Real-Time Predictability of Mutual Fund Performance Predictors

Yu Xia

McGill University

AFA 2022

Why mutual fund performance matters?

- Mutual funds are widely used
  - Actively managed funds hold 60% U.S. total net assets in equity
  - $\approx$ 50% U.S. households own mutual funds
- Hard to gauge their value added to investors
- Question:

Can investors gain from using predictors to select actively managed mutual funds in real time?

Why mutual fund performance matters?

- Mutual funds are widely used
  - Actively managed funds hold 60% U.S. total net assets in equity
  - $\approx$ 50% U.S. households own mutual funds
- Hard to gauge their value added to investors
- Question:

Can investors gain from using predictors to select actively managed mutual funds in real time?

Why mutual fund performance matters?

- Mutual funds are widely used
  - Actively managed funds hold 60% U.S. total net assets in equity
  - $\approx$ 50% U.S. households own mutual funds
- Hard to gauge their value added to investors
- Question:

Can investors gain from using predictors to select actively managed mutual funds in real time?

Main findings

- Short answer: Yes!
- Using two adaptive approaches to evaluating predictors in real time
  - Regression-based machine learning: $1.3 \sim 1.7\%$ p.a. real-time alphas
    - Li and Rossi (2020): ML based on stock holdings and stock characteristics.
    - DeMiguel et al. (2021): ML based on fund characteristics and performance.
    - My paper: fund characteristics, performance, and holding-based activeness.
  - Rule-based portfolio sorts: 2.5% p.a. real-time market-adjusted alpha
- Do investors react to predictive information? Yes!
  - Investor flow chases for predictive information
  - Reaction is generally stronger among more growth-oriented funds

Main findings

- Short answer: Yes!
- Using two adaptive approaches to evaluating predictors in real time
  - Regression-based machine learning: $1.3 \sim 1.7\%$ p.a. real-time alphas
    - Li and Rossi (2020): ML based on stock holdings and stock characteristics.
    - DeMiguel et al. (2021): ML based on fund characteristics and performance.
    - My paper: fund characteristics, performance, and holding-based activeness.
  - Rule-based portfolio sorts: 2.5% p.a. real-time market-adjusted alpha
- Do investors react to predictive information? Yes!
  - Investor flow chases for predictive information
  - Reaction is generally stronger among more growth-oriented funds

Main findings

- Short answer: Yes!
- Using two adaptive approaches to evaluating predictors in real time
  - Regression-based machine learning: $1.3 \sim 1.7\%$ p.a. real-time alphas
    - ★ Li and Rossi (2020): ML based on stock holdings and stock characteristics.
    - ★ DeMiguel et al. (2021): ML based on fund characteristics and performance.
    - ★ My paper: fund characteristics, performance, and holding-based activeness.
  - Rule-based portfolio sorts: 2.5% p.a. real-time market-adjusted alpha
- Do investors react to predictive information? Yes!
  - Investor flow chases for predictive information
  - Reaction is generally stronger among more growth-oriented funds

Main findings

- Short answer: Yes!
- Using two adaptive approaches to evaluating predictors in real time
  - Regression-based machine learning: $1.3 \sim 1.7\%$ p.a. real-time alphas
    - Li and Rossi (2020): ML based on stock holdings and stock characteristics.
    - DeMiguel et al. (2021): ML based on fund characteristics and performance.
    - My paper: fund characteristics, performance, and holding-based activeness.
  - Rule-based portfolio sorts: 2.5% p.a. real-time market-adjusted alpha
- Do investors react to predictive information? Yes!
  - Investor flow chases for predictive information
  - Reaction is generally stronger among more growth-oriented funds

Main findings

- Short answer: Yes!
- Using two adaptive approaches to evaluating predictors in real time
  - Regression-based machine learning: $1.3 \sim 1.7\%$ p.a. real-time alphas
    - ★ Li and Rossi (2020): ML based on stock holdings and stock characteristics.
    - ★ DeMiguel et al. (2021): ML based on fund characteristics and performance.
    - ★ My paper: fund characteristics, performance, and holding-based activeness.
  - Rule-based portfolio sorts: 2.5% p.a. real-time market-adjusted alpha
- Do investors react to predictive information? Yes!
  - Investor flow chases for predictive information
  - Reaction is generally stronger among more growth-oriented funds

Main findings

- Short answer: Yes!
- Using two adaptive approaches to evaluating predictors in real time
  - Regression-based machine learning: $1.3 \sim 1.7\%$ p.a. real-time alphas
    - ★ Li and Rossi (2020): ML based on stock holdings and stock characteristics.
    - ★ DeMiguel et al. (2021): ML based on fund characteristics and performance.
    - ★ My paper: fund characteristics, performance, and holding-based activeness.
  - Rule-based portfolio sorts: 2.5% p.a. real-time market-adjusted alpha
- Do investors react to predictive information? Yes!
  - Investor flow chases for predictive information
  - Reaction is generally stronger among more growth-oriented funds

List of fund performance predictors

Category	Predictor	Study
Characteristics-Based	Expense Ratio (ER)	Elton et al. (1993)
	Fund Flow (Flow)	Zheng (1999)
	Fund Size (Size)	Chen et al. (2004)
Performance-Based	One-Year Return (Ret1y)	Hendricks et al. (1993)
	Carhart Alpha (Car1y)	Carhart (1997)
	One-Month Return (Ret1m)	Bollen & Busse (2004)
	Return Gap (RG)	Kacperczyk et al. (2006)
Activeness	Turnover (TR)	Elton et al. (1993)
	Active Share (AS)	Cremers & Petajisto (2009)
	R-Squared (R^2)	Amihud & Goyenko (2013)
	Active Weight (AW)	Doshi et al. (2015)
	Fund Duration (Dur)	Cremers & Petajisto (2016)

Methodology 1: Machine learning

- Non-ML benchmark: OLS with objective function

$$\min_{\theta} \mathcal{L}(\theta) = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (\underbrace{r_{i,t+1}}_{\text{Net Fund Return}} - \underbrace{x'_{i,t}}_{\text{Predictors}} \theta)^2$$

- ML: Balance between fit & robustness; allow real-time selection
- Shrinkage/sparsity/both: Ridge, LASSO, elastic net, sparse group LASSO

$$\min_{\theta} \mathcal{L}(\theta; \cdot) = \mathcal{L}(\theta) + \phi(\theta; \cdot)$$

  where $\phi(\theta; \cdot)$ is a penalty term.
- Dimension reduction: PCR, PLS
  - OLS after transforming & reducing predictor space to principal components

Methodology 1: Machine learning

- Non-ML benchmark: OLS with objective function

$$\min_{\theta} \mathcal{L}(\theta) = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (\underbrace{r_{i,t+1}}_{\text{Net Fund Return}} - \underbrace{x_{i,t}'}_{\text{Predictors}} \theta)^2$$

- ML: Balance between fit & robustness; allow real-time selection
- Shrinkage/sparsity/both: Ridge, LASSO, elastic net, sparse group LASSO

$$\min_{\theta} \mathcal{L}(\theta; \cdot) = \mathcal{L}(\theta) + \phi(\theta; \cdot)$$

  where $\phi(\theta; \cdot)$ is a penalty term.
- Dimension reduction: PCR, PLS
  - OLS after transforming & reducing predictor space to principal components

Methodology 1: Machine learning

- Non-ML benchmark: OLS with objective function

$$\min_{\theta} \mathcal{L}(\theta) = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (\underbrace{r_{i,t+1}}_{\text{Net Fund Return}} - \underbrace{x_{i,t}^{\prime}}_{\text{Predictors}} \theta)^2$$

- ML: Balance between fit & robustness; allow real-time selection
- Shrinkage/sparsity/both: Ridge, LASSO, elastic net, sparse group LASSO

$$\min_{\theta} \mathcal{L}(\theta; \cdot) = \mathcal{L}(\theta) + \phi(\theta; \cdot)$$

where $\phi(\theta; \cdot)$ is a penalty term.

- Dimension reduction: PCR, PLS
  - OLS after transforming & reducing predictor space to principal components

Methodology 1: Machine learning

- Non-ML benchmark: OLS with objective function

$$\min_{\theta} \mathcal{L}(\theta) = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (\underbrace{r_{i,t+1}}_{\text{Net Fund Return}} - \underbrace{x'_{i,t}}_{\text{Predictors}} \theta)^2$$

- ML: Balance between fit & robustness; allow real-time selection
- Shrinkage/sparsity/both: Ridge, LASSO, elastic net, sparse group LASSO

$$\min_{\theta} \mathcal{L}(\theta; \cdot) = \mathcal{L}(\theta) + \phi(\theta; \cdot)$$

where $\phi(\theta; \cdot)$ is a penalty term.

- Dimension reduction: PCR, PLS
  - OLS after transforming & reducing predictor space to principal components

Methodology 2: Rule-based portfolio sorts

- Fund selection rule
  - Quintile of one predictor (single sort, "$Car1y, 5$"), or
  - Quintile of two predictors (dependent double sort, "$R^2, 1$ & $Car1y, 5$")
- Potential gains:

Nonlinearity, interaction, few parametric restrictions
(+ pros of other ML such as trees, neural networks)

+

Easy to understand
(− cons of other opaque ML)

Methodology 2: Rule-based portfolio sorts

- Fund selection rule
  - Quintile of one predictor (single sort, "$Car1y, 5$"), or
  - Quintile of two predictors (dependent double sort, "$R^2, 1$ & $Car1y, 5$")
- Potential gains:

Nonlinearity, interaction, few parametric restrictions
(+ pros of other ML such as trees, neural networks)

+

Easy to understand
(− cons of other opaque ML)

Methodology 2: Rule-based portfolio sorts

- Fund selection rule
  - Quintile of one predictor (single sort, "$Car1y, 5$"), or
  - Quintile of two predictors (dependent double sort, "$R^2, 1$ & $Car1y, 5$")

- Potential gains:

Nonlinearity, interaction, few parametric restrictions
(+ pros of other ML such as trees, neural networks)

+

Easy to understand
(− cons of other opaque ML)

Methodology 2: Rule-based portfolio sorts

- Fund selection rule
  - Quintile of one predictor (single sort, "$Car1y, 5$"), or
  - Quintile of two predictors (dependent double sort, "$R^2, 1$ & $Car1y, 5$")

- Potential gains:

Nonlinearity, interaction, few parametric restrictions
(+ pros of other ML such as trees, neural networks)

+

Easy to understand
(− cons of other opaque ML)

Real-time evaluation

- Expanding window starts with 7 years

Methods	Training	Validation	Real-Time	Tuning Parameter
ML	5 (yrs) $\uparrow$	1	1	$\phi(\theta;\cdot)$ or # PCs
Rule-Based	6 $\uparrow$	0	1	None
	5 $\uparrow$	1	1	# Top Rules

Validation for tuning parameters

- ML
  - Training: Estimate parameters
  - Real-time: Pick funds w/ highest predicted net-of-fee return
- Rule-based
  - Training: Determine top rules
  - Real-time: Pick funds using top rules

Real-time evaluation

- Expanding window starts with 7 years

Methods	Training	Validation	Real-Time	Tuning Parameter
ML	5 (yrs) $\uparrow$	1	1	$\phi(\theta; \cdot)$ or # PCs
Rule-Based	6 $\uparrow$	0	1	None
	5 $\uparrow$	1	1	# Top Rules

Validation for tuning parameters

- ML
  - Training: Estimate parameters
  - Real-time: Pick funds w/ highest predicted net-of-fee return
- Rule-based
  - Training: Determine top rules
  - Real-time: Pick funds using top rules

Real-time evaluation

- Expanding window starts with 7 years

Methods	Training	Validation	Real-Time	Tuning Parameter
ML	5 (yrs) $\uparrow$	1	1	$\phi(\theta;\cdot)$ or # PCs
Rule-Based	6 $\uparrow$	0	1	None
	5 $\uparrow$	1	1	# Top Rules

Validation for tuning parameters

- ML
  - Training: Estimate parameters
  - Real-time: Pick funds w/ highest predicted net-of-fee return
- Rule-based
  - Training: Determine top rules
  - Real-time: Pick funds using top rules

Real-time evaluation

- Expanding window starts with 7 years

Methods	Training	Validation	Real-Time	Tuning Parameter
ML	5 (yrs) $\uparrow$	1	1	$\phi(\theta; \cdot)$ or # PCs
Rule-Based	6 $\uparrow$	0	1	None
	5 $\uparrow$	1	1	# Top Rules

Validation for tuning parameters

- ML
  - Training: Estimate parameters
  - Real-time: Pick funds w/ highest predicted net-of-fee return
- Rule-based
  - Training: Determine top rules
  - Real-time: Pick funds using top rules

Data

- Fund returns & characteristics from CRSP
- Fund stock holdings from Thomson Reuters to construct predictors
- Mutual fund sample: 1994 - 2016
  - Complete fund characteristics data from 1994 in CRSP
  - Active share and fund duration up to 2015

Data

- Fund returns & characteristics from CRSP
- Fund stock holdings from Thomson Reuters to construct predictors
- Mutual fund sample: 1994 - 2016
  - Complete fund characteristics data from 1994 in CRSP
  - Active share and fund duration up to 2015

Summary statistics

Predictor	Obs.	Mean	SD	AR(1)
ER (%)	900	1.17	0.36	0.95
Flow ($M)	900	-1.20	108.04	0.78
Size ($M)	900	6.12	1.67	0.97
Ret1y (%)	900	10.81	12.41	0.92
Car1y (%)	900	-0.05	0.90	0.84
Ret1m (%)	900	0.87	2.38	$\boxed{0.10}$
RG (%)	900	-0.01	1.26	0.13
TR (%)	900	75.74	61.18	0.93
AS	900	0.81	0.15	0.96
R^2	900	0.91	0.07	0.94
AW	900	0.79	0.21	0.93
Dur (yrs)	900	5.64	3.49	0.96

Main results: Real-time performance

	Avg. Return	CAPM α	FF3 α	C4 α
		Panel A: Benchmark		
OLS	0.56	-0.07	-0.14	-0.12
		Panel B: Machine Learning		
Ridge	0.58	-0.04	-0.11	-0.11
LASSO	0.74**	0.14**	0.11**	0.11**
EN	0.74**	0.14**	0.11**	0.11**
PCR	0.61	0.00	-0.08	-0.09
PLS	0.55	-0.07	-0.14	-0.12
SGL	0.68*	0.07	0.03	0.03
Validation		**Panel C: Rule-Based**		
No	0.79**	0.21*	0.11	0.08
Yes	0.70*	0.11	0.02	-0.01

- Monthly net-of-fee returns in percentage

Does macro information explain performance?

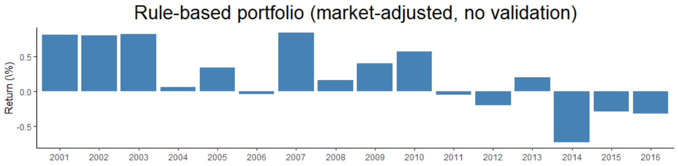
- Empirical specification (Ferson and Schadt, 1996):

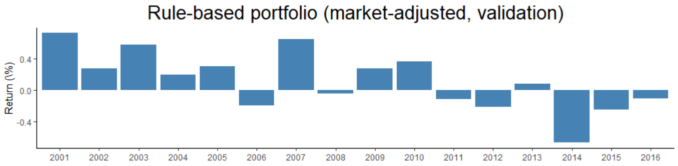
$$R_t - R_{f,t} = \alpha + (\beta + B' z_{t-1})(R_{M,t} - R_{f,t}) + sR_{SMB,t} + hR_{HML,t} + mR_{MOM,t} + \epsilon_t,$$

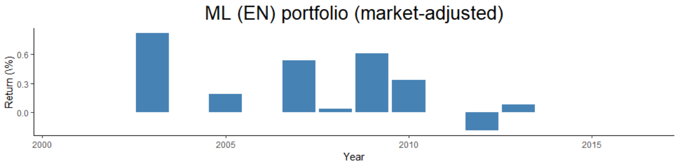
z_{t-1}: lagged macroeconomic variables.

	α	β	B (MKT × TB)	B (MKT × DY)	B (MKT × TS)	B (MKT × DS)	s	h	m
					Panel A: Benchmark				
OLS	-0.05	1.01***	0.58**	-0.08	0.05**	0.07	0.30***	0.03	-0.02
					Panel B: Machine Learning				
Ridge	-0.04	0.99***	0.45**	-0.08	0.05**	0.08	0.30***	0.04	-0.01
EN	0.12**	1.00***	0.07	0.00	0.01	0.03	0.13***	-0.02	0.00
PCR	-0.10	1.00***	0.37**	0.13**	0.02	-0.05	0.31***	0.04	0.04
PLS	-0.05	1.00***	0.57**	-0.08	0.04**	0.06	0.29***	0.04	-0.02
SGL	0.05	1.00***	0.34***	0.06	0.04***	-0.02	0.18***	0.01	0.00
Validation					**Panel C: Rule-Based**				
No	0.10	0.96***	0.35	0.04	0.02	0.01	0.39***	0.05	0.11**
Yes	0.00	0.98***	0.19	0.05	0.03	-0.01	0.36***	0.03	0.10**

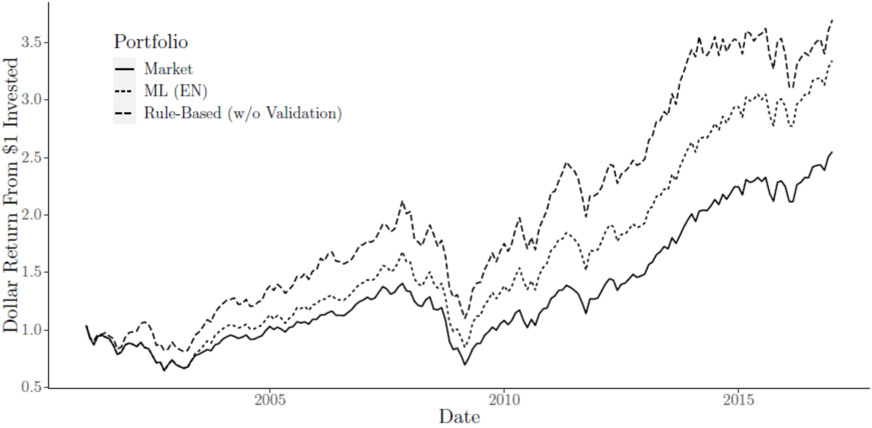
Time variations in performance



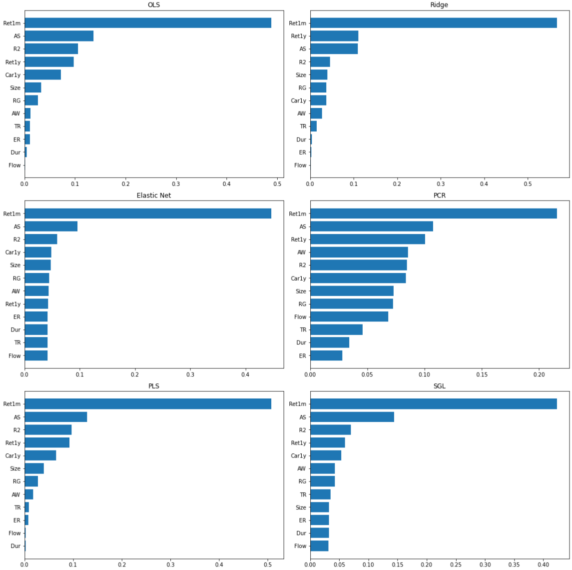




Investment value



Which predictor matters? - ML



- **Key predictor: one-month return (Ret1m)**

Which predictor matters? - Rule-based

Ranking	2001	2002	...	2011	...	2015	2016
1	TR, 5 & Ret1m, 5	Car1y, 5 & Ret1m, 5	...	Ret1m, 5 & Ret1y, 5	...	Ret1m, 5 & Ret1y, 5	Ret1m, 5 & Ret1y, 5
2	Car1y, 5 & Ret1m, 5	TR, 5 & Ret1m, 5	...	AW, 2 & R2, 1	...	Car1y, 5 & Ret1m, 5	Car1y, 5 & Ret1m, 5
3	Ret1m, 5 & Ret1y, 5	Ret1m, 5 & Ret1y, 5	...	Car1y, 5 & Ret1m, 5	...	AW, 2 & R2, 1	AW, 2 & R2, 1

- One-year return: 24/48 rules (in total)
- One-month return: 17/48 rules
- R^2: 7/48 rules

Which predictor matters? - Rule-based

Ranking	2001	2002	...	2011	...	2015	2016
1	TR, 5 &	Car1y, 5 &	...	Ret1m, 5 &	...	Ret1m, 5 &	Ret1m, 5 &
	Ret1m, 5	Ret1m, 5		Ret1y, 5		Ret1y, 5	Ret1y, 5
2	Car1y, 5 &	TR, 5 &	...	AW, 2 &	...	Car1y, 5 &	Car1y, 5 &
	Ret1m, 5	Ret1m, 5		R2, 1		Ret1m, 5	Ret1m, 5
3	Ret1m, 5 &	Ret1m, 5 &	...	Car1y, 5 &	...	AW, 2 &	AW, 2 &
	Ret1y, 5	Ret1y, 5		Ret1m, 5		R2, 1	R2, 1

- One-year return: 24/48 rules (in total)
- One-month return: 17/48 rules
- R^2: 7/48 rules

Do investors use predictive information? - Part I

- **Direct** investor reaction measure: Monthly fund flow

$$F_{i,t+1} = \frac{TNA_{i,t+1}}{TNA_{i,t}} - (1 + R_{i,t+1})$$

- Key independent variable: Predictor-implied performance (PIP)
  - Capture predictive information embedded in each predictor

Construction

- Main specification:

$$F_{i,t+1} = b_0 + b_\alpha PureAlpha_{i,t} + b_P PIP^P_{i,t} + \sum_j b_j FACTOR_{i,j,t} + \theta' \underbrace{X_{i,t}}_{\text{incl. predictor itself}} + \eta_{t+1} + \epsilon_{i,t+1}$$

  - $PIP^P_{i,t}$: after weighting past 18-month components

Do investors use predictive information? - Part I

- **Direct** investor reaction measure: Monthly fund flow

$$F_{i,t+1} = \frac{TNA_{i,t+1}}{TNA_{i,t}} - (1 + R_{i,t+1})$$

- Key independent variable: Predictor-implied performance (PIP)
  - Capture predictive information embedded in each predictor

Construction

- Main specification:

$$F_{i,t+1} = b_0 + b_\alpha PureAlpha_{i,t} + b_P PIP^P_{i,t} + \sum_j b_j FACTOR_{i,j,t} + \theta' \underbrace{X_{i,t}}_{\text{incl. predictor itself}} + \eta_{t+1} + \epsilon_{i,t+1}$$

  - $PIP^P_{i,t}$: after weighting past 18-month components

Do investors use predictive information? - Part I

- **Direct** investor reaction measure: Monthly fund flow

$$F_{i,t+1} = \frac{TNA_{i,t+1}}{TNA_{i,t}} - (1 + R_{i,t+1})$$

- Key independent variable: Predictor-implied performance (PIP)
  - Capture predictive information embedded in each predictor

Construction

- Main specification:

$$F_{i,t+1} = b_0 + b_\alpha PureAlpha_{i,t} + b_P PIP^P_{i,t} + \sum_j b_j FACTOR_{i,j,t} + \theta' \underbrace{X_{i,t}}_{\text{incl. predictor itself}} + \eta_{t+1} + \epsilon_{i,t+1}$$

  - $PIP^P_{i,t}$: after weighting past 18-month components

Do investors use predictive information? - Part I

- **Direct** investor reaction measure: Monthly fund flow

$$F_{i,t+1} = \frac{TNA_{i,t+1}}{TNA_{i,t}} - (1 + R_{i,t+1})$$

- Key independent variable: Predictor-implied performance (PIP)
  - Capture predictive information embedded in each predictor

Construction

- Main specification:

$$F_{i,t+1} = b_0 + b_\alpha PureAlpha_{i,t} + b_P PIP^P_{i,t} + \sum_j b_j FACTOR_{i,j,t} + \theta' \underbrace{X_{i,t}}_{\text{incl. predictor itself}} + \eta_{t+1} + \epsilon_{i,t+1}$$

  - $PIP^P_{i,t}$: after weighting past 18-month components

Do investors use predictive information? - Part II

- Barber et al. (2016): Investors most likley use CAPM

Asset Pricing Model: CAPM						
Monthly Flow	**Predictor P**					
	Size	**RG**	**AS**	**R^2**	**AW**	**Dur**
Pure Alpha	0.632***	0.625***	0.647***	0.651***	0.625***	0.633***
PIP^P	0.520***	0.506	0.301	0.153	0.776***	0.466***
Size	-0.166***	-0.166***	-0.176***	-0.172***	-0.165***	-0.168***
RG		21.591**				
AS			-0.692***			
R^2				1.347***		
AW					0.035	
Dur						0.007
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	152,756	152,756	152,756	152,756	152,756	152,756
Adj. R-Squared	0.026	0.026	0.026	0.026	0.026	0.026

Do investors use predictive information? - Part III

Asset Pricing Model: CAPM						
Monthly Flow	**Predictor P**					
	Size	**RG**	**AS**	**R^2**	**AW**	**Dur**
Panel A: Aggressive Growth						
Pure Alpha	0.654***	0.654***	0.656***	0.662***	0.648***	0.655***
PIP^P	0.701***	0.899***	0.807***	0.643***	0.055	0.731***
Obs.	16,530	16,530	16,530	16,530	16,530	16,530
Panel B: Growth						
Pure Alpha	0.750***	0.737***	0.760***	0.759***	0.737***	0.741***
PIP^P	0.355	0.364	0.422*	0.373	0.678***	0.361
Obs.	80,637	80,637	80,637	80,637	80,637	80,637
Panel C: Growth and Income						
Pure Alpha	0.855***	0.847***	0.874***	0.872***	0.858***	0.851***
PIP^P	0.824	1.078	-0.805	-0.248	0.206	0.569
Obs.	36,859	36,859	36,859	36,859	36,859	36,859

Conclusions

- **Can investors gain from using predictors in real time?**
  - Yes! Regression-based ML (only with sparsity) gives $1.3 \sim 1.7\%$ p.a. alphas
  - Short-term one-month return matters the most
- Do investors react to predictive information?
  - Yes! Great variations in using predictive information
  - Investor reaction is stronger for more growth-oriented funds
- Why does real-time predictability exist?
  - Not due to lack of investor attention
  - But compensation for intensive search algorithms to find skilled managers.
  - Empirical support for Garleanu & Pedersen (2018)

Conclusions

- **Can investors gain from using predictors in real time?**
  - Yes! Regression-based ML (only with sparsity) gives $1.3 \sim 1.7\%$ p.a. alphas
  - Short-term one-month return matters the most
- Do investors react to predictive information?
  - Yes! Great variations in using predictive information
  - Investor reaction is stronger for more growth-oriented funds
- Why does real-time predictability exist?
  - Not due to lack of investor attention
  - But compensation for intensive search algorithms to find skilled managers.
  - Empirical support for Garleanu & Pedersen (2018)

Conclusions

- **Can investors gain from using predictors in real time?**
  - Yes! Regression-based ML (only with sparsity) gives $1.3 \sim 1.7\%$ p.a. alphas
  - Short-term one-month return matters the most
- Do investors react to predictive information?
  - Yes! Great variations in using predictive information
  - Investor reaction is stronger for more growth-oriented funds
- Why does real-time predictability exist?
  - Not due to lack of investor attention
  - But compensation for intensive search algorithms to find skilled managers.
  - Empirical support for Garleanu & Pedersen (2018)

Conclusions

- **Can investors gain from using predictors in real time?**
  - Yes! Regression-based ML (only with sparsity) gives $1.3 \sim 1.7\%$ p.a. alphas
  - Short-term one-month return matters the most
- Do investors react to predictive information?
  - Yes! Great variations in using predictive information
  - Investor reaction is stronger for more growth-oriented funds
- Why does real-time predictability exist?
  - Not due to lack of investor attention
  - But compensation for intensive search algorithms to find skilled managers.
  - Empirical support for Garleanu & Pedersen (2018)

Conclusions

- **Can investors gain from using predictors in real time?**
  - Yes! Regression-based ML (only with sparsity) gives $1.3 \sim 1.7\%$ p.a. alphas
  - Short-term one-month return matters the most
- Do investors react to predictive information?
  - Yes! Great variations in using predictive information
  - Investor reaction is stronger for more growth-oriented funds
- Why does real-time predictability exist?
  - Not due to lack of investor attention
  - But compensation for intensive search algorithms to find skilled managers.
  - Empirical support for Garleanu & Pedersen (2018)

Future work

- Economic gains for long-term investors
- Who incorporate predictive information?
  - Investor heterogeneity: sophisticated v.s. unsophisticated
- When do investors acquire information?
  - Uncertainty may affect investors' attention or willingness

Thank You!

Future work

- Economic gains for long-term investors
- Who incorporate predictive information?
  - Investor heterogeneity: sophisticated v.s. unsophisticated
- When do investors acquire information?
  - Uncertainty may affect investors' attention or willingness

Thank You!

Future work

- Economic gains for long-term investors
- Who incorporate predictive information?
  - Investor heterogeneity: sophisticated v.s. unsophisticated
- When do investors acquire information?
  - Uncertainty may affect investors' attention or willingness

Thank You!

Future work

- Economic gains for long-term investors
- Who incorporate predictive information?
  - Investor heterogeneity: sophisticated v.s. unsophisticated
- When do investors acquire information?
  - Uncertainty may affect investors' attention or willingness

Thank You!

Future work

- Economic gains for long-term investors
- Who incorporate predictive information?
  - Investor heterogeneity: sophisticated v.s. unsophisticated
- When do investors acquire information?
  - Uncertainty may affect investors' attention or willingness

Thank You!

Future work

- Economic gains for long-term investors
- Who incorporate predictive information?
  - Investor heterogeneity: sophisticated v.s. unsophisticated
- When do investors acquire information?
  - Uncertainty may affect investors' attention or willingness

Thank You!

PIP construction - Part I

- Why PIP?
  - Predictors for managerial skills or behavioral biases? Not clear
  - Solution:

    Step 1. Capture skill by similarity in return pattern with a portfolio of funds sorted on predictor (extend Cohen et al., 2005)

    $$\text{Portfolio: Funds with high predictor value} - \text{funds with low predictor value}$$

    Step 2. Use fitted performance as skill component captured by predictor (PIP)

PIP construction - Part I

- Why PIP?
  - Predictors for managerial skills or behavioral biases? Not clear
  - Solution:

    Step 1. Capture skill by similarity in return pattern with a portfolio of funds sorted on predictor (extend Cohen et al., 2005)

    Portfolio: Funds with high predictor value − funds with low predictor value

    Step 2. Use fitted performance as skill component captured by predictor (PIP)

PIP construction - Part I

- Why PIP?
  - Predictors for managerial skills or behavioral biases? Not clear
  - Solution:

    Step 1. Capture skill by similarity in return pattern with a portfolio of funds sorted on predictor (extend Cohen et al., 2005)

    $$\text{Portfolio: Funds with high predictor value} - \text{funds with low predictor value}$$

    Step 2. Use fitted performance as skill component captured by predictor (PIP)

PIP construction - Part II

1. Estimate performance similarity & risk exposure:

$$R_{i,\tau} - R_{f,\tau} = \alpha_{i,t} + \gamma_{i,t}^{P} R_{\tau}^{P} + \sum_{j} \beta_{i,j,t} f_{j,\tau} + \epsilon_{i,\tau},\ \tau \in \{t-1, \ldots, t-60\}.$$

Illustration

2. Decompose returns into three performance components:

$$R_{i,t} - R_{f,t} = \underbrace{\widehat{\alpha}_{i,t}}_{\text{pure alpha}} + \underbrace{\widehat{\gamma}_{i,t}^{P} R_{t}^{P}}_{\text{predictor } P\text{-implied performance (PIP)}} + \underbrace{\sum_{j} \widehat{\beta}_{i,j,t} f_{j,t}}_{\text{risk premia}}.$$

Back

PIP construction - Part II

1. Estimate performance similarity & risk exposure:

$$R_{i,\tau} - R_{f,\tau} = \alpha_{i,t} + \gamma_{i,t}^{P} R_{\tau}^{P} + \sum_{j} \beta_{i,j,t} f_{j,\tau} + \epsilon_{i,\tau},\ \tau \in \{t-1, \ldots, t-60\}.$$

Illustration

2. Decompose returns into three performance components:

$$R_{i,t} - R_{f,t} = \underbrace{\widehat{\alpha}_{i,t}}_{\text{pure alpha}} + \underbrace{\widehat{\gamma}_{i,t}^{P} R_{t}^{P}}_{\textbf{predictor } P\textbf{-implied performance (PIP)}} + \underbrace{\sum_{j} \widehat{\beta}_{i,j,t} f_{j,t}}_{\text{risk premia}}.$$

Back

Portfolio of funds sorted on predictor P = Size

Predictor	Value	Group of MFs
$Size$	Low	MF1
	$\vdots$	$\vdots$
	High	MF5

- R_t^{Size}(i.e., R_t^P) $= Ret_t^{MF5} - Ret_t^{MF1}$

Back