.92	0.21	-0.04	0.00	0.98
High Expense	0.93	0.42	-0.27	0.01	0.97
Low Turnover	0.91	0.20	0.04	-0.02	0.98
High Turnover	0.95	0.37	-0.28	0.08	0.97

Empirical Results

- ***First skill dimension: First-Dollar Alpha***

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Volatility (Ann.)	Skewness	Kurtosis	Negative Skill	Positive Skill	10%	90%
All Funds	3.5	3.6	6.2	5.8	12.5	87.5	-0.4	7.8
Investment Styles								
Small-cap	5.1	3.4	10.5	1.5	4.6	95.4	1.2	9.7
Large-cap	2.0	2.4	5.9	3.9	17.7	82.3	-0.8	4.9
Growth	3.6	4.1	5.6	5.1	15.5	84.5	-0.8	8.5
Value	3.4	3.8	6.4	4.3	15.3	84.7	-0.6	8.4
Fund Characteristics								
Low Expense	2.2	3.5	5.0	6.4	23.4	76.6	-1.7	6.1
High Expense	4.6	4.3	5.9	3.1	12.0	88.0	-0.4	9.9
Low Turnover	3.3	4.0	5.1	4.8	16.0	84.0	-0.9	8.2
High Turnover	3.8	5.0	5.5	5.4	17.2	82.8	-1.2	9.2

Empirical Results

- ***Second skill dimension: Size Coefficient***

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Volatility (Ann.)	Skewness	Kurtosis	Negative Skill	Positive Skill	10%	90%
All Funds	1.5	1.5	8.1	7.2	14.1	85.9	-0.2	3.4
Investment Styles								
Small-cap	2.0	1.5	8.2	3.7	9.1	90.9	0.1	4.1
Large-cap	1.0	1.0	7.4	2.7	17.1	82.9	-0.3	2.3
Growth	1.6	1.8	7.1	4.9	17.1	82.9	-0.5	3.9
Value	1.4	1.5	7.2	4.7	18.0	82.0	-0.4	3.3
Fund Characteristics								
Low Expense	0.9	1.3	2.9	13.2	24.4	75.6	-0.6	2.4
High Expense	1.9	2.0	7.8	4.1	14.5	85.5	-0.3	4.3
Low Turnover	1.3	1.7	3.7	8.0	19.8	80.2	-0.6	3.3
High Turnover	1.7	2.2	4.6	5.7	18.5	81.5	-0.7	4.2

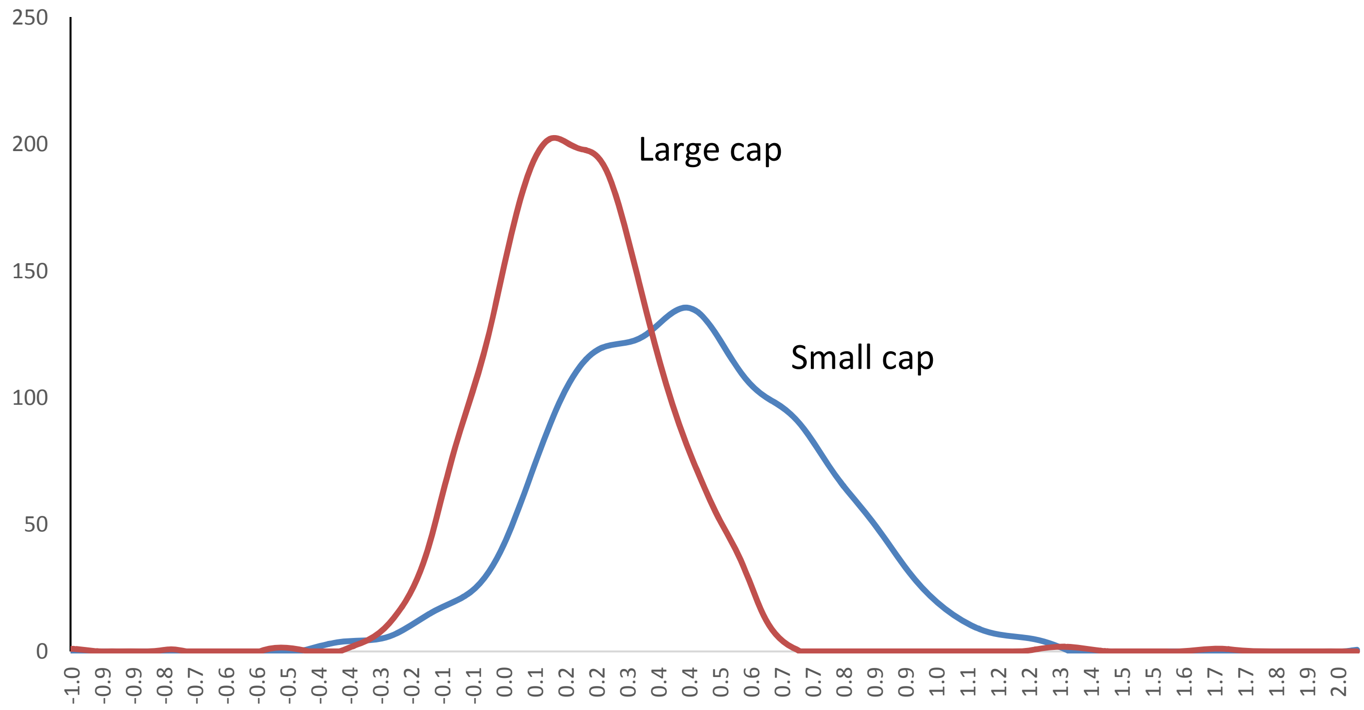
Empirical Results

- ***The two skill dimensions***
  - In short:
    - Strong evidence of investment skills
    - Strong evidence of capacity constraints
    - Large heterogeneity
    - Highly non-normal skill distributions
  - In addition:
    - The two skill measures are positively correlated

      High fd alpha comes with high size coefficient
    - Aggregation of the two dimensions is necessary to measure overall skill level

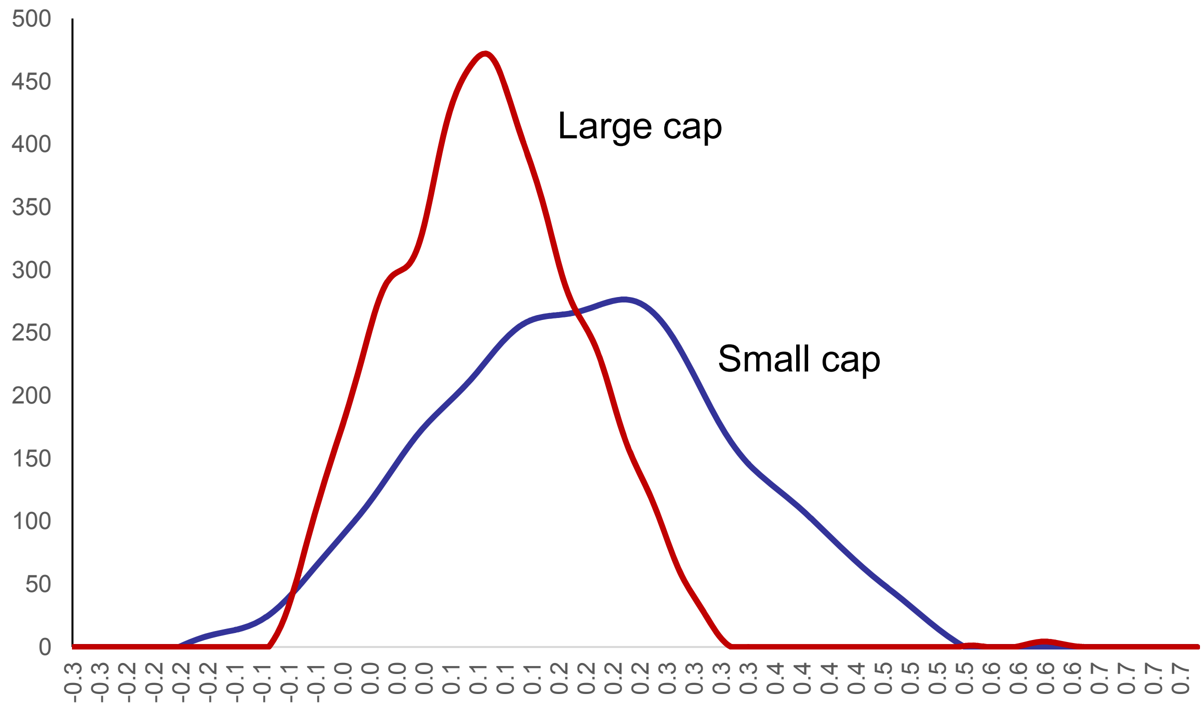
Empirical Results

- ***First Dollar Alpha***
  - Small cap versus large cap funds



Empirical Results

- ***Size Coefficient***
  - Small cap versus large cap funds



Empirical Results

- ***Aggregating skill: the value added (lifecycle)***

Panel A: Lifecycle Value Added

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Volatility (Ann.)	Skewness	Kurtosis	Negative Skill	Positive Skill	10%	90%
All Funds	1.9	9.3	8.8	14.5	35.9	64.1	-2.3	11.1
Investment Styles								
Small-cap	4.1	11.6	9.8	19.9	34.8	65.2	-2.9	15.4
Large-cap	0.0	7.8	5.5	8.8	47.8	52.2	-3.1	5.2
Growth	2.3	11.3	10.5	16.8	38.0	62.0	-3.4	14.0
Value	1.5	11.9	2.1	6.1	41.5	58.5	-3.4	8.9
Fund Characteristics								
Low Expense	5.5	21.2	9.8	9.2	30.8	69.2	-2.5	31.1
High Expense	1.3	7.7	1.6	6.7	39.7	60.3	-2.9	6.9
Low Turnover	5.8	19.1	8.9	2.5	12.4	87.6	-0.2	26.4
High Turnover	1.0	13.3	4.2	7.8	44.9	55.1	-5.9	10.9

Empirical Results

- ***Aggregating skill: the value added (steady state)***

Panel B: Steady State Value Added

	Moments				Proportions (%)		Quantiles (Ann.)	
	Mean (Ann.)	Volatility (Ann.)	Skewness	Kurtosis	Negative Skill	Positive Skill	10%	90%
All Funds	7.1	15.1	9.6	21.5	27.3	72.7	-1.7	25.6
Investment Styles								
Small-cap	8.4	14.2	8.0	13.7	17.8	82.2	-1.3	24.3
Large-cap	4.9	12.1	11.7	24.4	31.7	68.3	-1.6	21.0
Growth	8.8	19.4	10.1	20.8	29.2	70.8	-2.3	33.9
Value	6.7	14.3	7.0	10.8	24.7	75.3	-1.5	25.4
Fund Characteristics								
Low Expense	10.3	30.8	8.3	16.2	32.5	67.5	-3.1	44.4
High Expense	5.5	10.3	9.0	13.5	23.0	77.0	-1.5	19.9
Low Turnover	12.9	28.5	8.5	12.9	16.5	83.5	-0.8	51.2
High Turnover	6.0	13.8	8.2	9.4	26.5	73.5	-2.1	23.0

Empirical Results

- ***Aggregating skill: optimal value added***
  - Key prediction of neoclassical BG model:
    - Funds optimize their value added
  - Optimal value added:

$$va_i{}^* = \frac{(a_i)^2}{4b_i}$$

  - We can apply the non-parametric approach
    - Estimate the distribution of the optimal value added
    - Test the prediction of the model

Empirical Results

- ***Aggregating skill: optimal value added***

	Mean (Ann.)			Abs. Difference		Rel. Difference (%)	
	Optimal	Value Added Lifecycle	Value Added Steady State	Value Added Lifecycle	Value Added Steady State	Value Added Lifecycle	Value Added Steady State
All Funds	13.4	1.3	9.0	12.2	4.5	9.5	66.8
Investment Styles							
Small-cap	12.7	3.1	9.1	9.6	3.6	24.2	71.5
Large-cap	12.7	-0.8	6.9	13.4	5.8	-6.2	54.3
Growth	15.0	0.9	11.0	14.1	4.1	6.1	73.0
Value	14.3	0.9	9.2	13.4	5.1	6.1	64.1
Fund Characteristics							
Low Expense	23.5	4.9	14.7	18.6	8.8	20.8	62.6
High Expense	10.4	0.6	6.9	9.8	3.5	5.6	66.6
Low Turnover	27.3	4.8	16.7	22.5	10.6	17.7	61.3
High Turnover	14.2	-0.9	7.6	15.1	6.6	-6.3	53.7

Empirical Results

- ***The value added***
  - In short:
    - Funds generate substantial profits
    - Large heterogeneity
    - Highly non-normal skill distributions
  - In addition:
    - Profits are not far from optimal levels once funds reach their average size
    - Consistent with prediction of the neoclassical BG model

The Skill Measures

- ***Quid gross alpha?***
  - – Widely used as a skill measure in previous work
  - – Issue: it does not control for differences in size
    - Fund 1 more skilled than fund 2 on all dimensions…

      Yet they could have the same alphas if size 1>size 2
  - – Neoclassical theory: gross alpha=fees

    So gross alpha related to skill (a_i,b_i,va_i) only if funds

    choose specific fees (i.e., specific fund sizes)

The Skill Measures

- ***Quid gross alpha?***
  - Empirically, we find that this is not the case
    - Funds like to operate at a larger size
    - Thus, some of them charge “low” fees
  - This implies that the gross alpha is not very informative about skill
    - Moderate correlation with fd alpha
    - Zero correlation with size coefficient
    - Zero correlation with value added

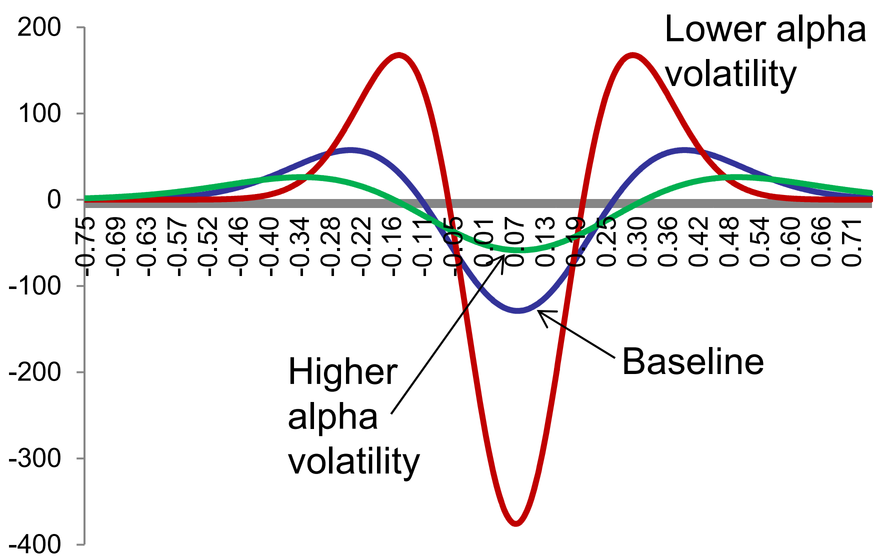
Conclusion

- **In this paper, we examine mutual fund skill**
  - Four measures (skill dimensions + value added)
- **We apply a novel non-parametric approach**
  - No parametric specification + joint skill measurement
- **Overall results**
  - Managers are skilled at detecting profitable trades… but unskilled at overriding capacity constraints
  - Overall, the fund profits from exploiting skill are large and close to the optimal level in the long run
  - Funds have skills and bargaining power

Additional Material

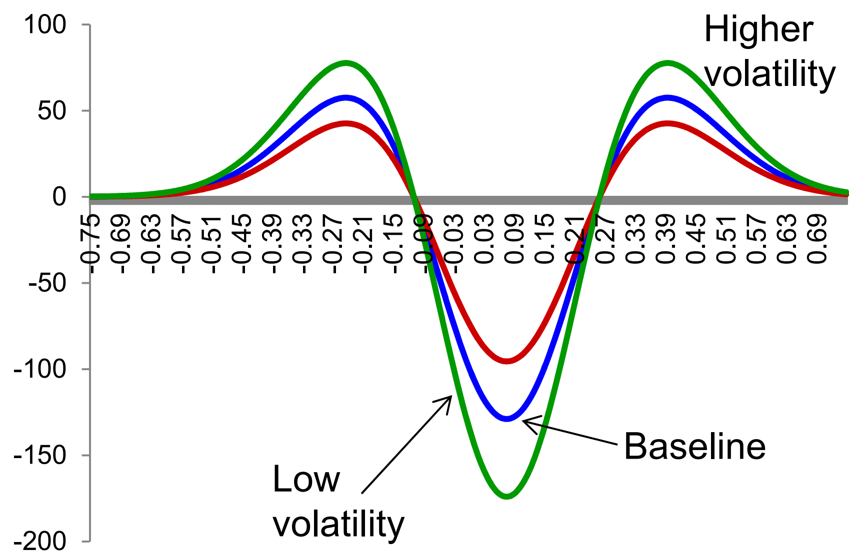
- **Comparative static analysis**

Change in volatility σ_α of α (+/- 30%)



Intuition: σ_α increases rel. importance of α over noise increases

Change in average variance σ_s of $\hat{\alpha}$ (+/- 30%)



Intuition: σ_s increases rel. importance of noise over α increases

Additional Material

- **Relations between the skill measures**

  – Neoclassical model of Berk and Green (2004)

  (i) skilled managers maximize profits

  (ii) investors compete for performance (α=fees(fe))

  – Total revenue (r_i): $a_i q_i$

  – Total cost (c_i): $b_i(q_i)^2$

  – Optimal size: $q_i^* = a_i/2b_i$ obtained with fees=fe^*

  – Value added or total profits (va_i):

$$r_i^* - c_i^* = a_i q_i^* - b_i(q_i^*)^2 = (a_i)^2/4b_i$$

Additional Material

– Key insight from the model: fees are irrelevant

- Low fees: additional flows q_i-q_i^* are passively indexed so that r_i^* -c_i^* is unchanged
- High fees: sell index short and invest the proceeds q_i^* -q_i in the fund so that r_i^* -c_i^* is unchanged

– However, fees affect the fund size and thus the relations between skill measures

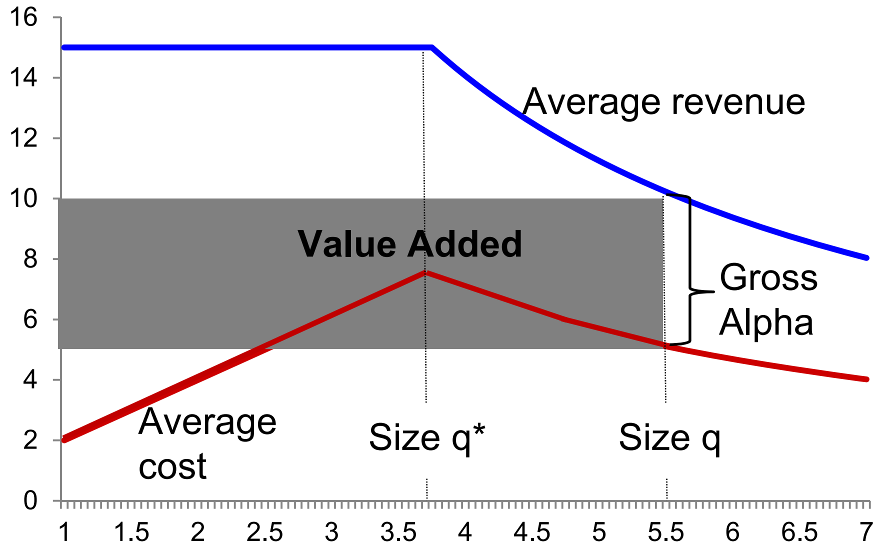
– Let us look at four different compensation schemes

Additional Material

Compensation Scheme	Scheme I	Scheme II	Scheme III	Scheme IV
Managers set fees such that	The size is optimal q_i^*	The size is optimal[2] $(q_i^*)^2$	The size is the same q	The size is arbitrary q_i^a
Gross alpha	$(r_i^* - c_i^*)/q_i^*$ $=a_i/2$	$(r_i^* - c_i^*)/(q_i^*)^2$ $=b_i$	$(r_i^* - c_i^*)/q$ $=(a_i)^2/4qb_i$	$(r_i^* - c_i^*)/q_i^a$ $=(a_i)^2/4q_i^a b_i$
Does the Gross Alpha Measure Skill?	YES	YES	YES	NO
	First-Dollar Alpha (1st dimension)	Size Coefficient (2nd dimension)	Value Added	
Predictions for Size/Fees	Moderate cross-fund variation in fees/size	Huge size Tiny fees	Same Size for all funds	Large cross-fund variation in size

Additional Material

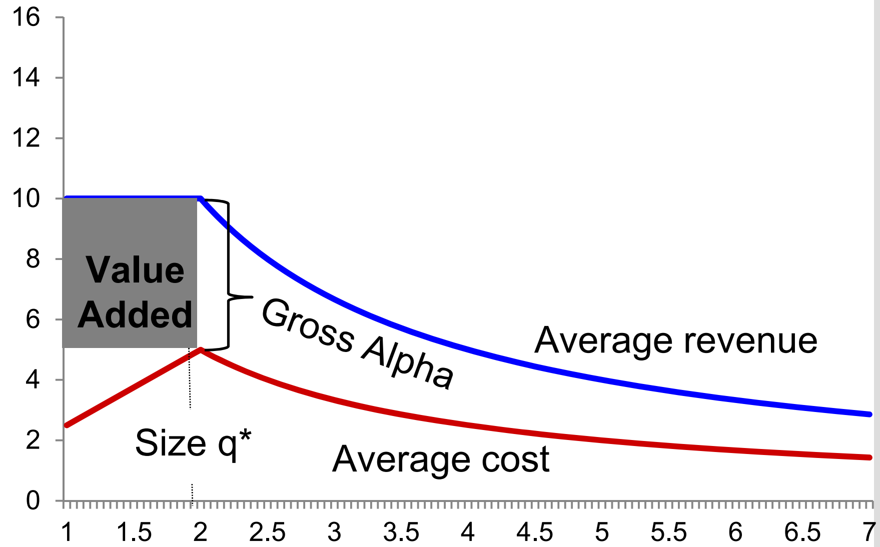
Manager 1
(choose low fees=fe)



Fees are low, so money flows until size q is reached:

- alpha=fe
- q^* is actively managed
- $q-q^*$ is passively managed

Manager 2
(same as manager 1 (fe))



Here again, we have alpha=fe
But this condition is achieved at a smaller size q^*