Man vs. Machine:

Quantitative and Discretionary Equity Management

Simona Abis *

September 14, 2022

Abstract

I use machine learning to categorize US active equity mutual funds as quants (reliant on computer models and fixed-rules) or discretionaries (reliant on human judgment). I then formulate hypotheses of how their skills, holdings and returns might differ, based on the conjecture that quants might have more learning capacity but less flexibility to adapt to changing market conditions than discretionaries. Consistent with those hypotheses, I find that quants hold more stocks, specialize in stock picking, and engage in more overcrowded trades. Discretionaries hold lesser-known stocks, switch between picking and timing and display higher active returns than quants in recessions.

Keywords: Investment Management, Quantitative Mutual Funds, Machine Learning, Rational Inattention.

*Columbia University, 665 W 130th St, New York, NY 10027. author e-mail: simona.abis@gsb.columbia.edu. I wish to thank Joël Peress for his support and invaluable advice. I am grateful for the very helpful comments from Lily Fang, Thierry Foucault, Sergei Glebkin, Jillian Grennan, Denis Gromb, Marcin Kacperczyk, John Kuong, Hugues Langlois, Yiming Ma, Harry Mamaysky, Giorgia Piacentino, Alberto Rossi, Laura Veldkamp, Daniel Wolfenzon, Rafael Zambrana and Qi Zeng. I also thank seminar participants at Bocconi University, Boston College, Columbia GSB, Georgetown McDonough, Harvard Business School, Hong Kong University of Science and Technology, Imperial College Business School, INSEAD brownbags, London Business School, London School of Economics, Rice Jones, Stanford GSB, UCLA Anderson, U. of Chicago Booth, U. of Illinois, U. of Texas Austin McCombs, U. of Virginia Darden, Washington U. in St. Louis Olin; and conference attendees at the AQR Institute Academic Symposium at LBS, FIRS The Financial Intermediation Research Society, HEC PhD Workshop, INSEAD-Wharton Consortium, LabEx PhD consortium, the Maryland Junior Finance Conference, the Nova-BPI Asset Management Conference and the WAPFIN conference. I am responsible for any errors.

INTERNET APPENDIX

A Training Sample

The following section contains details regarding the collection of the pre-classified PIS sample utilized for training the random forest algorithm to recognize quantitative or discretionary strategy descriptions. The sample classifications were collected through IRB # AAAS3044, which contains the full procedure utilized in selecting classifiers, explaining the task at hand and handling the collected data.

A.1 Expert Selection

Students interested in participating in the training sample labeling had to provide answers to the following list of questions:

1. Which degree are you currently enrolled in?
2. How many Finance courses have you taken in your Academic career?
   - List the most relevant courses
3. Have you ever participated in a Research Project in Finance?
   - If yes, what project (if more than one, mention the most relevant)?
4. Have you ever worked in the Finance industry for an investment company or another related job?
   - If yes, what was your role (if more than one, mention the most relevant)?
5. Are you an English language native speaker?
6. If not, please answer the following questions:
   - Have you completed a degree in an English-speaking country?
     - If yes, what type of degree (e.g. High School, Bachelor, Masters, ...)?

- If not, have you studied for at least 3 years in an English-speaking country? What type of degree were you studying for (e.g. High School, Bachelor, Masters, ...)?

- Have you ever worked in an English-speaking country?
  - If yes, for how long?

Of the 12 selected studetns, 7 had direct experience in asset management and 6 had prior work experience as Research Assistants in academic projects in Finance. The students' prior work experiences included: investment banking, quantitative research, data science, investment analysis. The list of programs that the students were attending at the time of the classification includes: Statistics, Data Science, Financial Engineering, Financial Economics, Business Analytics.

B Classification

B.1 Random Forest Algorithm

The random forest is a type of decision tree classifier, specifically it is an ensemble of decision trees, whose final classification is decided by majority voting.

Binary decision tree classifiers consist in breaking down the data into subsets by asking a series of questions inferred from the training sample. More specifically at the root of the tree the algorithm computes the information gain obtained by splitting the data according to each feature in the features matrix. The feature that determines the highest information gain is chosen and the training sample is split into two sub-samples according to the chosen feature. The second iteration includes two nodes, for each of the two nodes the same procedure is repeated and so on iteratively until all samples at each node belong to the same class. In order to avoid over-fitting a limit to the maximum depth of the tree is usually set – this is referred to as pruning. In a binary tree the information gain which is maximized at every split is defined as the difference in impurity between the parent node and the two children nodes:

$$IG(Node_{parent}, f) = I(Node_{parent}) - \frac{N_{left}}{N_p} I(Node_{leftchild}) - \frac{N_{left}}{N_p} I(Node_{rightchild})$$

where $IG(Node_{parent}, f)$ is the information gain at the parent node for feature f and N_p, N_{right} and N_{left} are the number of samples in the parent node, the right and the left children nodes respectively.

The impurity measure $I(node)$ used in this paper is entropy:

$$I(node) = -\sum_{i=1}^{class} p(i|node) log_2 p(i|node)$$

where $p(i|node)$ is the proportion of samples in the node that belong to class i – hence entropy is 0 when all samples in a node belong to the same class and it is maximum when samples are uniformly distributed across classes.

Figure B.1 represents graphically a single decision tree pruned at 3 branches. The representation was created utilizing the pre-classified samples belonging to the training sample.[1]

When utilizing a single decision tree, strong over-fitting issues are generally encountered. A way to minimize these issues is to employ a random forest.

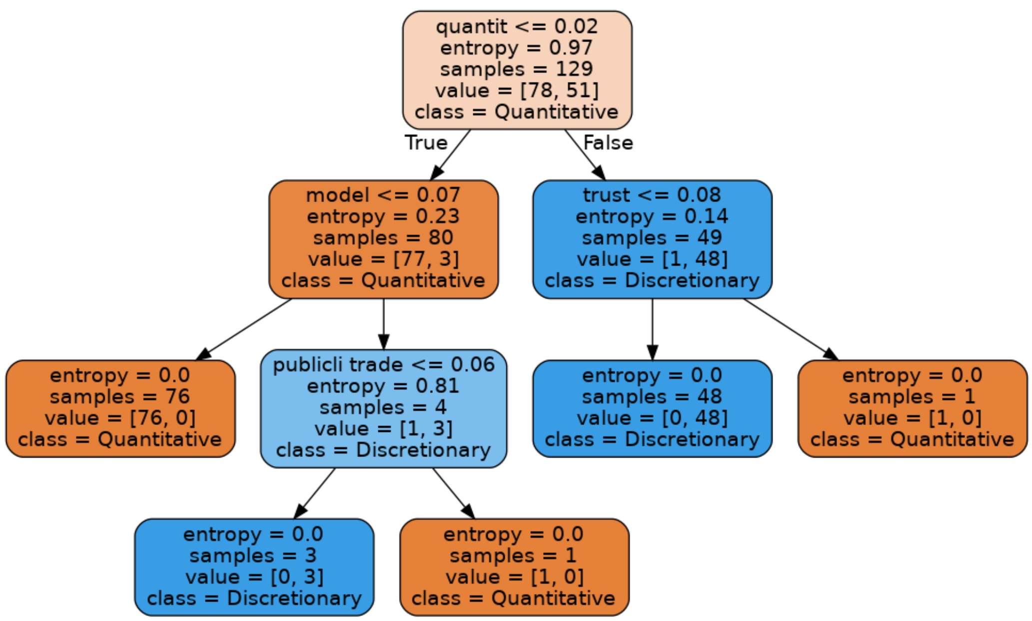
The random forest algorithm can be described in four steps:

1. Randomly choose a sample of size n from the training sample with replacement
2. Grow a decision tree from the selected sample. At each node:
   (a) Randomly select f features from the features matrix without replacement
   (b) Among the selected subset of features (f) choose the one that maximizes the information gain and split the node according to that feature
3. Repeat the above 2 steps k times and record the classification of each item in the training sample by each of the k trees
4. Use majority voting to assign a final classification to each item in the training sample

The key parameter to be chosen is the number of trees in the random forest (k), the larger the number the better the prediction but the higher the computational burden. In this paper $1,000$ trees were used.

[1] This was created for illustration purposes only. This tree was not used in creating the classification.

Figure B.1: **Single decision tree example:** Graphical illustration of the decision making process of a single decision tree. The tree was constructed using the 129 pre-classified sections utilized in training the random forest algorithm (78 pre-classified as *Discretionary* and 51 pre-classified as *Quantitative*)—i.e. it excludes the pre-classified sections belonging to the test sample. Each box displays the stemmed feature (word or bi-gram) with the highest entropy, on the bases of which splitting is decided. Each box additionally displays the computed entropy (to be minimized), the number of total samples, and the number of samples assigned a quantitative or discretionary label. The tree is pruned to have a depth of 3. This tree is just for illustration purposes. This specific tree was not utilized in the creation of the classification utilized in the paper.

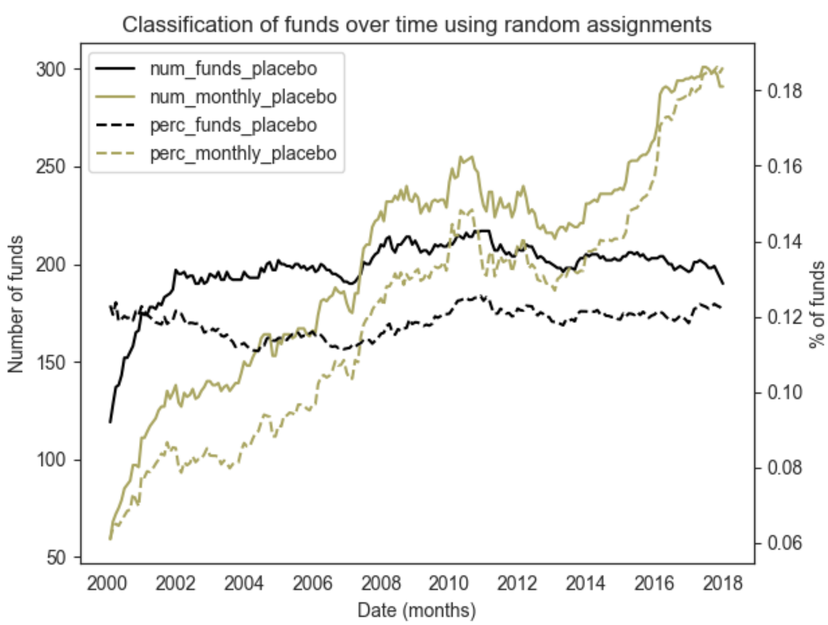


C Placebo Classifications

C.1 Construction

The first placebo (*funds_ placebo*) takes as input the overall percentage of observations classified as being quantitative (12.15%), then randomly selects 12.15% of funds from the overall sample (386 funds) and it permanently assigns a value of 1 to those funds (0 otherwise). The second (*months_ placebo*) is constructed by randomly assigning a value of 1 every month to as many observations as those classified to belong to quantitative funds in that month (0 otherwise). Figure C.1 shows both the number of funds and the percentage of funds assigned to the two placebo classifications over time.

Figure C.1: **Placebo classifications:** Number of funds (left index) and percentage of funds (right index) belonging to the two placebo classifications over time.



C.2 Regressions

I test all model predictions using the two placebo classifications instead of the quantitative classification constructed using the random forest algorithm. All results are included next.

Table C.1: **Age, size and style:** Dependent variables: funds' age ($ln(Age)$), size ($ln(TNA)$), turnover ratio ($Turnover$), the amount of cash held in percentage of TNA ($Cash$), and style ($Market$, $Size$, $Value$, Mom, $Invest.$ and $Profit.$). Independent variables: funds_placebo (F – Panel 1) and months_placebo (M – Panel 2) dummies and their interaction with a dummy identifying NBER recessions periods ($Recess$); expense ratio ($Expenses$); turnover ratio ($Turnover$); growth in net fund flows ($FlowsGrowth$); the volatility in net fund flows growth ($FlowsVol$); total fund loads ($Loads$); fund style, size and age as listed above. Variables construction is detailed in Table 1; summary statistics are reported in Table 2 in the paper's body. Control variables have been omitted for brevity. Each regression excludes the control variable corresponding to the dependent variable. All regressions include month fixed–effects (FE); standard errors are clustered at the month and fund level (Cl).

	ln(Age) (1)	ln(TNA) (2)	Turnover (3)	Cash (4)	Market (5)	Size (6)	Value (7)	Mom (8)	Invest. (9)	Profit. (10)
Constant	3.97***	3.78***	69.39***	3.51***	0.74***	-0.16***	0.02	0.12***	0.21***	0.11***
	(54.36)	(16.95)	(8.68)	(9.13)	(98.60)	(-3.96)	(0.82)	(5.37)	(5.02)	(4.28)
F	0.03	-0.11	-1.30	0.08	-0.00	0.03*	-0.00	0.00	-0.00	-0.00
	(0.95)	(-1.24)	(-0.46)	(0.57)	(-0.75)	(1.69)	(-0.65)	(0.32)	(-0.49)	(-0.55)
F X Recess	0.02	-0.04	-2.13	0.07	-0.00	0.01	-0.00	0.00	-0.01	0.00
	(0.76)	(-1.06)	(-0.98)	(0.41)	(-0.77)	(1.06)	(-0.28)	(0.39)	(-0.95)	(0.13)
Adjusted R2	0.28	0.31	0.12	0.08	0.94	0.15	0.35	0.18	0.16	0.30
Obs	334,821	334,821	334,821	317,247	334,821	334,821	334,821	334,821	334,821	334,821
Constant	3.98***	3.77***	69.12***	3.53***	0.74***	-0.16***	0.02	0.12***	0.21***	0.11***
	(54.72)	(16.90)	(8.63)	(9.15)	(98.11)	(-3.86)	(0.79)	(5.39)	(5.00)	(4.28)
M	-0.00	-0.00	0.37	-0.01	-0.00	0.00	0.00	-0.00**	-0.00	-0.00
	(-0.24)	(-0.05)	(1.08)	(-0.46)	(-0.52)	(0.61)	(0.64)	(-2.59)	(-0.19)	(-0.67)
M X Recess	-0.00	-0.02	-0.25	-0.00	0.00*	-0.01*	0.00	0.00	0.00	-0.00
	(-0.30)	(-0.91)	(-0.21)	(-0.04)	(1.67)	(-1.69)	(0.81)	(0.84)	(0.39)	(-1.34)
Adjusted R2	0.28	0.31	0.12	0.08	0.94	0.15	0.35	0.18	0.16	0.30
Obs	334,821	334,821	334,821	317,247	334,821	334,821	334,821	334,821	334,821	334,821

t statistics in parentheses; * p<0.10, ** p<0.05, *** p<0.01

Table C.2: **Flows and Fees:** Dependent variables: growth in net fund flows (*FlowGrowth*), the volatility of net fund flows growth (*FlowVol*), funds' expense ratio (*Expenses*), management fees (*Mgmt*), Actual 12b1 fees (12*b*1) and fund loads (*Loads*). Variables construction is detailed in Table 1; summary statistics are reported in Table 2. Independent variables and all regression specifications are the same as described in Table C.1. Control variables are omitted for brevity.

	Flow		*Fees*			
	Growth	*Vol*	*Expenses*	*Mgmt*	12*b*1	*Loads*
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	2.700***	0.135***	1.630***	0.632***	0.478***	-0.025***
	(13.48)	(29.18)	(41.36)	(16.96)	(13.12)	(-16.64)
F	-0.007	0.001	-0.012	-0.003	-0.014	-0.000
	(-0.11)	(0.74)	(-0.74)	(-0.22)	(-0.97)	(-0.37)
F X Recess	-0.069	-0.000	0.002	0.018*	-0.002	0.000
	(-0.51)	(-0.10)	(0.22)	(1.82)	(-0.60)	(0.23)
Adjusted R2	0.06	0.09	0.35	0.07	0.31	0.19
Obs	334,821	334,821	334,821	334,481	228,769	334,821
Constant	2.698***	0.135***	1.629***	0.632***	0.476***	-0.025***
	(13.48)	(29.17)	(41.25)	(16.97)	(13.09)	(-16.68)
M	-0.006	-0.000	-0.003	-0.002	-0.002	0.000
	(-0.23)	(-0.25)	(-1.45)	(-1.46)	(-1.53)	(0.66)
M X Recess	0.004	0.000	0.001	0.002	0.003	-0.000
	(0.08)	(0.13)	(0.18)	(0.30)	(0.96)	(-1.52)
Adjusted R2	0.06	0.09	0.35	0.07	0.31	0.19
Obs	334,821	334,821	334,821	334,481	228,769	334,821

t statistics in parentheses; * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table C.3: **Prediction 1: High Ability Funds—funds placebo:** Graphical representation of regression coefficients. Dependent variables: macro-timing in recessions with respect to innovations in industrial production ($IndPro$ (R), Panel 1) or non-farm payrolls ($NFPay$ (R), Panel 2); characteristic-timing in recessions ($DGTW$ (R), Panel 3); stock picking in recessions ($Picking$ (R), Panel 4); characteristic-timing in expansions ($DGTW$ (E), Panel 5); and macro-timing in expansions with respect to innovations in industrial production ($IndPro$ (E), Panel 6) or non-farm payrolls ($NFPay$ (E), Panel 7). Independent variables: a dummy identifying the funds_placebo classification ($F\_placebo$), dummy variables identifying the top q% of funds with the highest timing ability in the previous recession ($TopT\_R$) or picking ability in the previous expansion ($TopP\_E$); and the interaction between the $F\_placebo$ dummy (in interactions abbreviated with: F) and the high-ability dummies ($TopP\_E \times F$, $TopT\_R \times F$). Variables construction is detailed in Table **??**; summary statistics are reported in Table **??**. Coefficients are estimated using high-ability cut-offs of $q = 10\%$ to $q = 25\%$. Bars indicate 95% confidence intervals. Control variables and all regression specifications are the same as described in Table C.1. Control variables are omitted for brevity.

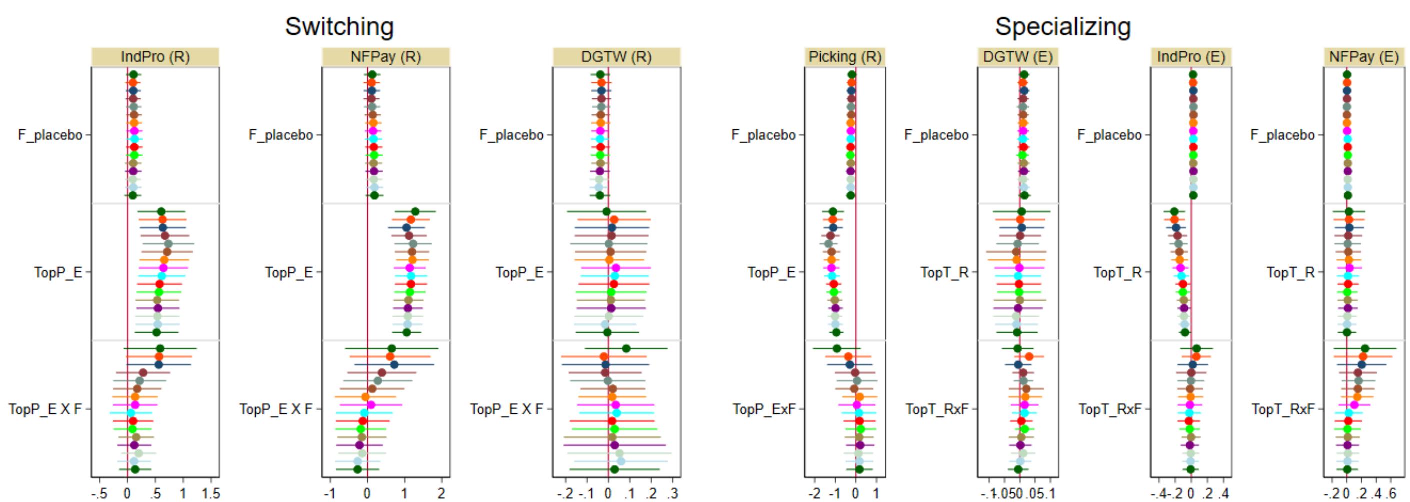


Table C.4: **Prediction 1: High Ability Funds—months placebo:** Graphical representation of regression coefficients. Dependent variables: macro-timing in recessions with respect to innovations in industrial production ($IndPro$ (R), Panel 1) or non-farm payrolls ($NFPay$ (R), Panel 2); characteristic-timing in recessions ($DGTW$ (R), Panel 3); stock picking in recessions ($Picking$ (R), Panel 4); characteristic-timing in expansions ($DGTW$ (E), Panel 5); and macro-timing in expansions with respect to innovations in industrial production ($IndPro$ (E), Panel 6) or non-farm payrolls ($NFPay$ (E), Panel 7). Independent variables: a dummy identifying the months_placebo classification ($M\_placebo$), dummy variables identifying the top q% of funds with the highest timing ability in the previous recession ($TopT\_R$) or picking ability in the previous expansion ($TopP\_E$); and the interaction between the $M\_placebo$ dummy (in interactions abbreviated with: M) and the high-ability dummies ($TopP\_E \times M$, $TopT\_R \times M$). Variables construction is detailed in Table **??**; summary statistics are reported in Table **??**. Coefficients are estimated using high-ability cut-offs of $q = 10\%$ to $q = 25\%$. Bars indicate 95% confidence intervals. Control variables and all regression specifications are the same as described in Table C.1. Control variables are omitted for brevity.

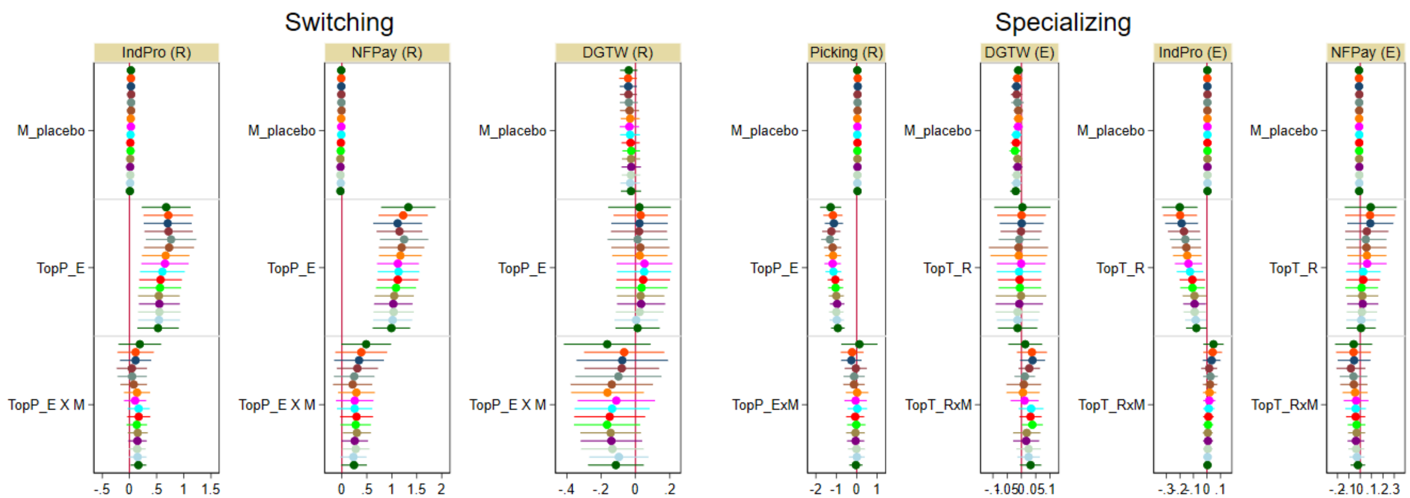


Table C.5: **Prediction 1: Average Picking and Timing Abilities:** Dependent variables: stock picking (SP_SUE), macro-timing (MT_IndPro, MT_NFPay) and characteristics-timing (CT_DGTW) abilities. Variables construction is detailed in Table 1; summary statistics are reported in Table 2 in the paper's body. Independent variables and all regression specifications are the same as described in Table C.1. Control variables are omitted for brevity.

	SP_SUE (1)	MT_IndPro (2)	MT_NFPay (3)	CT_DGTW (4)
Constant	0.0618	0.2799***	-0.1038	-0.0998
	(0.36)	(3.18)	(-0.81)	(-0.80)
F	-0.0150	0.0185	0.0158	0.0117
	(-0.61)	(1.03)	(0.64)	(1.15)
F X Recess	-0.2153*	0.1513*	0.1607	-0.0289
	(-1.70)	(1.96)	(1.36)	(-1.25)
Adjusted R2	0.18	0.37	0.38	0.57
Obs	334,374	293,325	293,325	334,821
Constant	0.0541	0.2845***	-0.0974	-0.0969
	(0.32)	(3.22)	(-0.76)	(-0.77)
M	0.0106	0.0058	-0.0123*	-0.0102
	(0.81)	(1.06)	(-1.72)	(-1.09)
M X Recess	0.0043	0.0279	0.0363	-0.0379
	(0.09)	(0.95)	(0.90)	(-1.51)
Adjusted R2	0.18	0.37	0.38	0.57
Obs	334,374	293,325	293,325	334,821

t statistics in parentheses; * p<0.10, ** p<0.05, *** p<0.0

Table C.6: **Prediction 2: Portfolio Diversification:** Dependent variables: the natural log of the number of stocks held ($ln(NStocks)$), and of the volatility of fund returns ($ln(RetVol)$). Independent variables and all regression specifications are the same as described in Table C.1; control variables are omitted for brevity. Model (6) also controls for the percentage of cash held ($Cash$). Variables construction is detailed in Table 1; summary statistics are reported in Table 2 in the paper's body.

	$ln(NStocks)$		$ln(RetVol)$	
			$24m$	$36m$
	(1)	(2)	(3)	(4)
Constant	4.581***	4.695***	1.234***	1.291***
	(43.45)	(43.08)	(50.70)	(49.51)
F	0.013	0.018	0.012*	0.010
	(0.30)	(0.41)	(1.87)	(1.49)
F X Recess	-0.055***	-0.049**	-0.009	-0.007
	(-2.88)	(-2.30)	(-1.15)	(-1.00)
Cash %		-0.018***		
		(-8.22)		
Adjusted R2	0.18	0.19	0.77	0.77
Obs	334,821	317,247	307,996	287,174
Constant	4.583***	4.698***	1.236***	1.293***
	(43.42)	(43.07)	(50.55)	(49.31)
M	-0.004	-0.005	-0.000	0.000
	(-1.10)	(-1.34)	(-0.24)	(0.23)
M X Recess	0.001	0.000	-0.005	-0.005
	(0.08)	(0.03)	(-1.52)	(-1.35)
Cash %		-0.018***		
		(-8.22)		
Adjusted R2	0.18	0.19	0.77	0.77
Obs	334,821	317,247	307,996	287,174

t statistics in parentheses; * p<0.10, ** p<0.05, *** p<0.01

Table C.7: **Prediction 3: Information Gap—funds placebo:** Dependent variables: TNA-weighted average of the natural logarithm of market capitalization ($ln(MktCap)$), of age in months (Age), of the natural logarithm of monthly media mentions ($ln(News)$), and of the number of analysts following the stocks held ($Analysts$). Independent variables are the same as described in Table C.1, plus the illiquidity of stocks held—measured using Amihud ratio—($Illiquidity$). The key variable of interest is the funds_placebo (F). Control variables are omitted for brevity. Models (4), (6), and (8) additionally control for $ln(MktCap)$. Model (1) excludes from the control variables set the size style ($Size$). All regression specifications are the same as described in Table C.1. Variables construction is detailed in Table 1; summary statistics are reported in Table 2 of the paper's main body.

	Stock Characteristics							
	ln(MktCap)		*Age*		*ln(News)*		*Analysts*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	10.260***	9.561***	348.774***	-116.258***	3.884***	-1.877***	5.880***	-7.366***
	(39.36)	(68.61)	(27.71)	(-7.69)	(36.91)	(-16.66)	(20.30)	(-10.18)
F	-0.139	-0.040	-3.724	-1.798	-0.028	-0.004	-0.059	-0.004
	(-1.30)	(-0.79)	(-0.98)	(-0.56)	(-0.93)	(-0.30)	(-0.70)	(-0.08)
F X Recess	-0.060	-0.028	-5.404*	-4.053	-0.026	-0.009	0.003	0.042
	(-1.21)	(-0.82)	(-1.69)	(-1.51)	(-0.90)	(-0.48)	(0.04)	(0.64)
Illiquidity	-0.938***	-0.316***	7.374**	22.724***	-0.048**	0.142***	-0.424***	0.013
	(-3.07)	(-4.30)	(2.37)	(5.35)	(-2.58)	(3.50)	(-4.38)	(0.19)
Size		-3.987***	-200.759***	-6.821	-2.301***	0.103**	-6.355***	-0.831**
		(-126.47)	(-51.99)	(-1.33)	(-67.21)	(2.17)	(-26.15)	(-2.36)
Stock_Size				48.641***		0.603***		1.385***
				(42.39)		(61.18)		(18.97)
Adjusted R2	0.11	0.60	0.50	0.69	0.52	0.85	0.63	0.76
Obs	334,817	334,817	334,817	334,817	334,652	334,652	334,817	334,817

t statistics in parentheses; * p<0.10, ** p<0.05, *** p<0.0

Table C.8: **Prediction 3: Information Gap—months placebo:** Dependent variables: TNA-weighted average of the natural logarithm of market capitalization ($ln(MktCap)$), of age in months (Age), of the natural logarithm of monthly media mentions ($ln(News)$), and of the number of analysts following the stocks held ($Analysts$). Independent variables are the same as described in Table C.1, plus the illiquidity of stocks held—measured using Amihud ratio—($Illiquidity$). The key variable of interest is the months_placebo (M). Control variables are omitted for brevity. Models (4), (6), and (8) additionally control for $ln(MktCap)$. Model (1) excludes from the control variables set the size style ($Size$). All regression specifications are the same as described in Table C.1. Variables construction is detailed in Table 1; summary statistics are reported in Table 2 of the paper's main body.

	Stock Characteristics							
	ln(MktCap)		*Age*		*ln(News)*		*Analysts*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	10.239***	9.554***	348.210***	-116.574***	3.880***	-1.877***	5.871***	-7.365***
	(39.19)	(68.36)	(27.64)	(-7.71)	(36.83)	(-16.69)	(20.28)	(-10.18)
M	0.005	0.008	-0.605	-0.975**	-0.000	-0.005*	0.007	-0.004
	(0.53)	(1.25)	(-1.10)	(-2.10)	(-0.06)	(-1.76)	(0.45)	(-0.28)
M X Recess	0.021	-0.002	1.350	1.449	0.002	0.003	0.010	0.013
	(1.06)	(-0.11)	(0.91)	(1.20)	(0.17)	(0.51)	(0.29)	(0.57)
Illiquidity	-0.946***	-0.318***	7.139**	22.607***	-0.049***	0.142***	-0.427***	0.014
	(-3.06)	(-4.39)	(2.41)	(5.32)	(-2.71)	(3.48)	(-4.43)	(0.19)
Size		-3.988***	-200.883***	-6.849	-2.302***	0.103**	-6.357***	-0.831**
		(-126.75)	(-52.12)	(-1.33)	(-67.13)	(2.17)	(-26.16)	(-2.36)
Stock_Size				48.651***		0.603***		1.385***
				(42.40)		(61.18)		(18.97)
Adjusted R2	0.11	0.60	0.50	0.69	0.52	0.85	0.63	0.76
Obs	334,817	334,817	334,817	334,817	334,652	334,652	334,817	334,817

t statistics in parentheses; * p<0.10, ** p<0.05, *** p<0.0

Table C.9: **Prediction 4: Commonality:** Dependent variables: holdings commonality ($Comm$), active commonality ($AComm$), and the natural log of dispersion ($Log(Disp)$). Independent variables and all regression specifications are the same as described in Table C.1. Control variables are omitted for brevity. Variables construction is detailed in Table 1; summary statistics are reported in Table 2 in the paper's body.

	$Comm$	$AComm$	$ln(Disp)$
	(1)	(2)	(3)
Constant	9.892***	3.463***	0.115
	(14.87)	(5.50)	(0.88)
F	-0.355*	-0.348**	0.030
	(-1.77)	(-2.34)	(0.63)
F X Recess	-0.093	-0.100	0.042
	(-0.52)	(-0.61)	(1.02)
R2	0.55	0.53	0.16
Obs	334,821	334,821	334,817
Constant	9.838***	3.408***	0.119
	(14.79)	(5.41)	(0.92)
M	0.007	0.016	0.002
	(0.28)	(0.70)	(0.45)
M X Recess	-0.038	-0.037	-0.001
	(-0.52)	(-0.67)	(-0.07)
R2	0.55	0.53	0.16
Obs	334,821	334,821	334,817

t statistics in parentheses; * p<0.10, ** p<0.05, *** p<0.01

Table C.10: **Prediction 5: Active Returns:** Dependent variables: 24–months (Models (1) and (2)) and 36–months (Models (3) and (4)) rolling Fama-French 6 AR, constructed monthly and averaged within a quarter. Independent variables are the same as in Table C.1, but they are lagged. Control variables are omitted for brevity. Variables construction is detailed in Table 1; summary statistics are reported in Table 2 in the paper's body. Models (1) and (3) utilize OLS regressions. Models (2) and (4) utilize quantile regressions for the 50^{th} quantile. All regressions are run at the quarterly frequency and include quarter fixed–effects (FE); standard errors are clustered at the quarter and fund level (Cl). For quantile regressions that is achieved through bootstrapping with 1000 repetitions.

	$\overline{AR} - ff6$			
	$24m$ OLS (1)	$24m$ $q50$ (2)	$36m$ OLS (3)	$36m$ $q50$ (4)
Constant	0.140*	1.100***	0.106	0.930***
	(1.93)	(20.22)	(1.45)	(17.12)
F	0.009	0.005	-0.002	-0.005
	(0.94)	(0.92)	(-0.22)	(-0.75)
F X Recess	-0.001	-0.012	0.009	-0.001
	(-0.02)	(-0.50)	(0.53)	(-0.03)
Obs	102,869	102,869	96,155	96,155
Constant	0.142*	1.105***	0.106	0.928***
	(1.97)	(20.23)	(1.45)	(17.07)
M	-0.000	0.001	0.005	-0.004
	(-0.04)	(0.18)	(0.71)	(-0.61)
M X Recess	-0.007	-0.003	-0.015	-0.019
	(-0.23)	(-0.11)	(-0.49)	(-0.57)
Obs	102,869	102,869	96,155	96,155

t statistics in parentheses; * p<0.10, ** p<0.05, *** p<0.01