期末总结

石宛青

（武汉大学经济与管理学院）

2025 年 1 月 15 日

本学期工作

- 论文：
  - 重新整理修改绿债论文，更正了数据处理问题和代码规范性。
  - 更新 climate 论文的数据，规范代码，增加了新机制与进一步分析，参加会议，整理了修改意见，准备进行进一步修改。
- 数据与复现：
  - 爬取了绿债市场的所有债券公告文本，准备构建绿色特质信息指标，可用于绿债中解决机制中一个一直不行的结果
  - 复现了《Pricing of sustainability-linked bonds》（JFE，2024），阅读相关文献，找研究价值。
- 感悟 & 提升 & 下一步：
  - 数据和代码的规范性非常重要！便于复现和修改。项目代码管理能力提升较多。
  - 自主找到有价值的选题的能力还比较差。扩展阅读其他领域文章如 ML. 目前的太局限了。
  - 复现不够；时间管理。

Real-time Machine Learning in the Cross-Section of Stock Returns

Bin Li, Alberto G. Rossi, Xuemin Yan, Lingling Zheng

(JFE,2025)

石宛青

（武汉大学金融系）

2025 年 01 月 15 日

Motivation

- attention on： ML on asset pricing, return prediction——common: construct long-short investment strategies, with highly profitable.
- an important issue has yet to be fully addressed： real-time implementability
  - use published anomaly variables as predictors, assuming investors knew at training period's start, despite most being discovered years later.
  - natural for econometrician(measure risk premium), but for real-time investor?
  - look-ahead bias： eg: asset growth anomaly of Cooper et al. (2008)
  - publication bias (Harvey et al., 2016):strong in-sample performance——hindsight overly optimistic assessment——prior studies overstated for real-time investors.
- How to solve this problem?——change predictors

Motivation

- this paper: examine ML strategies based on a "universe" of over 18,000 fundamental signals.

  1. no look-ahead bias: Fundamental analysis, dating back to Graham and Dodd (1934)——investors naturally consider as return predictors
  2. economists highlighted economic intuition behind return predictors in ML context (Arnott et, 2019)——related to firm cash flows and valuations
  3. construct a "universe" using permutational arguments——real-time investors don't know which signals turn out to be significant ex-post

Question

- How does the real-time ML method, based on the "universe" of fundamental signals, perform in predicting stock returns and investment outcomes?
  - While positive and significant, the out-of-sample performance of these strategies is significantly weaker than those documented by prior studies.
  - We find qualitatively similar results when examining a “universe” of past-return-based signals.

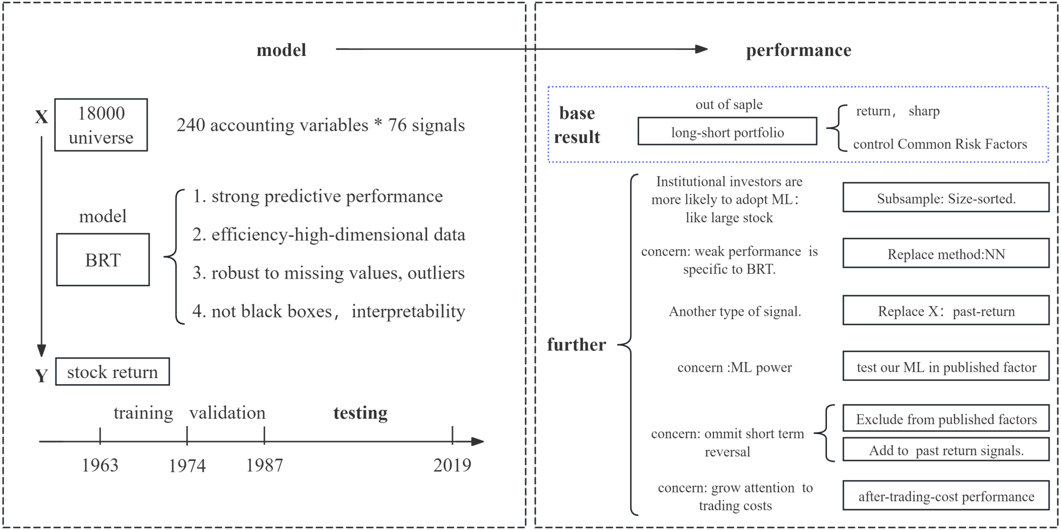
Contribution

- contributes to literature employing ML methods in empirical asset pricing.
  - prior: large economic gains to investors using machine learning strategies
  - expand: perspective of real-time investors——significantly less profitable
- contributes on accounting variables on stock return prediction using ML
  - prior: a fundamental deviation index from 105 signals (Avramov et al.,2022)
  - prior: accounting variables on stock return prediction (Haugen and Baker, 1996)
  - expand: focus on 'real-time'—— "universe" using permutational arguments

Contribution

- contributes to literature on the limitations of ML methods in invest strategies.
  - prior: profitability of strategies is derived from difficult-to-arbitrage stocks or limits-to-arbitrage period(Avramov et al.,2023)
  - Validation： out-of-sample predictability is significantly weaker among large stocks. expand: discovered anomaly may not be implementable in real-time.
- contributes to literature on ML not work as well in finance as in other areas
  - prior: the lack of data (time series), low signal-to-noise ratio, adaptive nature of financial markets.(Arnott et al.,2019)
  - expand: modest performance of real-time ML strategies could be a manifestation

Design


model
performance
X
18000 universe
240 accounting variables * 76 signals
model
BRT
1. strong predictive performance
2. efficiency-high-dimensional data
3. robust to missing values, outliers
4. not black boxes， interpretability
Y
stock return
training
validation
testing
1963
1974
1987
2019
base result
out of saple
long-short portfolio
return， sharp
control Common Risk Factors
further
Institutional investors are more likely to adopt ML： like large stock
Subsample: Size-sorted.
concern: weak performance is specific to BRT.
Replace method:NN
Another type of signal.
Replace X： past-return
concern :ML power
test our ML in published factor
concern: ommit short term reversal
Exclude from published factors
Add to past return signals.
concern: grow attention to trading costs
after-trading-cost performance

Design-Stock Sample and Fundamental Signals

- CRSP + Compustat： NYSE, AMEX, and NASDAQ common stocks
- sample spans from July 1963 to June 2019, 15035 stocks.
- **universe:240 accounting variables** $\times$ **76 signals** (Yan and Zheng， 2017)

#	*Variable*	*Description*	*Missing Rate*	*Start Year*
1	ACCHG	Accounting changes - cumulative effect	39.29%	1988
2	ACO	Current assets other total	0.76%	1963
3	ACOX	Current assets other sundry	2.20%	1963
4	ACT	Current assets - total	2.13%	1963
5	AM	Amortization of intangibles	33.03%	1965
6	AO	Assets – other	0.06%	1963

#	*Description*	#	*Description*	#	*Description*	#	*Description*	#	*Description*
1	X/AT	16	Δ in X/AT	31	%Δ in X/AT	46	ΔX/LAGAT	61	%Δ in X - %Δ in AT
2	X/ACT	17	Δ in X/ACT	32	%Δ in X/ACT	47	ΔX/LAGACT	62	%Δ in X - %Δ in ACT
3	X/INVT	18	Δ in X/INVT	33	%Δ in X/INVT	48	ΔX/LAGINVT	63	%Δ in X - %Δ in INVT
4	X/PPENT	19	Δ in X/PPENT	34	%Δ in X/PPENT	49	ΔX/LAGPPENT	64	%Δ in X - %Δ in PPENT
5	X/LT	20	Δ in X/LT	35	%Δ in X/LT	50	ΔX/LAGLT	65	%Δ in X - %Δ in LT
6	X/LCT	21	Δ in X/LCT	36	%Δ in X/LCT	51	ΔX/LAGLCT	66	%Δ in X - %Δ in LCT

- 18240 (240 $\times$ 76)——18113 meaningful signals, and some are redundant.

Design-Methodology

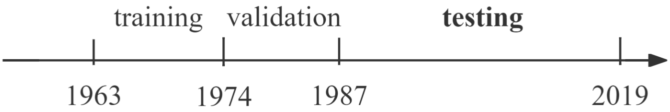
- Prediction Equation:

$$R_{i,t+1} = f\left(\mathbf{x}_{i,t} \mid \boldsymbol{\theta}\right) + \epsilon_{i,t+1}$$

  - $R_{i,t+1}$:annual excess return for stock i from July of year t to June of year t + 1
  - pair accounting variables in year t − 1
  - scale $\mathbf{x}_{i,t}$ ranks to the interval [1, +1].
- why annual excess returns?
  - X are constructed from annual financial statements and are updated annually
  - number of signals larger than prior,Predict annual returns is more efficient than monthly
- ML v.s. LR
  - 1) variable selection, 2) model combination, 3) regularization,4)nonlinearities
  - use “unmined” versus “known” predictors: advantages and disadvantages

Design-Methodology

- BRT——Regression Trees + Boosting



  - refit model annually:training sample + 1, maintain validation period at 12 years.
  - Every year, we generate return forecasts for all the stocks
  - then construct decile portfolios based on the predicted returns
  - hold the long-short portfolios for 12 months

- replace missing values with the cross-sectional median——0

Design-Performance Evaluation

- alpha: CAPM 1, FF 3, Carhart 4, FF 5, FF 5 + Mom, and Q factor models

$$
\begin{aligned}
r_t &= \alpha + \beta\, MKT_t + \epsilon_t \\
r_t &= \alpha + \beta\, MKT_t + s\, SMB_t + h\, HML_t + \epsilon_t \\
r_t &= \alpha + \beta\, MKT_t + s\, SMB_t + h\, HML_t + u\, UMD_t + \epsilon_t \\
r_t &= \alpha + \beta\, MKT_t + s\, SMB_t + h\, HML_t + r\, RMW_t + c\, CMA_t + \epsilon_t \\
r_t &= \alpha + \beta\, MKT_t + s\, SMB_t + h\, HML_t + r\, RMW_t + c\, CMA_t + u\, UMD_t + \epsilon_t \\
r_t &= \alpha + \beta\, MKT_t + s\, SMB_t + r\, ROE_t + i\, IA_t + \epsilon_t
\end{aligned}
$$

Result-Baseline

	Equal-weighted					Value-weighted				
Rank	*Pred*	*Avg*	*t-stat*	*SD*	*SR*	*Pred*	*Avg*	*t-stat*	*SD*	*SR*
1 (Low)	-0.04	-0.01	-0.05	7.51	-0.01	0.00	0.40	1.30	6.18	0.22
2	0.30	0.49	1.58	6.22	0.28	0.30	0.58	2.39	5.53	0.36
3	0.49	0.65	2.12	6.02	0.37	0.50	0.58	2.63	4.73	0.42
4	0.64	0.74	2.65	5.64	0.46	0.64	0.75	3.21	4.60	0.56
5	0.73	0.76	2.72	5.45	0.48	0.73	0.60	2.34	4.61	0.45
6	0.80	0.90	3.23	5.44	0.58	0.80	0.66	2.99	4.51	0.51
7	0.88	0.90	3.18	5.57	0.56	0.88	0.68	2.68	4.93	0.48
8	0.97	0.96	3.17	5.43	0.62	0.97	0.49	1.81	4.82	0.35
9	1.12	0.93	2.84	5.78	0.56	1.11	0.64	2.20	5.15	0.43
10 (High)	1.69	0.94	2.55	6.71	0.48	1.61	0.80	2.51	5.96	0.47
10-1	1.74	0.95	6.63	3.26	1.02	1.61	0.40	2.34	4.68	0.30

- BRT Gu et al.(2020) EW monthly long-short portfolio return of 2.14% per month and SR of 1.73.Chen et al.(2022):2.60;Freyberger et al.(2020): 2.75;
- economic gains to real-time investors from using ML are smaller than prior.

Result-Controlling for Common Risk Factors

	Equal Weighted											
Rank	CAPM		FF3		Carhart		FF5		FF5+Mom		Q	
	alpha	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*
L(ow)	-0.84	-3.72	-0.76	-4.91	-0.71	-4.65	-0.44	-3.43	-0.43	-3.33	-0.30	-1.78
2	-0.21	-1.00	-0.19	-1.50	-0.17	-1.48	-0.09	-0.74	-0.09	-0.72	0.02	0.17
3	-0.07	-0.34	-0.06	-0.64	-0.06	-0.62	-0.03	-0.36	-0.03	-0.34	0.07	0.70
4	0.05	0.29	0.03	0.50	0.03	0.48	0.02	0.23	0.02	0.24	0.09	1.29
5	0.10	0.58	0.07	1.02	0.07	1.06	0.02	0.32	0.03	0.38	0.09	1.16
6	0.24	1.23	0.18	2.37	0.23	2.80	0.11	1.44	0.15	1.91	0.19	2.03
7	0.23	1.26	0.21	3.05	0.20	2.74	0.22	2.79	0.21	2.65	0.26	3.73
8	0.30	1.59	0.28	3.17	0.29	3.34	0.28	3.23	0.29	3.34	0.35	3.84
9	0.22	1.33	0.22	2.22	0.29	3.08	0.35	3.25	0.39	3.74	0.45	4.87
H(igh)	0.18	0.84	0.25	1.60	0.37	2.64	0.58	3.87	0.65	4.43	0.68	4.67
H-L	1.01	6.30	1.01	6.35	1.08	6.43	1.03	5.42	1.08	5.60	0.98	5.11

- Results remain economically significant after control for risk exposures, but weaker than prior

Result-Different Market Capitalizations

Sub-Samples	Returns		SR	CAPM		FF3		Carhart		FF5		FF5+MOM		Q	
Equal Weight															
	Avg	*t-stat*		*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*
Large Stocks	0.63	2.93	0.57	0.71	3.17	0.74	3.11	0.77	3.70	0.93	3.39	0.93	3.63	0.90	3.27
Small Stocks	1.13	6.14	1.16	1.18	6.48	1.20	6.21	1.22	5.98	1.23	5.48	1.24	5.35	1.14	5.10
Value Weight															
	Avg	*t-stat*		*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*
Large Stocks	0.27	1.23	0.20	0.35	1.37	0.41	1.77	0.52	2.33	0.60	2.39	0.67	2.58	0.66	2.63
Small Stocks	1.16	5.50	1.04	1.18	5.58	1.21	5.29	1.20	5.08	1.27	5.12	1.25	4.87	1.16	4.41

- benefits of machine learning strategies may be even more limited for institutional investors.

Result-Neural Networks

Equal Weight

Method	Returns		SR	CAPM		FF3		Carhart		FF5		FF5+MOM		Q	
	Avg	*t-stat*		*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*
NN1	0.66	3.19	0.61	0.70	3.02	0.65	3.26	0.83	5.09	0.73	3.27	0.86	4.45	0.74	3.21
NN2	0.85	4.18	0.92	1.02	5.04	0.93	5.33	0.83	4.93	0.61	4.36	0.56	4.01	0.57	3.56
NN3	0.31	1.76	0.40	0.38	1.99	0.41	2.34	0.37	2.23	0.40	2.83	0.37	2.67	0.37	2.49
NN4	0.41	2.07	0.47	0.37	1.62	0.41	1.98	0.44	2.34	0.62	3.67	0.63	3.79	0.61	3.55
NN5	0.09	0.48	0.10	0.06	0.28	0.10	0.53	0.13	0.74	0.12	0.82	0.14	1.00	0.26	1.74

Value Weighted

Method	Returns		SR	CAPM		FF3		Carhart		FF5		FF5+MOM		Q	
	Avg	*t-stat*		*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*
NN1	0.01	0.03	0.01	0.01	0.03	-0.01	-0.04	0.23	1.18	0.15	0.59	0.32	1.46	0.24	0.99
NN2	0.04	0.15	0.03	0.21	0.79	0.08	0.40	-0.01	-0.04	-0.23	-1.17	-0.27	-1.36	-0.25	-1.13
NN3	-0.01	-0.02	-0.01	0.07	0.28	0.06	0.23	0.16	0.71	-0.10	-0.41	-0.01	-0.06	0.00	-0.01
NN4	0.20	0.60	0.15	0.13	0.30	0.23	0.69	0.22	0.72	0.54	2.02	0.51	1.94	0.54	1.80
NN5	0.20	0.75	0.16	0.11	0.37	0.15	0.56	0.17	0.68	0.12	0.56	0.14	0.67	0.25	1.07

- Real-time performance of ML strategies is more modest than that portrayed by prior.

Result-Past-return Signals

Method	Returns		SR	CAPM		FF3		Carhart		FF5		FF5+Mom		Q	
	Avg	*t-stat*		*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*	*alpha*	*t-stat*
							Equal Weight								
BRT	1.38	4.93	1.04	1.63	5.86	1.56	6.90	1.09	6.62	1.09	4.82	0.78	5.83	0.78	3.89
NN1	1.17	6.61	1.33	1.32	7.56	1.32	7.81	0.93	9.33	1.14	5.99	0.87	7.96	0.88	4.83
NN2	1.40	7.46	1.45	1.53	7.81	1.50	8.55	1.10	9.69	1.20	5.24	0.93	6.86	0.96	4.31
NN3	1.30	6.34	1.42	1.42	6.90	1.40	7.62	1.01	8.81	1.16	5.74	0.89	7.60	0.93	4.82
NN4	0.74	5.64	0.81	0.75	5.95	0.76	5.75	0.48	2.53	0.62	2.98	0.43	1.99	0.43	1.74
NN5	0.10	0.50	0.11	-0.02	-0.07	0.09	0.54	0.02	0.14	0.33	2.71	0.27	2.00	0.29	2.18
							Value Weight								
BRT	0.78	2.41	0.46	1.17	4.00	1.07	4.23	0.63	3.05	0.56	2.37	0.28	1.55	0.28	1.36
NN1	0.87	4.29	0.72	0.98	4.54	1.02	5.08	0.57	3.34	0.90	4.23	0.58	3.14	0.69	3.08
NN2	1.06	4.34	0.78	1.21	4.73	1.23	4.91	0.77	4.49	1.00	3.36	0.68	3.66	0.83	2.81
NN3	1.00	3.79	0.80	1.09	4.22	1.12	4.35	0.63	3.76	0.98	3.32	0.63	3.70	0.78	2.44
NN4	0.38	1.68	0.31	0.39	1.67	0.43	1.97	0.10	0.37	0.39	1.50	0.16	0.58	0.27	0.93
NN5	0.32	1.30	0.26	0.19	0.76	0.32	1.73	0.17	0.88	0.56	3.20	0.43	2.53	0.49	2.17

- the past 120 months, excluding the most recent month
- results based on past-return signals are broadly consistent with those based on fundamental signals.

Result-ML Implementation

Method	Returns Avg	Returns t-stat	SR	CAPM alpha	CAPM t-stat	FF3 alpha	FF3 t-stat	Carhart alpha	Carhart t-stat	FF5 alpha	FF5 t-stat	FF5+Mom alpha	FF5+Mom t-stat	Q alpha	Q t-stat
Equal Weight															
BRT	3.57	8.91	2.35	3.67	9.36	3.69	9.13	3.37	8.61	3.66	7.85	3.43	8.05	3.35	6.71
NN1	3.78	9.74	2.69	3.86	9.87	3.90	10.06	3.60	9.48	3.78	9.34	3.57	9.10	3.67	8.82
NN2	3.87	9.93	2.95	3.94	10.00	3.96	10.08	3.66	9.43	3.78	9.60	3.57	9.36	3.61	9.00
NN3	3.94	10.09	2.84	4.03	10.31	4.05	10.40	3.75	9.57	3.90	9.39	3.69	9.23	3.79	8.98
NN4	2.94	9.59	2.27	3.01	9.78	2.99	9.78	2.78	7.82	2.89	8.98	2.75	7.71	2.71	7.88
NN5	1.70	5.92	1.33	1.62	5.29	1.69	5.44	1.71	5.37	1.77	5.41	1.78	5.37	1.90	5.50
Value Weight															
BRT	1.51	4.53	0.71	1.66	5.50	1.70	5.14	1.13	4.08	1.56	3.25	1.15	3.37	1.09	2.20
NN1	1.68	4.84	0.85	1.90	5.56	1.96	6.01	1.30	4.41	1.70	4.49	1.24	4.11	1.33	3.54
NN2	1.87	5.18	0.99	2.04	5.56	2.08	6.13	1.42	4.74	1.74	4.73	1.29	4.34	1.43	3.56
NN3	1.71	4.93	0.90	1.94	5.38	1.95	6.08	1.25	5.45	1.48	3.87	1.00	3.98	1.16	3.13
NN4	1.37	6.32	0.82	1.51	6.23	1.48	7.41	0.98	4.19	1.25	4.71	0.90	3.33	0.89	3.20
NN5	0.69	3.78	0.49	0.76	3.67	0.81	4.22	0.57	2.59	0.81	4.02	0.63	2.93	0.72	3.55

- Green, Hand, and Zhang (2017, GHZ):94 factors;Chen and Zimmermann (2022, CZ):207 factors
- our ML generating strong performance when use published predictors

Result-Short-term Reversal

Method	GHZ94 Ret.	GHZ94 t-stat	GHZ94 SR	GHZ93 Ret.	GHZ93 t-stat	GHZ93 SR
			Equal Weight			
BRT	3.57	8.91	2.35	3.04	9.11	1.80
NN1	3.78	9.74	2.69	3.22	9.99	2.46
NN2	3.87	9.93	2.95	3.37	10.68	2.30
NN3	3.94	10.09	2.84	3.27	10.57	2.48
NN4	2.94	9.59	2.27	2.30	8.91	1.84
NN5	1.70	5.92	1.33	0.72	3.65	0.62

Method	TECH120 Ret.	TECH120 t-stat	TECH120 SR	TECH119 Ret.	TECH119 t-stat	TECH119 SR
			Equal Weighted			
BRT	1.81	6.40	1.77	1.38	4.93	1.04
NN1	1.61	7.30	1.98	1.17	6.61	1.33
NN2	1.75	8.82	2.28	1.40	7.46	1.45
NN3	1.52	7.99	1.94	1.30	6.34	1.42
NN4	0.98	6.80	1.21	0.74	5.64	0.81
NN5	0.66	3.73	0.84	0.10	0.50	0.11

- omission of short-term reversal have a moderate impact on performance
- short-term reversal alone cannot explain the performance difference

Result-After-trading-cost Performance

Methods	a Gross Return	b Turnover (2-Sided)	c Ave. Spread Paid	d≈ b× c Return Reduction	e=a-d Net Return
			Equal Weight		
BRT	97	14	177	24	73
NN1	67	14	174	25	43
NN2	85	14	187	27	58
NN3	30	14	178	26	4
NN4	38	14	185	24	14
NN5	7	12	177	21	-14

Methods	Panle A: TECH119					Panel B: TECH120				
	a	b	c	d≈ b× c	e=a-d	a	b	c	d≈ b× c	e=a-d
					Equal Weight					
BRT	141	124	190	238	-97	190	144	232	338	-148
NN1	123	153	201	305	-182	167	156	212	330	-163
NN2	146	149	206	309	-163	182	155	219	340	-158
NN3	136	152	210	318	-182	157	156	216	338	-181
NN4	76	153	205	317	-241	104	160	213	340	-236
NN5	9	106	210	221	-212	69	130	213	277	-207

- net performance of ML strategies is positive for fundamental signals and negative for past-return signals.

result-Additional Results

- employs rolling windows instead of recursive windows
  - 0.83% （t = 4.29）， SR = 0.77,weaker but robust
- whether our results are robust to alternative training and validation periods.
  - 9 alternative: train:10-18.validation:10-14——robust
- examines different subsets of our universe of fundamental signals
  - 240 Accounting Variables:best performance in 180 or 210.
  - Subsets of Financial Ratio Configurations: weaker than universe
- compare in-sample performance with out-of-sample performance
  - in-sample R2， return >> out-of-sample
- whether performance varies with economic and market conditions.
  - investor sentiment, VIX, market liquidity, business cycle, market state

Question and idea

- Question：
  - Short-term Reversal 的一小节是否有必要?
  - 240 个会计变量的结果：0.95 （t= 6.63),SR = 1.02. 是否有必要使用 18000 universe
- Idea：
  - 投资者通常在开始投资时已经能够获取的变量包括：宏观 + 微观：财务报表：市值和股价；宏观经济；行业：股东结构与管理层信息；历史股息和分红记录：信用评级和债务信息。
  - 根据因子提出的时间来持续加入因子，动态的 ML?

Thanks!