Wharton
UNIVERSITY of PENNSYLVANIA

JACOBS LEVY EQUITY MANAGEMENT CENTER
for Quantitative Financial Research

The Statistical Limit of Arbitrage

Dacheng Xiu

The Statistical Limit of Arbitrage

Rui Da, Stefan Nagel, and Dacheng Xiu

University of Chicago

Frontiers in Quantitative Finance Conference

September 22, 2023

Introduction

- Rapid advances in machine learning should, in theory, make statistical arbitrage investing a lot easier and more lucrative.
  - See a recent survey: "Financial Machine Learning", Kelly and Xiu (2023, FnT Finance)
- Is there a ceiling for gains in machine-learned arbitrage?
  - Yes! And this ceiling is much lower than what investors might anticipate
  - Pricing errors (alphas) $\neq$ Arbitrage opportunities

Introduction

- Ross (1976)'s arbitrage pricing theory (APT) rules out near-arbitrage opportunities based on *true* parameters in the return generating process
  - Implicit assumption: arbitrageurs know true parameters, as in rational expectations asset pricing models following tradition of Lucas (1978)
- Yet, arbitrageurs in practice learn about profit opportunities from *statistical analysis* prior to investment.
- The challenge of learning about *a large number of* alphas induces a limit to arbitrage due to statistical learning.
- Statistical limits to arbitrage is absent from existing theories of limits to arbitrage.

Myron Scholes' Description of LTCM's Trading Strategy

— Illustration created by Bing powered by OpenAI's Dall-E technology

Challenge Facing Arbitrageurs: Real or Fake?

Summary of Theoretical Results

- We document a statistical limit to arbitrage in a linear asset pricing model where arbitrageurs are only allowed to employ a *feasible* trading strategy that relies on historical data to make inference on alpha signals.
- We derive the optimal Sharpe ratio achievable by any feasible arbitrage trading strategies.
- We shed light on a theoretical gap between the infeasible Sharpe ratio (pricing errors) and the feasible Sharpe ratio (investment opportunities).
- We show how arbitrageurs can design an "all-weather" feasible strategy that achieves the optimality.
- We examine alternative machine learning strategies that exploit multiple testing, shrinkage, and selection techniques.

Summary of Empirical Findings

- The cross-sectional R^2s of a 27-factor model we built akin to BARRA are rather low, with a time-series average 8.25% from Jan 1965 to Dec 2020.
- Among 12,415 test statistics in total, only 6.35% (0.63%) of the t-statistics are greater than 2.0 (3.0) in absolute values.
- Feasible arbitrage portfolios achieve a moderately low annualized Sharpe ratio, about 0.5, whereas the perceived (infeasible) Sharpe ratios are considerably higher, around 3.0.

Return Model

- N assets in the market. Their excess returns follow a linear factor model.
- For the talk: An N-dimensional return vector follows:

$$r_t = \alpha + u_t, \quad \text{where} \quad u_{i,t} \sim \mathcal{N}(0,1),$$

where alphas are i.i.d. across assets.

Classical Arbitrage Pricing Theory

- With perfect knowledge of α, the mean-variance optimal arbitrage portfolio is

$$w = \alpha,$$

which achieves a Sharpe ratio

$$S^\star := \sqrt{\alpha^\intercal \alpha}.$$

- Ingersoll (1984, JF) shows that the condition for absence of near-arbitrage:

$$S^\star \leq C, \quad \text{for some constant } C > 0.$$

- But near-arbitrage under knowledge of $\alpha \neq$ near-arbitrage w/o knowledge of α

A Simple Feasible Strategy

- But arbs do not observe true DGP. They infer α from a sample of size T:

$$\widehat{\alpha} = \alpha + \bar{u}, \quad \bar{u} \sim \mathcal{N}(0, 1/T).$$

- Suppose further they form an arbitrage portfolio with weights

$$\widehat{\omega} = \widehat{\alpha}.$$

Statistical Limits: Sharpe Ratio Gap

- OOS Sharpe ratio S achieved by $\widehat{\omega} = \widehat{\alpha}$ vs. $S^\star$ based on $w = \alpha$:

$$\frac{(S^\star)^2}{S^2} = 1 + \frac{1}{T} + \frac{N}{T(S^\star)^2}.$$

- The gap can be substantial when alphas sufficiently rare and/or small!

Statistical Limits: A Non-trivial Example

- Suppose alphas are drawn from:

$$\alpha_i \overset{i.i.d.}{\sim} \begin{cases} \mu & \text{with prob. } \rho/2 \\ -\mu & \text{with prob. } \rho/2 \\ 0 & \text{with prob. } 1-\rho \end{cases}, \quad 1 \le i \le N.$$

- Suppose only a small portion of assets have a nonzero yet small alpha:

$$\mu \sim T^{-1/2}, \qquad \rho \sim N^{-1/2}$$

- Suppose $N^{1/2}/T \to \infty$ and $T \to \infty$.
- Infeasible strategy $w = \alpha$ generates

$$S^{\star} \sim N^{1/2}/T \to \infty,$$

- Feasible strategy $w = \widehat{\alpha}$ generates

$$S \sim 1/T \to 0.$$

More Difficult Questions

- Is it possible that a smart arbitrageur can find a strategy that closes this gap?
- What is the optimal feasible strategy?

Impact of Feasibility Constraint

- Let $S(\widehat{w})$ be the Sharpe ratio generated by any $\widehat{w}$.
- Feasible strategy: function of historical data (returns from $t-T+1$ to t).
- Theorem 1: $\widehat{w}$ is feasible $\Longrightarrow$

$$S(\widehat{w}) \leq S^{\mathrm{OPT}} + o_{\mathrm{P}}(1), \quad (S^{\mathrm{OPT}})^2 := \mathrm{E}\big(\mathrm{E}(\alpha|\mathscr{G})^{\intercal}\mathrm{E}(\alpha|\mathscr{G})\big).$$

  Here $\mathscr{G}$ is the information set generated by historical returns.
- Not surprisingly, feasible Sharpe ratio smaller than infeasible one:

$$\mathrm{E}((S^{\star})^2) \geq \left(S^{\mathrm{OPT}}\right)^2.$$

- This provides a solution to the classical problem of optimal portfolio allocation when parameters (mean and covariances) are unknown.

S^{OPT}

$\rho \times 100$ \ $\mu/\sigma \times \sqrt{12}$	10.95	8.22	5.48	4.38	3.29	2.74	2.19	1.92	1.64	1.37	1.10	0.99	0.88	0.77	0.66	0.55	0.44	0.33	0.22	0.16	0.11
50.00	70.71	53.03	35.35	28.21	20.83	16.95	12.96	10.94	8.94	6.97	5.04	4.29	3.55	2.85	2.19	1.58	1.05	0.61	0.28	0.16	0.07
25.00	50.00	37.50	24.99	19.92	14.55	11.66	8.64	7.12	5.64	4.22	2.92	2.43	1.98	1.56	1.17	0.83	0.54	0.31	0.14	0.08	0.04
20.00	44.72	33.54	22.35	17.80	12.96	10.33	7.58	6.20	4.85	3.59	2.44	2.02	1.63	1.28	0.96	0.68	0.44	0.25	0.11	0.06	0.03
15.00	38.73	29.05	19.36	15.40	11.16	8.84	6.40	5.17	4.00	2.90	1.93	1.58	1.27	0.98	0.73	0.51	0.33	0.19	0.08	0.05	0.02
10.00	31.62	23.72	15.80	12.56	9.04	7.08	5.02	4.00	3.03	2.14	1.38	1.12	0.88	0.68	0.50	0.35	0.22	0.13	0.06	0.03	0.01
7.50	27.39	20.54	13.68	10.87	7.78	6.04	4.22	3.32	2.47	1.71	1.08	0.87	0.68	0.52	0.38	0.26	0.17	0.09	0.04	0.02	0.01
5.00	22.36	16.77	11.17	8.86	6.29	4.83	3.30	2.55	1.85	1.25	0.76	0.60	0.47	0.35	0.26	0.18	0.11	0.06	0.03	0.02	0.01
2.50	15.81	11.86	7.90	6.24	4.36	3.27	2.14	1.60	1.11	0.71	0.41	0.32	0.24	0.18	0.13	0.09	0.06	0.03	0.01	0.01	0.00
1.25	11.18	8.39	5.58	4.40	3.01	2.20	1.37	0.99	0.66	0.40	0.22	0.17	0.12	0.09	0.07	0.04	0.03	0.02	0.01	0.00	0.00
1.00	10.00	7.50	4.99	3.93	2.67	1.93	1.18	0.84	0.55	0.33	0.18	0.13	0.10	0.07	0.05	0.04	0.02	0.01	0.01	0.00	0.00
0.75	8.66	6.50	4.32	3.39	2.28	1.63	0.98	0.68	0.44	0.25	0.13	0.10	0.08	0.06	0.04	0.03	0.02	0.01	0.00	0.00	0.00
0.50	7.07	5.30	3.53	2.76	1.83	1.28	0.74	0.51	0.31	0.18	0.09	0.07	0.05	0.04	0.03	0.02	0.01	0.01	0.00	0.00	0.00
0.25	5.00	3.75	2.49	1.94	1.25	0.84	0.46	0.30	0.18	0.09	0.05	0.03	0.03	0.02	0.01	0.01	0.01	0.00	0.00	0.00	0.00
0.12	3.54	2.65	1.76	1.36	0.85	0.55	0.28	0.17	0.10	0.05	0.02	0.02	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00

Ratio $S^{\mathrm{OPT}}/S^{\star}$

$\rho \times 100$ \ $\mu/\sigma \times \sqrt{12}$	10.95	8.22	5.48	4.38	3.29	2.74	2.19	1.92	1.64	1.37	1.10	0.99	0.88	0.77	0.66	0.55	0.44	0.33	0.22	0.16	0.11
50.00	1.00	1.00	1.00	1.00	0.98	0.96	0.92	0.88	0.84	0.79	0.71	0.67	0.63	0.58	0.52	0.45	0.37	0.29	0.20	0.15	0.10
25.00	1.00	1.00	1.00	1.00	0.97	0.93	0.86	0.81	0.75	0.68	0.58	0.54	0.49	0.44	0.39	0.33	0.27	0.21	0.14	0.11	0.07
20.00	1.00	1.00	1.00	1.00	0.97	0.92	0.85	0.79	0.72	0.64	0.54	0.50	0.46	0.41	0.36	0.30	0.25	0.19	0.13	0.09	0.06
15.00	1.00	1.00	1.00	0.99	0.96	0.91	0.83	0.76	0.69	0.60	0.50	0.45	0.41	0.36	0.31	0.27	0.21	0.16	0.11	0.08	0.05
10.00	1.00	1.00	1.00	0.99	0.95	0.90	0.79	0.72	0.64	0.54	0.44	0.39	0.35	0.31	0.26	0.22	0.18	0.13	0.09	0.07	0.04
7.50	1.00	1.00	1.00	0.99	0.95	0.88	0.77	0.69	0.60	0.50	0.39	0.35	0.31	0.27	0.23	0.19	0.15	0.12	0.08	0.06	0.04
5.00	1.00	1.00	1.00	0.99	0.94	0.86	0.74	0.65	0.55	0.45	0.34	0.30	0.26	0.23	0.19	0.16	0.13	0.09	0.06	0.05	0.03
2.50	1.00	1.00	1.00	0.99	0.92	0.83	0.68	0.58	0.47	0.36	0.26	0.22	0.19	0.16	0.14	0.11	0.09	0.07	0.04	0.03	0.02
1.25	1.00	1.00	1.00	0.98	0.90	0.79	0.61	0.50	0.39	0.28	0.19	0.16	0.14	0.12	0.10	0.08	0.06	0.05	0.03	0.02	0.02
1.00	1.00	1.00	1.00	0.98	0.89	0.77	0.59	0.48	0.37	0.26	0.18	0.15	0.13	0.11	0.09	0.07	0.06	0.04	0.03	0.02	0.01
0.75	1.00	1.00	1.00	0.98	0.88	0.75	0.56	0.45	0.34	0.23	0.15	0.13	0.11	0.09	0.08	0.06	0.05	0.04	0.02	0.02	0.01
0.50	1.00	1.00	1.00	0.98	0.86	0.73	0.52	0.41	0.30	0.20	0.13	0.11	0.09	0.08	0.06	0.05	0.04	0.03	0.02	0.02	0.01
0.25	1.00	1.00	1.00	0.97	0.83	0.68	0.46	0.34	0.24	0.15	0.09	0.08	0.06	0.05	0.04	0.04	0.03	0.02	0.01	0.01	0.01
0.12	1.00	1.00	1.00	0.96	0.80	0.62	0.39	0.28	0.18	0.11	0.07	0.05	0.05	0.04	0.03	0.03	0.02	0.02	0.01	0.01	0.01

Optimal Feasible Strategy

- The optimal feasible strategy is simply $w^{\mathrm{OPT}} = \mathrm{E}(\alpha|\mathscr{G})$, which has closed-form given Bayes' rule, if arbitrageurs know the distribution of alphas.
- In practice we do not know the distribution.
- In this case a feasible strategy $\widehat{w}^{\mathrm{OPT}}$ can be constructed using Empirical Bayes via Tweedie's formula:

$$\widehat{w}_i^{\mathrm{OPT}} = \widehat{\alpha}_i + \underbrace{\frac{1}{T}\frac{d\log\widehat{p}(a)}{da}\bigg|_{a=\widehat{\alpha}_i}}_{\text{Bayes Shrinkage}},$$

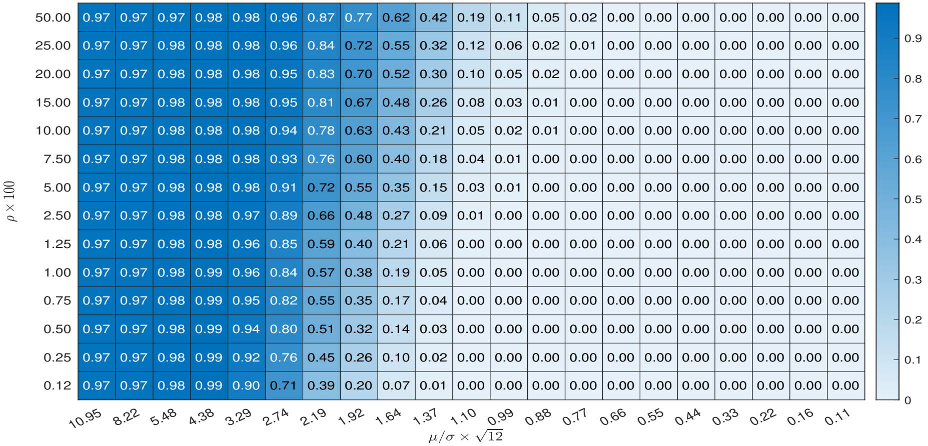
  where $\widehat{p}(a)$ is a nonparametric estimator of the marginal density of $\widehat{\alpha}_i$.
- We show the strategy $\widehat{w}^{\mathrm{OPT}}$ achieves the optimal feasible Sharpe S^{OPT} under (almost) arbitrary alpha distributions.

Ratio $S^{\mathrm{CSR}} / S^{\mathrm{OPT}}$

$\rho \times 100$ \ $\mu/\sigma \times \sqrt{12}$	10.95	8.22	5.48	4.38	3.29	2.74	2.19	1.92	1.64	1.37	1.10	0.99	0.88	0.77	0.66	0.55	0.44	0.33	0.22	0.16	0.11
50.00	1.00	0.99	0.98	0.97	0.97	0.97	0.98	0.98	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
25.00	0.99	0.98	0.96	0.95	0.93	0.93	0.95	0.96	0.97	0.98	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
20.00	0.99	0.98	0.95	0.93	0.92	0.91	0.93	0.94	0.95	0.97	0.98	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
15.00	0.98	0.97	0.94	0.91	0.89	0.88	0.89	0.91	0.92	0.94	0.96	0.97	0.98	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00
10.00	0.98	0.96	0.91	0.88	0.84	0.83	0.84	0.85	0.87	0.90	0.94	0.95	0.96	0.98	0.98	0.99	1.00	1.00	1.00	1.00	1.00
7.50	0.97	0.95	0.89	0.85	0.80	0.79	0.79	0.81	0.83	0.87	0.92	0.93	0.95	0.97	0.98	0.99	1.00	1.00	1.00	1.00	1.00
5.00	0.95	0.92	0.85	0.79	0.73	0.72	0.73	0.74	0.78	0.82	0.89	0.91	0.94	0.96	0.97	0.99	0.99	1.00	1.00	1.00	1.00
2.50	0.91	0.86	0.75	0.68	0.61	0.59	0.60	0.63	0.68	0.75	0.84	0.88	0.91	0.94	0.97	0.98	0.99	1.00	1.00	1.00	1.00
1.25	0.85	0.76	0.62	0.54	0.48	0.47	0.49	0.53	0.59	0.68	0.80	0.85	0.90	0.94	0.96	0.98	0.99	1.00	1.00	1.00	1.00
1.00	0.82	0.73	0.58	0.50	0.44	0.43	0.46	0.50	0.57	0.67	0.80	0.85	0.90	0.93	0.96	0.98	0.99	1.00	1.00	1.00	1.00
0.75	0.77	0.68	0.52	0.45	0.39	0.39	0.42	0.46	0.54	0.65	0.79	0.84	0.89	0.93	0.96	0.98	0.99	1.00	1.00	1.00	1.00
0.50	0.71	0.60	0.45	0.38	0.33	0.33	0.37	0.42	0.50	0.62	0.77	0.83	0.89	0.93	0.96	0.98	0.99	1.00	1.00	1.00	1.00
0.25	0.58	0.47	0.33	0.28	0.25	0.26	0.31	0.36	0.45	0.59	0.76	0.83	0.88	0.93	0.96	0.98	0.99	1.00	1.00	1.00	1.00
0.12	0.45	0.35	0.24	0.20	0.19	0.20	0.25	0.31	0.41	0.56	0.75	0.82	0.88	0.93	0.96	0.98	0.99	1.00	1.00	1.00	1.00

- Strategy: $\widehat{w} = \widehat{\alpha}$.

Ratio $S^{\text{BH}}/S^{\text{OPT}}$



- False Discovery Rate Control Strategy: applying Benjamini-Hochberg algo onto $\widehat{\alpha}_i$s.

Ratio $S^{\text{LASSO}} / S^{\text{OPT}}$

$\rho \times 100$ \ $\mu/\sigma \times \sqrt{12}$	10.95	8.22	5.48	4.38	3.29	2.74	2.19	1.92	1.64	1.37	1.10	0.99	0.88	0.77	0.66	0.55	0.44	0.33	0.22	0.16	0.11
50.00	1.00	0.99	0.98	0.98	0.97	0.97	0.98	0.98	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
25.00	1.00	0.99	0.98	0.97	0.97	0.97	0.98	0.98	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
20.00	1.00	0.99	0.98	0.97	0.97	0.97	0.98	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
15.00	1.00	0.99	0.98	0.97	0.96	0.97	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
10.00	1.00	0.99	0.98	0.97	0.96	0.97	0.98	0.99	0.99	0.98	0.98	0.98	0.98	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00
7.50	1.00	0.99	0.98	0.97	0.96	0.97	0.98	0.99	0.99	0.98	0.97	0.97	0.98	0.98	0.99	0.99	1.00	1.00	1.00	1.00	1.00
5.00	1.00	0.99	0.98	0.97	0.96	0.97	0.98	0.99	0.98	0.97	0.96	0.96	0.97	0.98	0.98	0.99	1.00	1.00	1.00	1.00	1.00
2.50	1.00	0.99	0.98	0.96	0.96	0.97	0.98	0.98	0.97	0.95	0.94	0.95	0.96	0.97	0.98	0.99	1.00	1.00	1.00	1.00	1.00
1.25	1.00	0.99	0.97	0.96	0.96	0.97	0.98	0.98	0.96	0.93	0.92	0.93	0.95	0.96	0.98	0.99	0.99	1.00	1.00	1.00	1.00
1.00	1.00	0.99	0.97	0.96	0.96	0.97	0.98	0.97	0.96	0.93	0.92	0.93	0.94	0.96	0.98	0.99	0.99	1.00	1.00	1.00	1.00
0.75	1.00	0.99	0.97	0.96	0.96	0.97	0.98	0.97	0.95	0.92	0.91	0.92	0.94	0.96	0.98	0.99	0.99	1.00	1.00	1.00	1.00
0.50	1.00	0.99	0.97	0.96	0.96	0.97	0.98	0.97	0.94	0.90	0.90	0.92	0.94	0.96	0.97	0.99	0.99	1.00	1.00	1.00	1.00
0.25	1.00	0.99	0.97	0.96	0.96	0.98	0.97	0.95	0.92	0.88	0.89	0.91	0.93	0.96	0.97	0.99	0.99	1.00	1.00	1.00	1.00
0.12	1.00	0.99	0.97	0.95	0.96	0.98	0.97	0.94	0.89	0.86	0.89	0.91	0.93	0.95	0.97	0.99	0.99	1.00	1.00	1.00	1.00

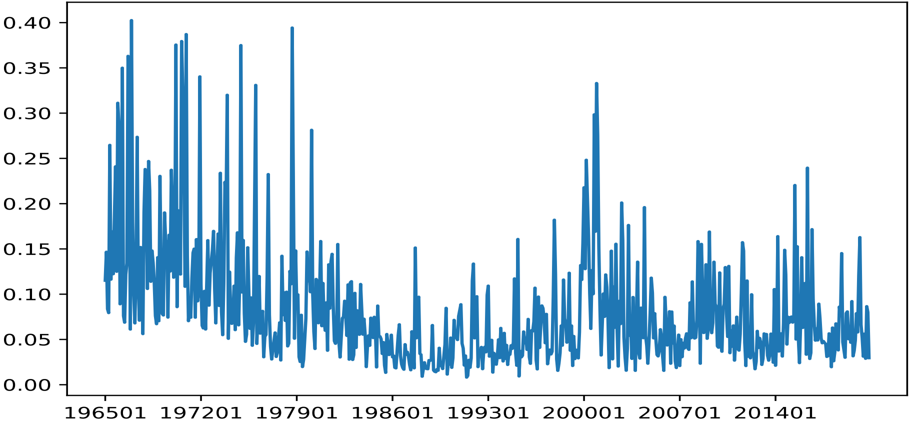
Colour scale: 0.88, 0.9, 0.92, 0.94, 0.96, 0.98

- Lasso Strategy: applying Lasso algo onto $\widehat{\alpha}_i$s.

Empirical Analysis of US Equities

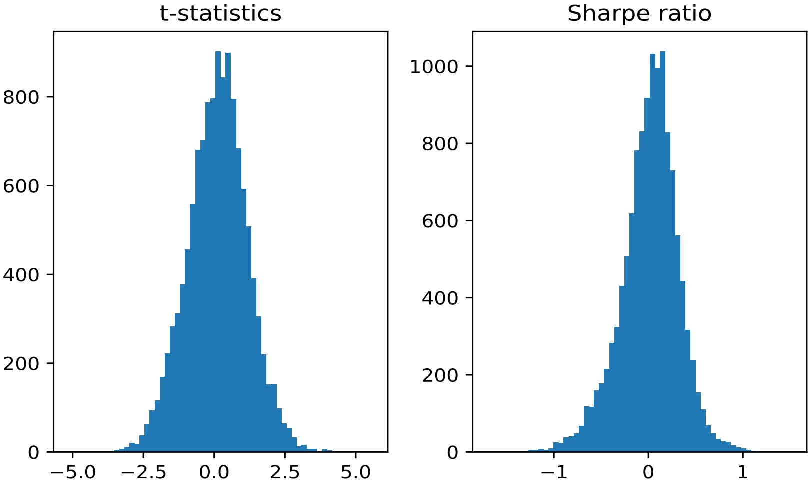
- We study US monthly equity returns from January 1965 to December 2020.
- We adopt a multi-factor model with 16 characteristics and 11 GICS sectors, including market beta, size, operating profits/book equity, book equity/market equity, asset growth, momentum, short-term reversal, industry momentum, illiquidity, leverage, return seasonality, sales growth, accruals, dividend yield, tangibility, and idiosyncratic risk.
- 10-year rolling window estimation, last 2 years as validation sample for tuning parameter selection

Time-series of the Cross-sectional R^2s



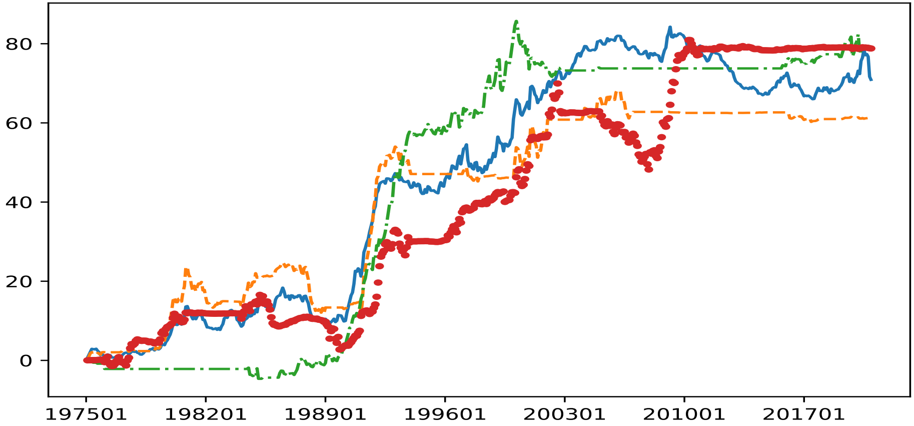
Lewellen (2015, CFR): 7.8%, 1964/05 - 2009/12; Gu, Kelly, and Xiu (2021, JoE): 12-14% 1987 - 2016

Rare and Weak Alphas



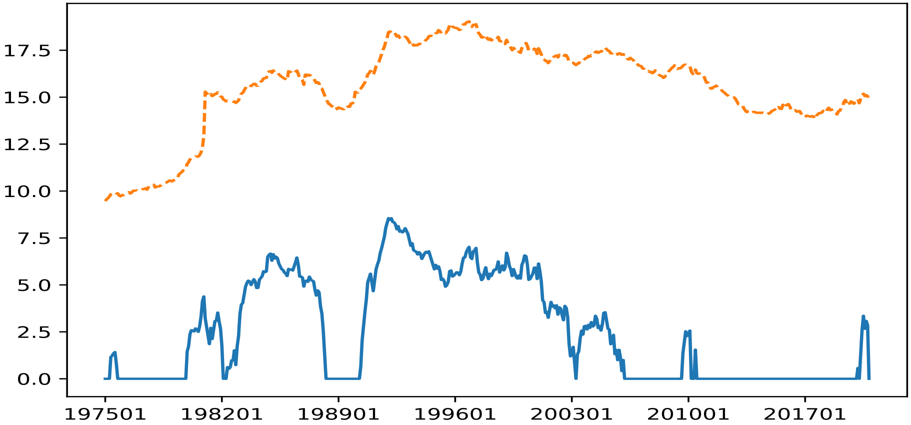
6.35% (0.63%) of the t-Stats > 2.0 (3.0); 0.505% alphas with a Sharpe ratio > 1.0.

Performance of Risk-Normalized Arbitrage Portfolios



Sharpe Ratios: OPT (red, 0.496), CSR (blue, 0.450), BH (green, 0.497), and LASSO (orange, 0.384)
In contrast, average of $\widehat{S}^{\star}$ is about 2.95, which is far greater than feasible Sharpe ratios, ~ 0.5.

Biased vs. Consistent Sharpe Ratio Estimators



$\widehat{S}^\star$ (blue) and $\widetilde{S}^\star$ (orange)

Average of $\widehat{S}^\star$ is about 2.95, which is far greater than feasible Sharpe ratios, ~ 0.5.

Conclusion

- Statistical limit to arbitrage: Widens the bounds in which mispricing can survive in presence of arbitrageurs.
- Existing empirical evidence provides "lower bound" of Sharpe ratios achievable with machine learning methods (based on ad-hoc choices). Our theoretical analysis provides an "upper bound" in a specific context (based on optimal strategy).
- The gap between feasible and infeasible Sharpe ratios will further increase if arbitrageurs face additional statistical challenges, e.g., model misspecification, omitted factors, weak factors, large non-sparse covariance matrix.