Machine Learning versus Economic Restrictions: Evidence from Stock Return Predictability

Doron Avramov, Si Cheng, and Lior Metzker

Working Paper, 2021.4

Lv Manni

2021.7.8

Contents

- Introduction
  - Background & Motivation
  - Research Problem
  - Contribution
- Model Setting
- Empirical Results
- Conclusion
- Extension

Backgrounds & Motivation

- Recent work has challenged the credibility of the predictable patterns of a plethora of anomalies. However, anomalous return patterns characterize expensive-to-buy and difficult-to-arbitrage stocks. Notably, it is increasingly difficult to exploit anomalies in recent years due to increased market liquidity and arbitrage activity.

- Counter to this “anomaly challenging” strand of literature, there has been an emerging body of work that reports phenomenal investment profitability based on signals generated by ML methods. Understanding the relevant economic mechanisms is essential.

➢ Whether ML methods clear sensible economic restrictions in empirical finance and face the same challenges?

➢ Is there any economic grounds of investment decisions advocated by the seemingly opaque ML methods?

Research Problem

- Whether machine learning methods sensible economic restrictions in empirical finance?
  - ➢ We find that machine learning methods often fail to clear standard economic restrictions, such as VW returns and excluding microcaps or distressed firms (expensive-to-buy).
  - ➢ The trading strategy is more profitable during periods of increasing limits to arbitrage, such as high investor sentiment, high market volatility, and low market liquidity (hard-to-arbitrage).
  - ➢ Given the high turnover and transaction costs, the strategy is hard to leave alpha on the table.
- Is there any economic grounds of investment decisions advocated by the seemingly opaque machine learning methods?
  - ➢ We find DL signals identify stocks in line with most anomaly-based trading strategies.

Contribution

- Our analysis initiates a protocol for future work to demonstrate the feasibility of trading profits, such as excluding difficult-to-arbitrage stocks and high limits-to-arbitrage market states, considering portfolio turnover and the corresponding transaction costs.
- Our paper provides evidence for the economic constraints and economic foundations of ML methods, enriches the academic and policy discussions surrounding the adoption of ML techniques in asset management, including the effectiveness and sustainability of new trading signals, the lack of transparency and economic interpretability in complex machine learning algorithms.

Outline

five ML models

ML and economic restrictions

cross-section: cheap to trade

- excluding microcaps
- only rated firms
- rated, exclude distressed firms

time series: limits-to-arbitrage

- high volatility
- high sentiment
- low liquidity

turnover and transaction costs

- leave alpha or not

ML and economic foundations

Whether stocks with similar ML signals share same characteristics that predict future returns

Model Setting

- two deep learning methods, nonlinear methods
  - GKX: 32-16-8 NN3, follow Gu, Kelly, and Xiu (2020)
  - CPZ: incorporate a no-arbitrage condition into multiple connected neural networks, including feed forward networks (FFNs), recurrent neural networks (RNNs) with long short-term memory (LSTM) cells, and a generative adversarial network (GAN) to estimate the SDF and its stock loadings, follow Chen, Pelger, and Zhu (2020)

- two conditional beta pricing models
  - IPCA: IPCA, follow Kelly, Pruitt, and Su (2019), linear
  - CA: the conditional autoencoder, extended from Gu, Kelly, and Xiu (2021), nonlinear ML methods

- method focused on equity portfolios
  - KNS: SDF and MVE portfolio, follow Kozak, Nagel, and Santosh (2020)

Model Setting

Sample: all NYSE/AMEX/Nasdaq stocks

- GKX, IPCA, and CA methods: 920 predictors, 1957 to 2017

➢ 94 firm characteristics that have been documented as anomalies+ 74 industry dummies based on the first two digits of SIC codes + 8 monthly macroeconomic predictors + 94 × 8 interactions between characteristics & macroeconomic variable

✓ for GKX and CA: 18 (train: 1957 to 1974) + 12 (validation:1975 to 1986) + 31 (test: 1987 to 2017), expanding

✓ for IPCA: skip the validation procedure and requires at least 120 months for the in-sample estimation, and forward rolling is performed on a monthly basis

- CPZ method: use 46 firm characteristics, 178 macroeconomic predictors, nonlinear interactions; 1967 to 2016; 20 (train: 1967 to 1986) + 5 (validation: 1987 to 1991) + 25 (test: 1992 to 2016), expanding

Model Setting

- Data Source:

➢ daily and monthly stock data obtained from the CRSP

➢ quarterly / annual financial data come from the COMPUSTAT database

- Subsamples with Economic Restrictions

1. excludes microcaps (stocks with a market capitalization smaller than the 20th NYSE size percentile)
2. includes only rated firms (firms with data on S&P long-term issuer credit ratings)
3. filter downgraded firms on the universe of rated firms

- Portfolios sort - using the four proposed ML signals

Empirical Results

- the standard OLS regression methodology with all 920 predictors used in the GKX estimation to predict the one-month-ahead stock return
- portfolio sorts while combining all the signals on individual anomalies

Panel A: Returns to Investment Strategies Sorted by OLS-Predicted Returns

	Equal-Weighted							Value-Weighted						
	Return	CAPM	FFC	FFC+PS	FF5	FF6	SY	Return	CAPM	FFC	FFC+PS	FF5	FF6	SY
Full Sample	1.215***	1.162***	1.268***	1.272***	1.521***	1.545***	1.475***	0.084	0.118	0.148	0.164	0.308	0.268	0.127
	(5.25)	(4.78)	(5.32)	(5.28)	(5.95)	(5.85)	(4.69)	(0.39)	(0.50)	(0.59)	(0.64)	(1.09)	(0.92)	(0.37)
Nonmicrocaps	0.472**	0.410*	0.368*	0.367*	0.565**	0.514**	0.409	0.013	0.060	0.118	0.128	0.237	0.223	0.074
	(2.37)	(1.84)	(1.72)	(1.66)	(2.42)	(2.19)	(1.47)	(0.07)	(0.30)	(0.58)	(0.61)	(1.04)	(0.95)	(0.25)
Credit Rating Sample	0.437*	0.361	0.503**	0.488**	0.458**	0.563**	0.632**	0.024	0.061	0.126	0.135	0.103	0.115	0.133
	(1.78)	(1.50)	(2.12)	(2.02)	(2.00)	(2.33)	(2.02)	(0.12)	(0.27)	(0.58)	(0.61)	(0.46)	(0.49)	(0.51)
Nondowngrades	0.506**	0.480**	0.542**	0.506**	0.412*	0.470**	0.550**	-0.008	0.035	0.092	0.078	0.082	0.065	0.032
	(2.39)	(2.13)	(2.41)	(2.29)	(1.96)	(2.08)	(1.97)	(-0.04)	(0.15)	(0.41)	(0.34)	(0.34)	(0.26)	(0.12)

Panel B: Returns to Investment Strategies Sorted by Individual Anomalies

	Equal-Weighted							Value-Weighted						
	Return	CAPM	FFC	FFC+PS	FF5	FF6	SY	Return	CAPM	FFC	FFC+PS	FF5	FF6	SY
Full Sample	0.366***	0.472***	0.342***	0.346***	0.308***	0.250***	0.221***	0.192***	0.306***	0.185***	0.189***	0.139***	0.092***	0.043
	(5.92)	(8.18)	(7.65)	(7.91)	(6.03)	(6.07)	(3.95)	(2.91)	(5.41)	(4.72)	(4.88)	(3.27)	(2.89)	(1.19)
Nonmicrocaps	0.233***	0.344***	0.214***	0.218***	0.167***	0.115***	0.079**	0.170***	0.272***	0.150***	0.151***	0.095**	0.049*	-0.003
	(3.71)	(6.11)	(5.98)	(6.21)	(4.40)	(4.60)	(2.16)	(2.74)	(4.87)	(3.97)	(4.02)	(2.57)	(1.87)	(-0.09)
Credit Rating Sample	0.302***	0.387***	0.263***	0.263***	0.244***	0.182***	0.121**	0.131**	0.215***	0.113***	0.111***	0.092**	0.045	-0.017
	(5.11)	(6.78)	(6.14)	(6.24)	(4.59)	(4.56)	(2.55)	(2.47)	(4.75)	(3.27)	(3.11)	(2.20)	(1.42)	(-0.52)
Nondowngrades	0.124***	0.197***	0.100***	0.103***	0.083**	0.039	-0.017	0.071	0.142***	0.045	0.046	0.038	-0.004	-0.071*
	(2.60)	(4.45)	(3.14)	(3.31)	(2.11)	(1.22)	(-0.44)	(1.48)	(3.28)	(1.43)	(1.44)	(0.93)	(-0.13)	(-1.93)

Empirical Results – NN3

Panel A: Returns to Investment Strategies Sorted by NN3-Predicted Returns

Equal-Weighted							Value-Weighted						
Return	CAPM	FFC	FFC+PS	FF5	FF6	SY	Return	CAPM	FFC	FFC+PS	FF5	FF6	SY
2.470***	2.742***	2.483***	2.497***	2.349***	2.236***	2.312***	1.556***	1.894***	1.338***	1.361***	1.206***	0.916***	0.769***
(9.03)	(9.78)	(9.02)	(9.27)	(9.35)	(8.06)	(7.91)	(4.53)	(5.64)	(5.14)	(5.31)	(4.66)	(4.08)	(3.03)

Panel B: Returns to Investment Strategies Sorted by NN3-Predicted Returns (Nonmicrocaps)

Equal-Weighted							Value-Weighted						
Return	CAPM	FFC	FFC+PS	FF5	FF6	SY	Return	CAPM	FFC	FFC+PS	FF5	FF6	SY
1.243***	1.625***	0.924***	0.940***	0.680***	0.353**	0.357	1.047***	1.389***	0.771***	0.771***	0.627**	0.312	0.179
(3.75)	(5.04)	(4.08)	(4.19)	(2.89)	(2.02)	(1.37)	(3.24)	(4.43)	(3.23)	(3.24)	(2.41)	(1.51)	(0.73)

Panel A: Returns to Investment Strategies Sorted by NN3-Predicted Returns (Credit Rating Sample)

Equal-Weighted							Value-Weighted						
Return	CAPM	FFC	FFC+PS	FF5	FF6	SY	Return	CAPM	FFC	FFC+PS	FF5	FF6	SY
1.374***	1.759***	1.073***	1.083***	1.048***	0.652***	0.548*	1.024***	1.344***	0.746***	0.743***	0.784***	0.433**	0.150
(4.08)	(5.75)	(4.62)	(4.59)	(3.63)	(3.16)	(1.93)	(3.18)	(4.40)	(3.31)	(3.23)	(2.70)	(2.05)	(0.59)

Panel B: Returns to Investment Strategies Sorted by NN3-Predicted Returns (Nondowngrades)

Equal-Weighted							Value-Weighted						
Return	CAPM	FFC	FFC+PS	FF5	FF6	SY	Return	CAPM	FFC	FFC+PS	FF5	FF6	SY
0.667**	0.997***	0.343	0.362*	0.323	-0.026	-0.205	0.723**	0.996***	0.424*	0.445**	0.517*	0.204	-0.129
(2.11)	(3.40)	(1.63)	(1.70)	(1.32)	(-0.13)	(-0.73)	(2.49)	(3.40)	(1.93)	(2.00)	(1.87)	(0.92)	(-0.52)

ML substantially improves the investment payoff compared to traditional methods. Both deliver lower payoffs in the presence of economic restrictions with a similar proportional magnitude.

Empirical Results - CPZ

Panel A: Returns to Investment Strategies Sorted by Risk Loadings

Equal-Weighted							Value-Weighted						
Return	CAPM	FFC	FFC+PS	FF5	FF6	SY	Return	CAPM	FFC	FFC+PS	FF5	FF6	SY
3.446***	3.415***	3.471***	3.473***	3.465***	3.492***	3.566***	2.183***	2.056***	1.910***	1.918***	1.869***	1.867***	1.983***
(11.25)	(11.17)	(8.87)	(9.28)	(8.94)	(8.04)	(8.01)	(6.37)	(5.68)	(5.42)	(5.66)	(5.29)	(4.86)	(5.30)

Panel B: Returns to Investment Strategies Sorted by Risk Loadings (Nonmicrocaps)

Equal-Weighted							Value-Weighted						
Return	CAPM	FFC	FFC+PS	FF5	FF6	SY	Return	CAPM	FFC	FFC+PS	FF5	FF6	SY
1.586***	1.645***	1.325***	1.332***	1.306***	1.161***	1.218***	1.083***	1.083***	0.688***	0.655***	0.720***	0.548**	0.548**
(6.80)	(6.98)	(5.86)	(6.09)	(6.42)	(5.13)	(5.06)	(4.28)	(4.06)	(2.87)	(2.79)	(2.93)	(2.23)	(2.27)

Panel A: Returns to Investment Strategies Sorted by Risk Loadings (Credit Rating Sample)

Equal-Weighted							Value-Weighted						
Return	CAPM	FFC	FFC+PS	FF5	FF6	SY	Return	CAPM	FFC	FFC+PS	FF5	FF6	SY
1.519***	1.522***	1.337***	1.340***	1.376***	1.270***	1.257***	0.812***	0.806***	0.480*	0.420	0.559*	0.423	0.339
(5.74)	(5.83)	(4.36)	(4.44)	(4.89)	(3.93)	(3.77)	(2.83)	(2.68)	(1.71)	(1.51)	(1.88)	(1.46)	(1.18)

Panel B: Returns to Investment Strategies Sorted by Risk Loadings (Nondowngrades)

Equal-Weighted							Value-Weighted						
Return	CAPM	FFC	FFC+PS	FF5	FF6	SY	Return	CAPM	FFC	FFC+PS	FF5	FF6	SY
1.413***	1.444***	1.131***	1.141***	1.214***	1.041***	0.994***	0.915***	0.969***	0.652**	0.597**	0.696**	0.574*	0.511
(4.85)	(5.15)	(3.78)	(3.88)	(4.53)	(3.50)	(2.88)	(2.91)	(3.01)	(2.15)	(2.00)	(2.14)	(1.82)	(1.64)

As before, machine learning-based investment payoffs deteriorate in the presence of sensible economic restrictions(60%).

Empirical Results – IPCA & CA

Panel A: Value-Weighted Returns to Investment Strategies Sorted by IPCA-Predicted Returns							
Full Sample		Nonmicrocaps		Credit Rating Sample		Nondowngrades	
Return	FF6	Return	FF6	Return	FF6	Return	FF6
0.945***	0.624***	0.901***	0.613***	0.893***	0.607***	0.733***	0.430**
(5.62)	(3.31)	(5.57)	(3.71)	(5.08)	(3.38)	(4.08)	(2.37)

Panel B: Value-Weighted Returns to Investment Strategies Sorted by CA2-Predicted Returns							
Full Sample		Nonmicrocaps		Credit Rating Sample		Nondowngrades	
Return	FF6	Return	FF6	Return	FF6	Return	FF6
1.159***	0.746***	1.105***	0.387**	0.874***	$\boxed{0.187}$	0.665**	$\boxed{0.048}$
(4.17)	(3.01)	(4.22)	(2.03)	(2.97)	(0.79)	(2.22)	(0.20)

Collectively, ML return patterns are concentrated in stocks that are relatively difficult to value and difficult to arbitrage.

Although IPCA underperforms for the full sample, it delivers consistent risk-adjusted performance even in the presence of economic restrictions.

All ML methods substantially improve the investment payoff compared to the traditional methods with and without economic restrictions.

Empirical Results

	Sharpe Ratio	Skewness	Excess Kurtosis	Maximum Drawdown	Return in Crisis	Turnover
Panel A: Sorted by NN3-Predicted Returns						
Full Sample	0.944	0.631	5.222	0.350	4.100	0.976
Nonmicrocaps	0.644	0.361	7.062	0.349	3.563	0.869
Credit Rating Sample	0.639	0.064	7.875	0.420	3.435	0.889
Nondowngrades	0.449	0.146	8.550	0.333	2.931	0.920
Panel B: Sorted by Risk Loadings						
Full Sample	1.225	1.063	5.932	0.209	0.472	1.664
Nonmicrocaps	0.839	0.326	1.582	0.246	0.677	1.625
Credit Rating Sample	0.566	0.267	1.440	0.407	-0.023	1.652
Nondowngrades	0.602	0.344	1.675	0.447	0.903	1.678
Panel C: Sorted by IPCA-Predicted Returns						
Full Sample	0.967	-0.449	4.805	0.203	0.574	1.186
Nonmicrocaps	0.978	-0.267	5.369	0.234	1.493	1.130
Credit Rating Sample	0.880	-0.219	4.221	0.315	0.474	1.164
Nondowngrades	0.697	-0.069	3.158	0.349	-0.640	1.184
Panel D: Sorted by CA2-Predicted Returns						
Full Sample	0.784	-0.077	2.418	0.202	-0.047	1.565
Nonmicrocaps	0.748	0.291	4.684	0.207	-0.529	1.478
Credit Rating Sample	0.522	-0.471	4.119	0.252	-1.796	1.542
Nondowngrades	0.387	-0.616	4.930	0.345	-0.167	1.571
Panel E: Market Portfolio						
Full Sample	0.527	-0.978	3.323	0.486	-6.954	0.089
Nonmicrocaps	0.530	-0.959	3.222	0.485	-6.907	0.086
Credit Rating Sample	0.543	-0.932	3.423	0.498	-6.747	0.080
Nondowngrades	0.682	-0.856	3.311	0.408	-6.615	0.084

全样本 盈亏平衡 交易成本	子样本 盈亏平衡 交易成本
0.94%	0.36%
1.12%	0.34%
0.53%	0.54%
0.48%	0.26%

Empirical Results

- Imposing economic restrictions significantly reduces the Sharpe ratio.
- ML methods are more positively skewed than the market portfolio.
- ML methods experience comparatively smaller drawdowns, are able to mitigate downside risk and protect investors from extreme crashes.
- Accounting for reasonable transaction costs would make it difficult for most ML signals to leave alpha on the table.

Empirical Results

Similar to CPZ, the KNS approach incorporates a no-arbitrage condition to estimate the SDF, while focus on equity portfolios that represent characteristics-based trading strategies.

train: September 1964 to December 2004

test: January 2005 to December 2017

Characteristics of SDF-Implied MVE Portfolios

	CAPM	FF6	Sharpe Ratio	SDF-Implied MVE Portfolio Weights										
				Mean	Std.Dev.	Min	5%	10%	25%	Median	75%	90%	95%	Max
Full Sample	3.662***	3.338***	2.318	0.083	1.338	-2.994	-2.343	-1.994	-0.912	0.341	0.964	1.687	1.895	3.182
	(6.01)	(5.90)												
Nonmicrocaps	1.543***	0.895***	0.977	0.084	0.447	-0.863	-0.666	-0.592	-0.238	0.072	0.407	0.647	0.741	1.431
	(3.88)	(2.87)												
Credit Rating Sample	1.418***	0.717*	0.898	-0.006	0.254	-0.678	-0.473	-0.382	-0.137	-0.003	0.187	0.326	0.387	0.419
	(2.97)	(1.93)												
Nondowngrades	1.308***	0.545	0.828	-0.022	0.248	-0.543	-0.448	-0.370	-0.217	0.004	0.135	0.293	0.376	0.595
	(2.92)	(1.59)												

Empirical Results

- We analyze the out-of-sample trading profits using predictive signals generated from five machine learning methods in subsamples with economic restrictions.

- IPCA underperforms deep learning models for the full sample but does not display a material deterioration of performance in subsamples with economic restrictions. Conversely, all other ML methods are opposite.

- As machine learning-based trading strategies require relatively high portfolio turnover, investors may need to further lower their expectations of achievable performance.

Empirical Results - Time-Varying Return Predictability

market state variables:

(1) investor sentiment (SENT), defined as the monthly Baker and Wurgler (2007) investor sentiment;

(2) realized market volatility(MKTVOL), defined as the standard deviation of daily CRSP VW index returns in a month;

(3) implied market volatility (VIX), defined as the monthly VIX index of implied volatilities of S&P 500 index options

(4) market illiquidity (MKTILLIQ), defined as the VW average of stock-level Amihud (2002) illiquidity for all stocks in a month

Empirical Results - Time-Varying Return Predictability

Value-Weighted FF6-Adjusted Returns of Machine Learning Portfolios								
	SENT		MKTVOL		VIX		MKTILLIQ	
	Low	High	Low	High	Low	High	Low	High
Panel A: Sorted by NN3-Predicted Returns								
Full Sample	0.732***	0.879**	0.641**	1.283***	0.218	1.662***	0.736***	1.075***
	(3.00)	(2.40)	(2.34)	(3.96)	(0.95)	(4.16)	(2.70)	(3.36)
Nonmicrocaps	0.334	0.095	0.062	0.814***	-0.050	0.747**	0.222	0.461
	(1.45)	(0.29)	(0.26)	(2.87)	(-0.20)	(2.03)	(0.84)	(1.51)
Credit Rating Sample	0.389	0.274	0.095	0.896***	0.025	0.808*	0.512**	0.351
	(1.49)	(0.91)	(0.44)	(2.93)	(0.11)	(1.96)	(1.97)	(1.15)
Nondowngrades	0.355	-0.161	0.054	0.597*	-0.060	0.582	0.294	0.120
	(1.43)	(-0.50)	(0.23)	(1.70)	(-0.24)	(1.25)	(1.21)	(0.36)
Panel B: Sorted by Risk Loadings								
Full Sample	1.372***	2.453***	1.352***	2.364***	0.898**	2.451***	1.150***	2.128***
	(3.83)	(3.78)	(3.64)	(4.09)	(2.30)	(3.80)	(2.73)	(4.53)
Nonmicrocaps	0.512**	0.681	0.423	0.665	0.321	0.386	0.527	0.432*
	(2.14)	(1.60)	(1.41)	(1.62)	(1.14)	(1.01)	(1.45)	(1.76)
Credit Rating Sample	0.269	0.658	0.448	0.527	0.097	0.337	0.691	0.073
	(0.84)	(1.35)	(1.12)	(1.12)	(0.31)	(0.69)	(1.63)	(0.20)
Nondowngrades	0.329	0.945*	0.453	0.699	0.031	0.651	0.847*	0.155
	(0.93)	(1.79)	(0.96)	(1.45)	(0.09)	(1.19)	(1.86)	(0.39)

Empirical Results - Time-Varying Return Predictability

	SENT		MKTVOL		VIX		MKTILLIQ	
	Low	High	Low	High	Low	High	Low	High
Panel C: Sorted by IPCA-Predicted Returns								
Full Sample	0.558**	0.667**	0.590***	0.694**	0.440**	0.763**	0.603**	0.645***
	(2.46)	(2.03)	(3.57)	(2.56)	(2.60)	(2.23)	(2.39)	(2.78)
Nonmicrocaps	0.606***	0.615**	0.566***	0.636**	0.478***	0.647**	0.759***	0.463**
	(3.06)	(2.15)	(3.28)	(2.59)	(2.98)	(2.25)	(3.34)	(2.23)
Credit Rating Sample	0.647***	0.512*	0.533***	0.653**	0.498**	0.649**	0.782***	0.447*
	(2.82)	(1.66)	(2.97)	(2.51)	(2.56)	(2.14)	(3.18)	(1.96)
Nondowngrades	0.535**	0.326	0.378**	0.451	0.478**	0.325	0.583**	0.292
	(2.37)	(1.10)	(2.04)	(1.60)	(2.23)	(0.96)	(2.22)	(1.26)
Panel D: Sorted by CA2-Predicted Returns								
Full Sample	0.748***	0.617	0.634**	0.606	0.453*	0.910**	0.359	0.787**
	(2.70)	(1.41)	(2.34)	(1.61)	(1.75)	(2.04)	(1.51)	(2.14)
Nonmicrocaps	0.584**	0.064	0.537**	0.254	0.524**	0.122	0.292	0.275
	(2.55)	(0.21)	(2.13)	(0.79)	(2.28)	(0.33)	(1.13)	(1.03)
Credit Rating Sample	0.304	-0.055	0.531	-0.074	0.508*	-0.300	-0.102	0.263
	(1.15)	(-0.15)	(1.49)	(-0.19)	(1.69)	(-0.67)	(-0.30)	(0.81)
Nondowngrades	0.204	-0.218	0.376	-0.170	0.193	-0.248	-0.331	0.245
	(0.73)	(-0.54)	(1.30)	(-0.45)	(0.73)	(-0.54)	(-0.89)	(0.84)

Value-Weighted FF6-Adjusted Returns of Machine Learning Portfolios

Empirical Results - Time-Varying Return Predictability

$$HML_t = \alpha_0 + \beta_1 High\,SENT_{t-1} + \beta_2 High\,MKTVOL_{t-1} + \beta_3 High\,MKTILLIQ_{t-1}$$
$$+ \beta_4 M_{t-1} + c_1' F_t + c_2' F_t \times High\,SENT_{t-1} + c_3' F_t \times High\,MKTVOL_{t-1}$$
$$+ c_4' F_t \times High\,MKTILLIQ_{t-1} + e_t$$

	Sorted by NN3-Predicted Returns					
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Constant	0.016	-0.453	0.865	1.252	1.491**	1.666*
	(0.03)	(-0.92)	(1.14)	(1.42)	(2.01)	(1.93)
High SENT	1.534**	1.710**	0.228	-0.005	-0.179	-0.257
	(2.43)	(2.51)	(0.58)	(-0.01)	(-0.37)	(-0.45)
High MKTVOL	0.791		0.959*		0.660	
	(1.24)		(1.93)		(1.23)	
High VIX		1.851***		1.647***		1.419**
		(2.85)		(3.28)		(2.57)
High MKTILLIQ	0.754	0.529	0.592	0.695	0.671	0.392
	(1.24)	(0.78)	(1.37)	(1.41)	(1.43)	(0.77)
DOWN			-0.691*	-0.803*	-0.311	-0.297
			(-1.71)	(-1.74)	(-0.77)	(-0.64)
TERM			-0.146	-0.194	-0.161	-0.154
			(-0.80)	(-0.95)	(-0.97)	(-0.81)
DEF			-0.245	-0.732	-0.795	-1.221*
			(-0.33)	(-0.94)	(-1.27)	(-1.78)

Empirical Results - Return Predictability in Recent Years

- Whether ML techniques have remained meaningful in recent years?

Panel A: Value-Weighted Returns to Investment Strategies Sorted by NN3-Predicted Returns							
	Return	CAPM	FFC	FFC+PS	FF5	FF6	SY
Full Sample	1.568***	2.012***	1.559***	1.555***	1.096***	1.181***	0.699**
	(3.07)	(4.63)	(5.01)	(5.20)	(3.95)	(4.32)	(2.27)
Nonmicrocaps	1.106**	1.583***	1.128***	1.116***	0.592**	0.687***	0.276
	(2.27)	(3.97)	(4.23)	(4.28)	(2.19)	(2.88)	(0.93)
Credit Rating Sample	1.182**	1.635***	1.167***	1.165***	0.700**	0.803***	0.288
	(2.37)	(4.22)	(4.14)	(4.04)	(2.36)	(3.14)	(0.87)
Nondowngrades	0.796*	1.188***	0.747***	0.768***	0.454	0.545*	-0.037
	(1.82)	(3.15)	(2.65)	(2.67)	(1.58)	(1.90)	(-0.11)

- Unlike individual anomalies, ML signals continue to predict cross-sectional stock returns in recent years for the full sample. On the other hand, anomalous return patterns are still confined within difficult-to-arbitrage stocks, and thus, practitioners should remain cautious in utilizing machine learning algorithms for real-time trading.

Empirical Results - Economic Foundations of ML

- Whether stocks with similar machine learning signals also share other characteristics that predict future returns?

➢ Compute the EW average of a comprehensive set of stock characteristics at the end of month for each portfolio.

- Despite their opaque nature, ML techniques successfully identify mispriced stocks with solid economic foundations

Stock Characteristics	Sorted by NN3-Predicted Returns			
	Low	High	HML	*t*-stat
Log (Price)	2.092	1.386	-0.706***	(-14.63)
Log (Size)	5.483	3.613	-1.870***	(-25.33)
Book-to-Market	0.780	1.391	0.612***	(10.62)
Log (Illiquidity)	1.220	3.791	2.571***	(22.70)
Beta	1.330	1.068	-0.263***	(-6.87)
1M Return	0.027	-0.012	-0.039***	(-10.22)
12M Momentum	-0.124	0.185	0.309***	(15.64)
IdioVol	0.076	0.084	0.008***	(4.19)
Absolute Accruals	0.103	0.107	0.003	(1.08)
Log (Age)	2.171	2.463	0.292***	(9.93)
Assets Growth	0.586	0.014	-0.572***	(-11.76)
ΔShares Outstanding	0.389	0.070	-0.319***	(-10.92)
Corporate Investment	-0.107	0.022	0.129***	(9.03)
Dividend-to-Price	0.012	0.009	-0.003***	(-3.61)
Gross Profitability	0.329	0.354	0.025	(1.57)
Leverage	1.308	1.832	0.524***	(4.18)
ROA	-0.024	-0.011	0.013***	(5.41)
ROE	-0.045	-0.017	0.028***	(5.87)
%Rated	0.236	0.093	-0.142***	(-13.48)
Credit Rating	11.350	12.047	0.697**	(2.44)
Analyst Coverage	4.142	1.432	-2.710***	(-14.99)
Analyst Dispersion	0.049	0.057	0.008	(0.68)
SUE	-0.019	-0.009	0.010***	(3.17)

Conclusion

- ML techniques face the usual challenge of cross-sectional return predictability. In particular, the anomalous return patterns concentrate in difficult-to-value and difficult-to-arbitrage stocks.
- The trading strategy is more profitable during periods of high market volatility and low market liquidity.
- ML signals also involve remarkably high turnover.
- Beyond economic restrictions, ML-based trading strategies nonetheless display smaller downside risk, yield considerable profit in the long positions, and remain viable in the post-2001 period and the crisis period.
- Black-box-like machine learning methods generate economically interpretable trading strategies.