Machine Learning vs. Dynamic Trading: Can Economics Help Explain AI?

Huahao Lu (PBCSF, Tsinghua)
Matthew Spiegel (Yale)
Hong Zhang (SMU)

Motivation:

the Triumph of Machine Learning

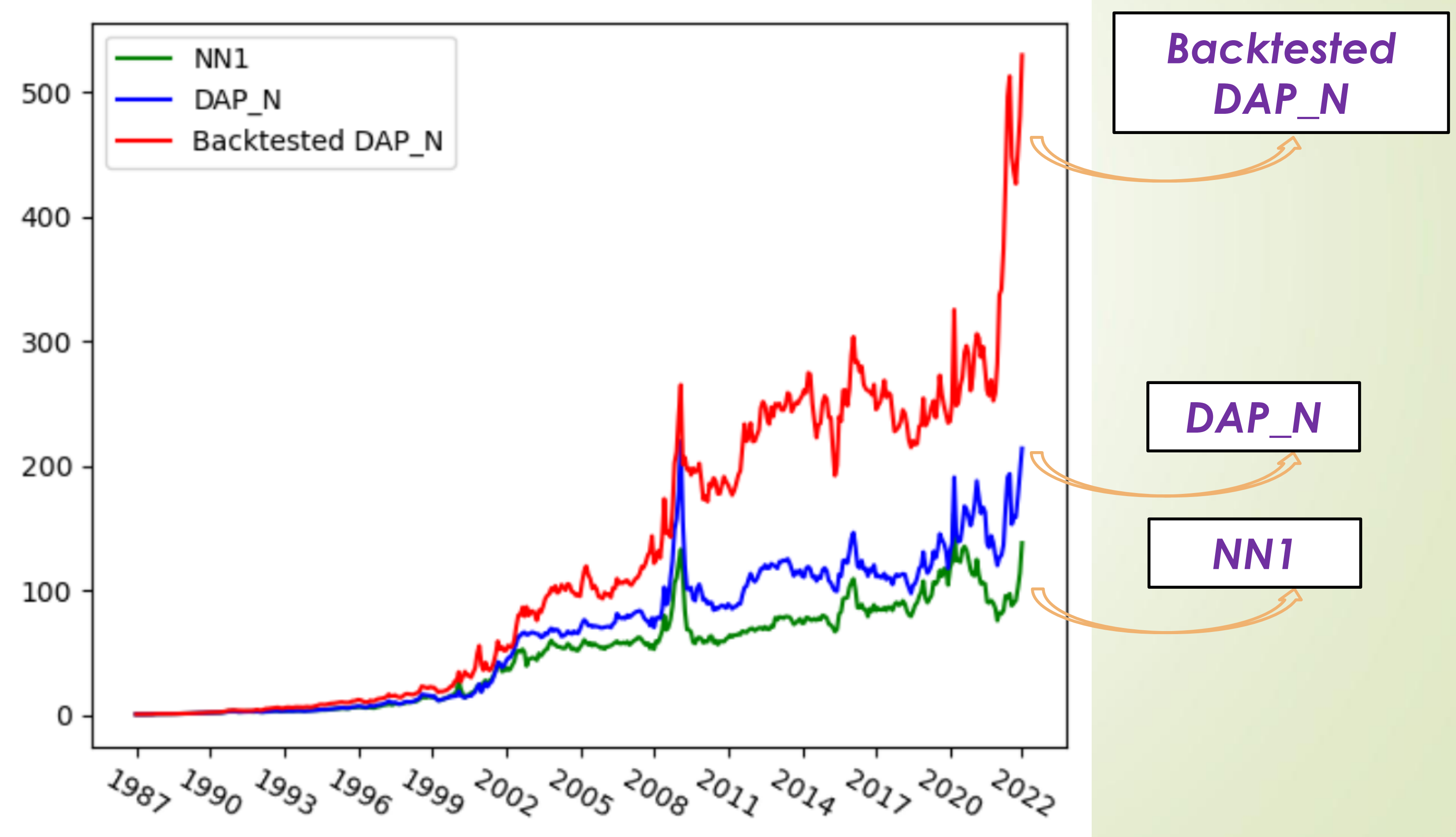
- Machine learning algorithms have performed remarkably well in financial markets.
  - Predicting stock returns: Freyberger, Neuhierl, and Weber (2020); Gu, Kelly, and Xiu (2020); Bryzgalova, Pelger, and Zhu (2020); Chen, Pelger, and Zhu (2023); Jensen, et al. (2022, trading-cost-adjusted portfolio optimization)
  - International: Leippold, Wang, and Zhou (2022, Chinese stock market); Li et al (2023, the global supply chain)
  - Other markets: Bianchi et al., (2021) for bond risk premium, Easley et al., (2021) in market microstructure, Filippou et al. (2022) for currencies, and Bali, et al., (2023) for option pricing. Van Binsbergen, Han, and Lopez-Lira, A. (2023, the conditional biases in earnings expectations)
  - Mutual/Hedge Fund Performance: Li and Rossi (2020), DeMiguel, et al., (2023), Kaniel et al., (2023). Wu et al., (2021, HF)
  - Surveys: Karolyi & Van Nieuwerburgh (2020); Kelly & Xiu (2023)

The Challenge:
How to explain the black box of ML (and its triumph)?

- Computer science research moves toward **interpretable AI** (e.g., Li et al. 2022 for a survey)
- The benefits of the highly parameterized nature of machine learning, such as "**complexity**" (e.g., Kelly, Malamud, Zhou 2024 JF; Didisheim, Ke, Kelly, and Malamud, 2024 WP).
- Our paper: are there known economic principles that can help explain the returns of ML models?
  - We employ an **economically motivated** and easy implementable trading strategy (DAPs) to interpret the performance of machine learning.
  - DAPs embody two critical aspects of machine learning:
    1) **nonlinearity**; and
    2) a **ReLU** activation function.

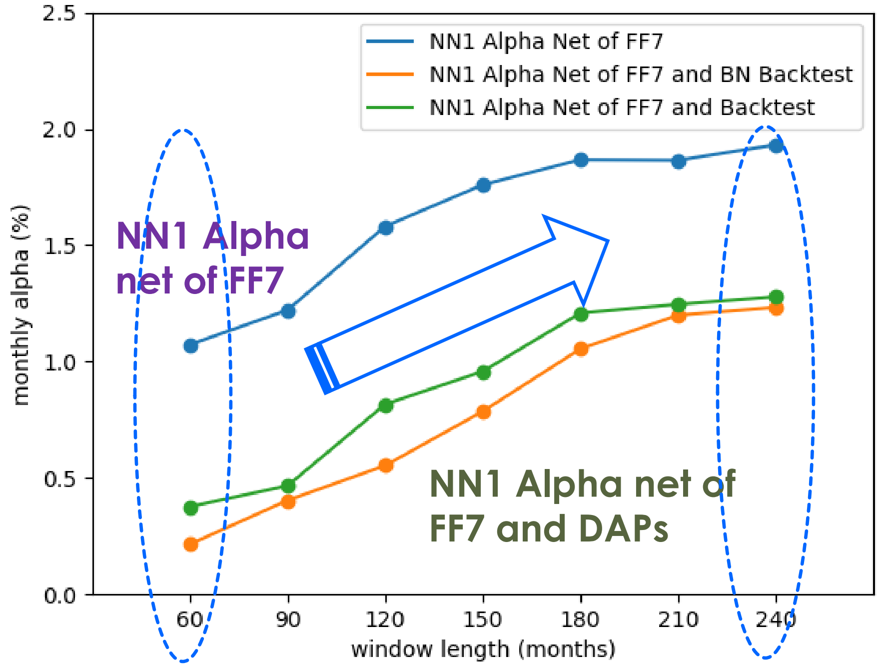
Main Findings (1)

- DAPs can help explain neural network-generated returns



Main Findings (2)

Main Predicting Period: 1987 – 2022



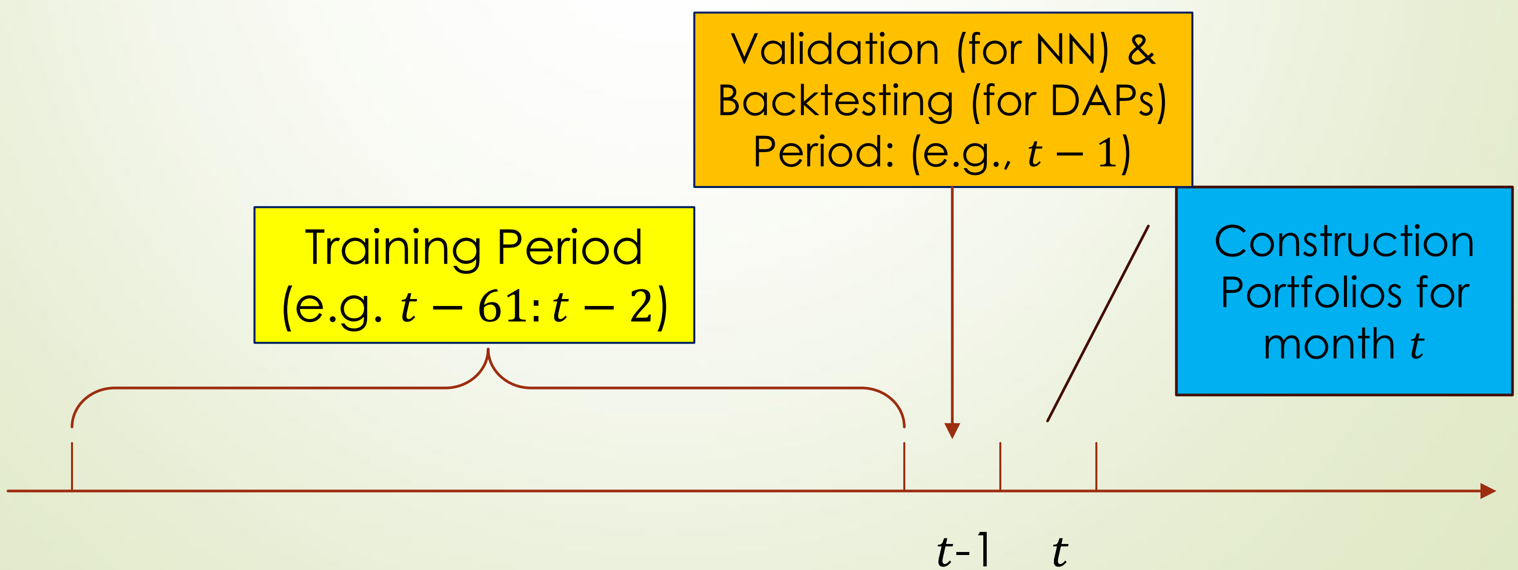
- DAPs consistently explain about 90 bps of neural network alphas.
  - DAPs absorb NN alphas with short training period (e.g., 60m).
  - With longer training periods, NNs generate better performance.
  - DAP can consistently explain a leading portion of NN alphas.
- Even at very long training period (240m), DAP can still absorb NN-generated alphas when:
  - excluding 20% microcap stocks;
  - NN and DAPs can only trade published anomalies.
- Two novel sources of ML returns:
  - A nonlinear strategy akin to DAPs
  - An aptitude for generating returns related to unpublished anomalies.

Roadmap

- Data sample
- The estimation frameworks: Neural Network and DAPs
- Empirical analyses
  - The Performance of DAPs and NNs
  - Using DAPs to explain NN returns
  - The explanatory of DAPs across training horizons
  - The post-publication test

1. Data and Portfolio Construction

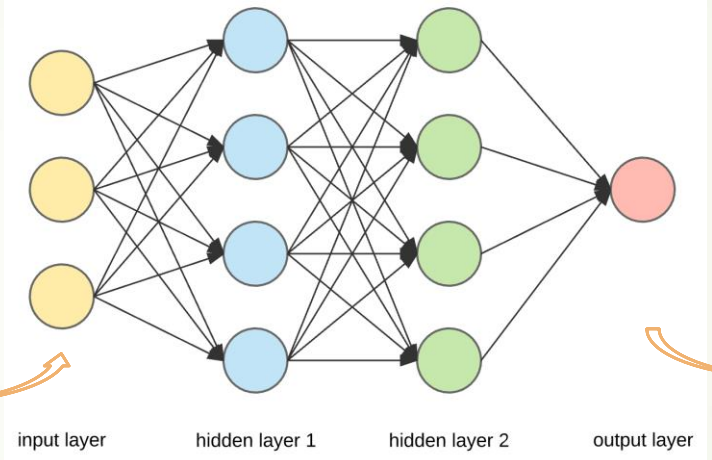
- We apply NN/DAP to US stocks with **153 anomalies** (Jensen, Kelly, and Pedersen, 2023 JF).
- In each month, we use NN/DAP predictions to construct 10 portfolios. Models are trained in a rolling window, with validation (NN)/backtesting (DAPs).
- We then analyze the out-of-sample NN/DAP portfolio monthly returns over **1987-2022**, the same prediction period as Gu, Kelly, and Xiu (2020).



2.1 Neural Network (NN)

- **Feedforward Neural Network**:
  - A multi-layer perceptron network consists of an input layer, an output layer, and one or more hidden layers.

Initial input: 153 characteristics of each stock



Output: return (ranks) of each stock

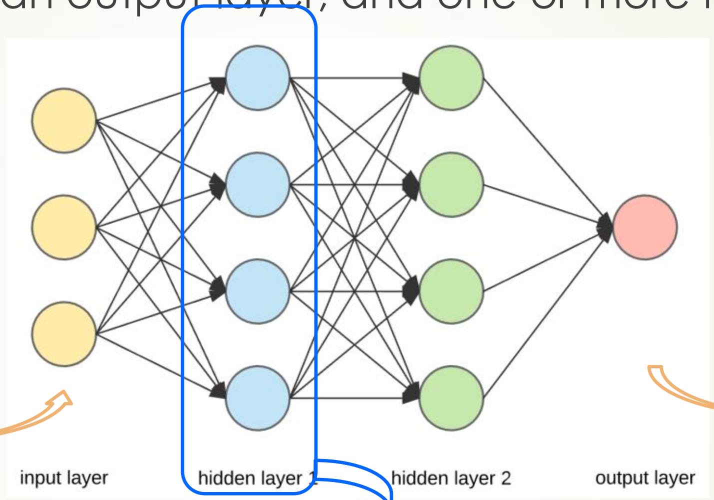
NN-based trading strategy:
1. We sort stocks into 10 deciles based on NN **predicted values** and calculate their value-weighted returns.
2. We then **long/short** the portfolios of high/low returns.

- Importantly: NN prediction functions are nonlinear (next PPT).

Neural Network (NN): continued

- **Feedforward Neural Network**:
  - A multi-layer perceptron network consists of an input layer, an output layer, and one or more hidden layers.

Initial input: 153 characteristics of each stock



Output: return (ranks) of each stock

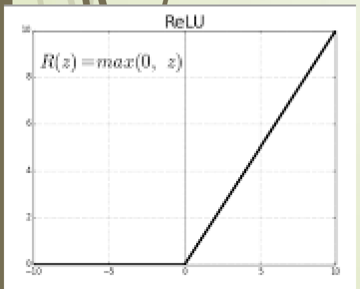
For the l-th layer: $\boxed{X^l} = g(W^{(l)T}\boxed{X^{(l-1)}} + b^{(l)})$

Output of the layer

input to the layer

$g(.)$: the non-linear activation function (ReLU)

$W^{(l)}$ and $b^{(l)}$: learnable parameters (weight + biases)



$$Relu(x) = \begin{cases} x, if\ x \geq 0 \\ 0, otherwise \end{cases}$$

2.2. A Hypothetical Example of DAPs

- Consider the above-market returns of 3 anomalies in two equally likely states (following Markov process).

	State X	State Y
A	3%	3%
B	6%	0
C	0	6%

- **Static strategies:** A, B, C or the average of A, B and C (3% return)
- **Dynamic Anomaly Portfolios (DAPs)**
  - **State-contingent**: A and B (but not C) in **State X**
  - **Signal weighted** (Kyle 1985; Admati 1985): investment weight in proportion to the value of the signal.

	Weights	Weights
A	1/3	1/3
B	2/3	0
C	0	2/3

5% ret (if perfect conditional information)

The Conceptual Framework of DAPs (1)

- The example fits the general principles of optimal trading rules (e.g., Kyle 1985; Admati 1985)
  - I.e., assuming that the k-th anomaly is associated with a dynamic risk premium in the conditional CAPM:

$$y_{kt} = \boxed{\underbrace{\boldsymbol{\alpha_{kt}}}_{\substack{time-varying \\ risk\ premium\ of \\ k^{th}\ characteristic}}} + \underbrace{\beta_{kt}}_{\substack{exposure\ to \\ the\ market}} \times \underbrace{X_t}_{\substack{market \\ return}} + \underbrace{\epsilon_{kt}}_{noise},$$

  - The **optimal investment weight** of anomalies is in proportion to the value of signals they offer (i.e., **conditional risk premium**):

$$w_{kt} \propto E[\boldsymbol{\alpha_{kt}} | t-1]$$

- **DAP returns:** $y_{Pt} = \sum_k c_t \times \boxed{\underbrace{\boldsymbol{\alpha}_{kt-1}}_{\substack{Conditional \\ Risk\ Premium \\ as\ inv-weight}}} \times \boxed{\underbrace{y_{kt}}_{\substack{Anomaly \\ Return \\ after\ investment}}}$

Nonlinearity 1: DAP returns ~ λ_{kt-1}^2

The Conceptual Framework of DAPs (2)

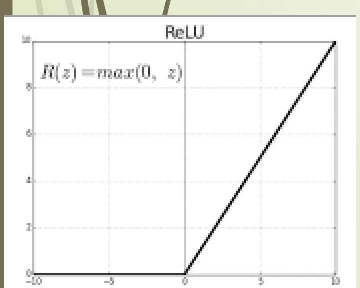
- The example also implies the source of estimation errors: **regime shifts** (Smith and Timmermann 2021a, b) may dynamically invalidate anomalies.
- Solution: back-testing (Mamaysky, Spiegel, and Zhang 2007, 2008), which is among the most robust predictors in the MF literature (e.g., Jones and Mo, 2021)
- Let $\hat{\boldsymbol{\alpha}}_{\mathbf{n}}$ = rolling-window risk premium (t-60:t-1); ($\hat{\boldsymbol{\alpha}}_{\mathbf{n}} > \mathbf{0}$)

  $\hat{\boldsymbol{\alpha}}_{\mathbf{n},-\mathbf{t}}$ = its most recent realization (in month $t-1$).

  Then backtest consider $\hat{\alpha}_n$ as a valid signal if $\hat{\alpha}_{n,-t} > 0$.
- In other words, back-testing introduces a ***ReLU*** activation function into the DAP framework

$$\boldsymbol{w_n = ReLU\left(\hat{\alpha}_n \times H(\hat{\alpha}_{n,-t})\right)},$$

where $H(\hat{\alpha}_{n,-t})$ is the sign function.



Construction of DAPs (1):
Mapping into a Stock-based DAP

- We first rewrite DAP returns at the stock-level, which allows us to derive the investment weight of a stock based on its exposure to anomalies:

$$y_{Pt} = \sum_k c_t \times \alpha_{kt-1} \times \underbrace{y_{kt}}_{\substack{Anomaly\\Return}} = \sum_k c_t \alpha_{kt-1} \underbrace{\left(\sum_i w_{i\in k,t} \times r_{it}\right)}_{\substack{Anomaly\ portfolio\\of\ stocks}}$$

$$= \sum_i \left\{ c_t \left(\sum_k \alpha_{kt-1} w_{i\in k,t} \right) r_{it} \right\},$$

- r_{it} is the return of stock i;
- $w_{i\in k,t}$ is the weight—determined by the **characteristic rank**—of the stock in the k-th anomaly.

The investment weight of the stock in DAP, which is the anomaly **alpha-weighted characteristic ranks** of stocks.

Construction of DAPs (2):
Mapping into a Stock-based DAP

- We construct four sets of DAPs based on the signals we use and the employment of the backtest
  - ***Base DAP***: conditional alphas (the H-L returns of anomalies) as signals; **no** backtesting;
  - ***Backtested DAP***: conditional alphas as signals; **with** backtesting;
  - ***DAP_N***: normalized conditional alphas (i.e., information ratio) as signals; **no** backtesting;
  - ***Backtested DAP_N***: using the conditional alpha as signals **with** backtesting;
- For each set:
  1. We sort stocks into 10 deciles based on their **DAP weights** and calculate the value-weighted returns.
  2. We then **long/short** the portfolios of high/low returns.
  3. Risk adjustment: momentum and short-term reversal enhanced **Fama-French seven-factor model**.

Roadmap

- Data sample
- The estimation frameworks: Neural Network and DAPs
- Empirical analyses
  - The Performance of DAPs and NNs
  - Using DAPs to explain NN returns
  - The explanatory of DAPs across training horizons
  - The post-publication test

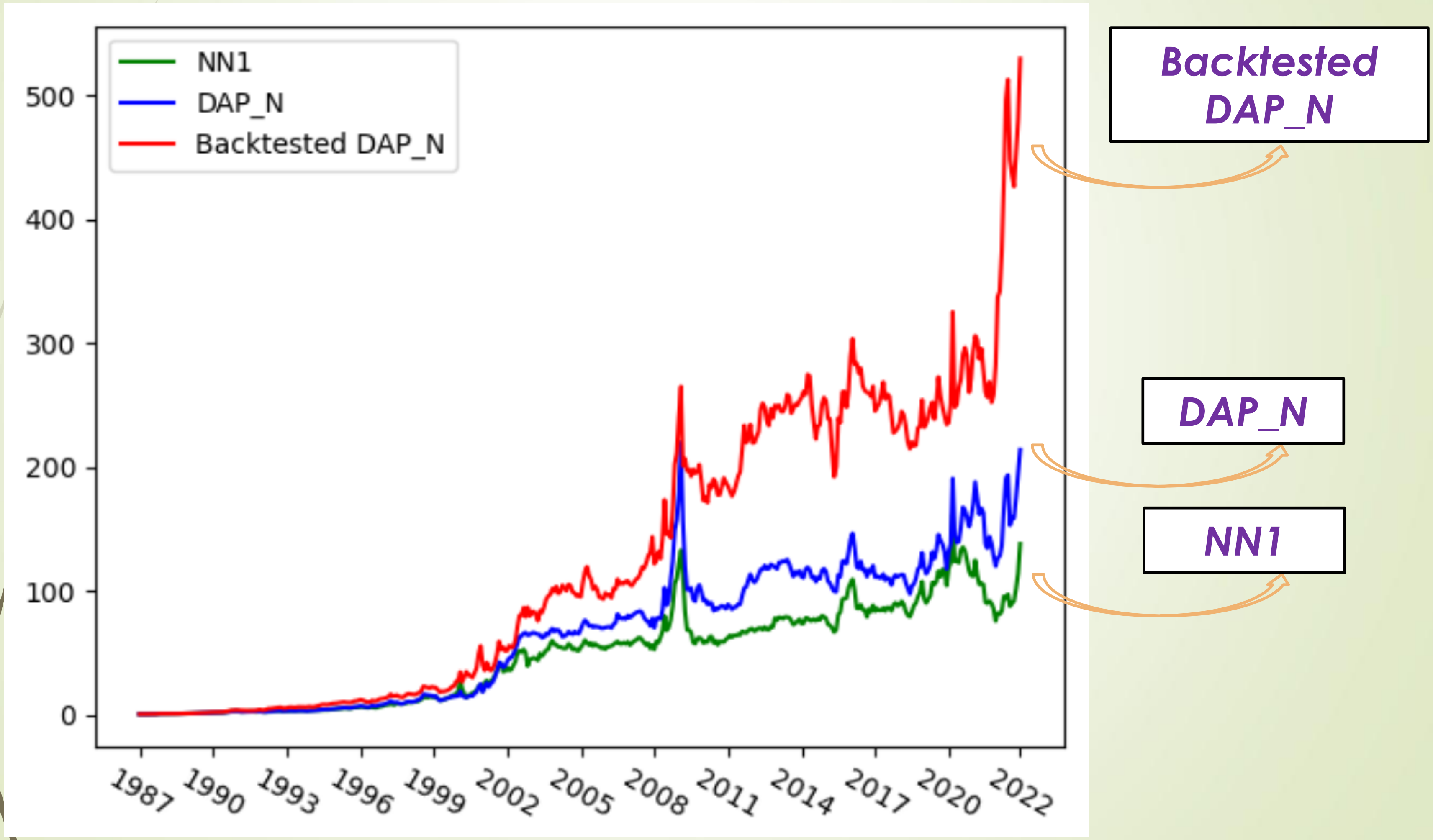
1. Out-of-sample monthly returns (60m training period; Excluding 20% microcaps)

	(1)	(2)	(3)	(4)	(5)	(6)
	Exclude Small Stocks (20%)					
	Low	High	H-L	FF3	FF5	FF7
Panel A. Returns Generated by Anomalies						
Avg Anomaly	-0.571	0.870***	1.441***	1.942***	1.271***	0.887***
	(-1.32)	(4.12)	(4.03)	(6.16)	(4.19)	(3.52)
Panel B. Returns Generated by DAP Portfolios						
Base DAP	-0.711	0.927***	1.638***	2.228***	1.573***	1.145***
	(-1.55)	(4.28)	(4.26)	(6.61)	(4.77)	(4.24)
Backtested DAP	-0.828*	1.060***	1.888***	2.384***	1.828***	1.823***
	(-1.87)	(4.30)	(4.87)	(6.73)	(5.27)	(7.12)
DAP_N	-0.510	0.976***	1.486***	1.994***	1.481***	1.150***
	(-1.22)	(4.61)	(4.48)	(6.83)	(5.24)	(4.84)
Backtested DAP_N	-0.605	1.084***	1.689***	2.144***	1.656***	1.658***
	(-1.50)	(4.82)	(5.17)	(7.31)	(5.82)	(7.71)
Panel C. Returns Generated by Neural Network Portfolios						
NN1	-0.368	1.029***	1.397***	1.846***	1.532***	1.072***
	(-0.91)	(4.03)	(4.08)	(5.72)	(4.69)	(3.95)
NN2	-0.376	0.946***	1.322***	1.799***	1.439***	0.964***
	(-0.90)	(3.86)	(3.78)	(5.52)	(4.35)	(3.47)
NN3	-0.330	0.900***	1.229***	1.632***	1.150***	0.678**
	(-0.80)	(3.63)	(3.49)	(4.87)	(3.41)	(2.38)

- Ave anomaly alpha: ~ 89 bps (or **11%** per year)
- The two backtested DAPs double the alphas: 1.82% and 1.66% (**24%** and **22%** when annualized)
- NN1 outperform other neural networks in this estimation. But NN performance is not impressive due to the short training window.

Time Series Plots:

Cumulative returns of H-L portfolios of DAPs and NN1 (60m training period; Excluding 20% microcaps)



2.1 **NN1 Returns** Explained by DAPs (60mon training)

	(1)	(2)	(3)	(4)	(5)
	Exclude Small Stocks (20%)				
Alpha	1.072***	0.223	0.261	0.081	0.136
	(3.95)	(1.19)	(1.00)	(0.44)	(0.52)
SMB	-0.107	0.054	0.011	0.169**	0.044
	(-1.09)	(0.80)	(0.12)	(2.58)	(0.50)
HML	-0.198*	-0.040	-0.085	0.045	-0.012
	(-1.68)	(-0.50)	(-0.78)	(0.57)	(-0.11)
RMW	-0.134	-0.424***	-0.202*	-0.214***	-0.206*
	(-1.08)	(-4.97)	(-1.78)	(-2.61)	(-1.84)
CMA	0.652***	-0.101	0.168	-0.282**	0.074
	(3.77)	(-0.83)	(1.02)	(-2.35)	(0.45)
MKT	-0.266***	0.002	-0.219***	-0.011	-0.189***
	(-4.04)	(0.05)	(-3.64)	(-0.25)	(-3.19)
MOM	0.859***	0.231***	0.720***	0.282***	0.725***
	(13.91)	(4.60)	(12.41)	(5.96)	(12.76)
SREV	-0.054	0.134**	0.544***	0.177***	0.552***
	(-0.65)	(2.39)	(5.56)	(3.22)	(5.86)
Base DAP		0.742***			
		(22.52)			
Backtested DAP			0.445***		
			(9.50)		
DAP_N				0.862***	
				(23.66)	
BackTested DAP_N					0.565***
					(10.29)
Adj R2	0.446	0.748	0.543	0.761	0.556
Observations	432	432	432	432	432

2.2 **Backtested DAP** returns explained by NNs (60mon training)

	(1)	(2)	(3)	(4)	(5)
	Exclude Small Stocks				
Alpha	1.823***	1.399***	1.496***	1.593***	1.406***
	(7.12)	(5.91)	(6.19)	(6.67)	(6.00)
SMB	-0.265***	-0.223***	-0.179**	-0.214**	-0.213**
	(-2.86)	(-2.64)	(-2.06)	(-2.49)	(-2.53)
HML	-0.255**	-0.177*	-0.167	-0.200*	-0.173*
	(-2.29)	(-1.74)	(-1.60)	(-1.93)	(-1.72)
RMW	0.153	0.206*	0.155	0.085	0.158
	(1.30)	(1.92)	(1.41)	(0.78)	(1.47)
CMA	1.087***	0.829***	0.868***	0.831***	0.775***
	(6.66)	(5.50)	(5.63)	(5.38)	(5.16)
MKT	-0.106*	-0.001	-0.006	-0.041	-0.000
	(-1.71)	(-0.02)	(-0.11)	(-0.71)	(-0.00)
MOM	0.312***	-0.027	0.019	0.019	-0.076
	(5.36)	(-0.43)	(0.29)	(0.29)	(-1.17)
SREV	-1.342***	-1.321***	-1.356***	-1.344***	-1.326***
	(-17.30)	(-18.73)	(-18.78)	(-18.70)	(-18.94)
NN1		0.395***			0.299***
		(9.50)			(4.25)
NN2			0.339***		-0.025
			(8.12)		(-0.33)
NN3				0.339***	0.176***
				(8.37)	(3.16)
Adj R2	0.616	0.683	0.667	0.670	0.690
Observations	432	432	432	432	432

3. The Explanatory Power of DAPs across Various Training Lengths

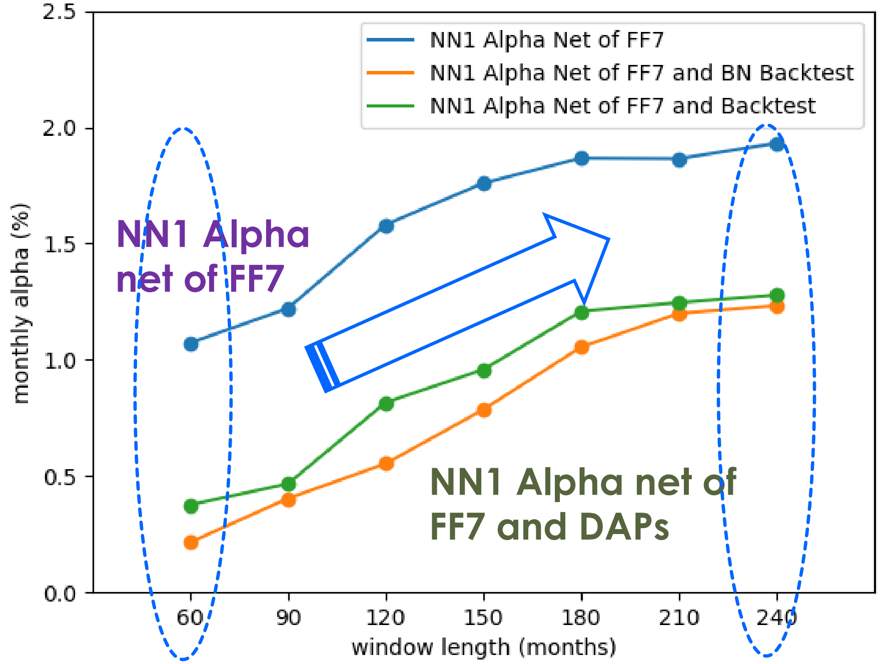
- We observe that DAP can absorb NN-generated alphas when the training period is short (60 months).
- But neural networks are known to perform better with larger data sets.
- Hence, we expand our previous tests to rolling training windows with different lengths (i.e., 60, 90, 120, …, 240 months).
  - We apply both DAPs and NNs to a similar length of training windows
  - Below we first illustrate an 240m example. We then document the general results.

3.1 NN1 Returns Explained by DAPs (240m training)

	(1)	(2)	(3)	(4)	(5)
	Exclude Small Stocks (20%)				
Alpha	1.930***	1.401***	1.253***	1.309***	1.164***
	(8.45)	(6.48)	(5.53)	(6.28)	(5.14)
SMB	-0.001	0.303***	0.091	0.252***	0.074
	(-0.02)	(3.68)	(1.17)	(3.30)	(0.97)
HML	0.315***	0.138	0.342***	0.152*	0.314***
	(3.17)	(1.48)	(3.71)	(1.71)	(3.44)
RMW	0.270**	-0.446***	0.064	-0.364***	0.017
	(2.58)	(-3.62)	(0.64)	(-3.36)	(0.17)
CMA	0.523***	0.071	0.086	-0.077	0.058
	(3.59)	(0.50)	(0.60)	(-0.56)	(0.40)
MKT	-0.094*	0.077	-0.044	0.089*	-0.052
	(-1.70)	(1.42)	(-0.85)	(1.73)	(-1.01)
MOM	0.630***	0.325***	0.543***	0.293***	0.548***
	(12.15)	(5.63)	(11.03)	(5.34)	(11.29)
SREV	0.098	0.143**	0.500***	0.125**	0.541***
	(1.42)	(2.25)	(6.24)	(2.05)	(6.72)
Base DAP		0.549***			
		(9.24)			
Backtested DAP			0.365***		
			(8.35)		
DAP_N				0.681***	
				(11.17)	
BackTested DAP_N					0.444***
					(8.95)
Adj R2	0.419	0.522	0.458	0.496	0.466
Observations	432	432	432	432	432

3.2 The Explanatory Power of DAPs across Various Training Lengths

Main Predicting Period: 1987 – 2022



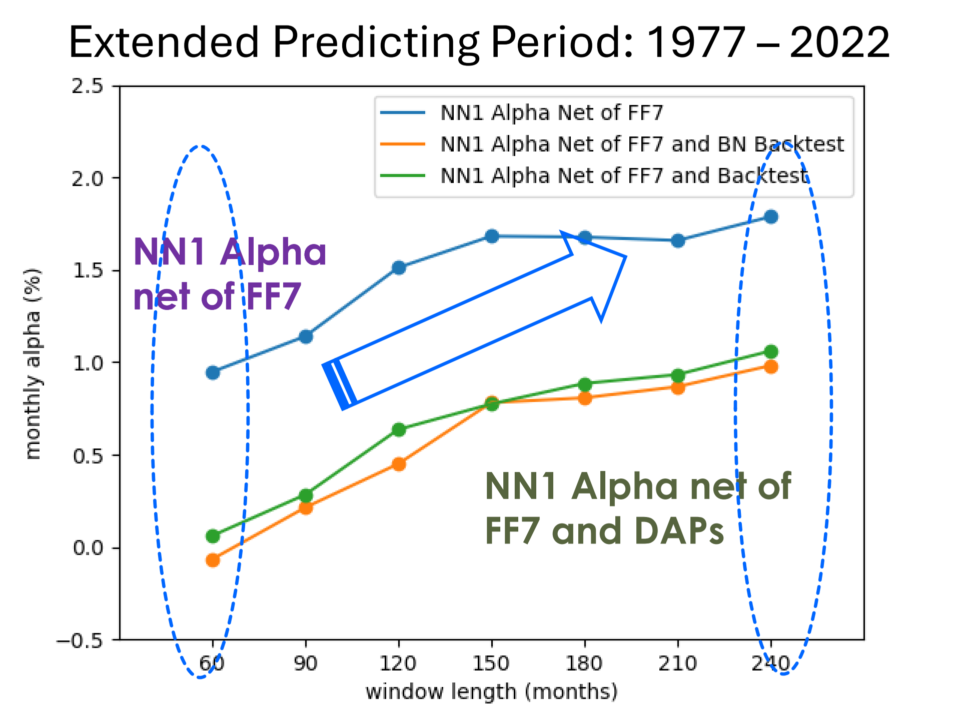
- DAP can absorb NN-generated alphas when the training period is short (60 months).
- With longer training periods, NNs generate better performance.
- DAP can consistently explain a leading portion of NN alphas (monthly 0.7~0.8% risk-adjusted).

3.3. The Explanatory Power of DAPs across Various Training Lengths (2)

Monthly NN1 Alphas (in %)	Estimation Window (months)						
	60	90	120	150	180	210	240
NN2 Alphas Net of FF7	1.07	1.22	1.58	1.76	1.87	1.86	1.93
	(3.95)	(4.71)	(6.20)	(7.04)	(8.05)	(8.12)	(8.45)
Net of FF7 and Backtested DAP	0.23	0.37	0.7	0.92	1.06	1.14	1.24
	(0.90)	(1.48)	(2.87)	(3.85)	(4.77)	(5.06)	(5.48)
Net of FF7 and Backtested DAP_N	0.14	0.29	0.49	0.86	1.05	1.11	1.15
	(0.54)	(1.16)	(2.06)	(3.68)	(4.68)	(4.97)	(5.08)
Explained by Backtested DAP	0.84	0.85	0.88	0.84	0.81	0.72	0.69
Explained by Backtested DAP_N	0.93	0.93	1.09	0.9	0.82	0.75	0.78

The average explained NN1 alphas are **0.804%** and **0.886%**, respectively, by backtested DAPs and BN backtested DAPs.

3.4 The Explanatory Power of DAPs in extended testing period: 1977-2022



- We observe similar patterns of DAPs in explaining NN1 alphas in the extended testing period.
- The average explained NN1 alphas are **0.827%** and **0.911%**, respectively, by backtested DAPs and BN backtested DAPs.

Monthly NN1 Alphas (in %)	Estimation Window (months)						
	60	90	120	150	180	210	240
NN2 Alphas Net of FF7	0.95	1.14	1.51	1.68	1.68	1.66	1.79
	(4.11)	(5.30)	(7.18)	(8.21)	(8.78)	(8.70)	(9.52)
Net of FF7 and Backtested DAP	0.06	0.28	0.64	0.77	0.88	0.93	1.06
	(0.28)	(1.33)	(3.05)	(3.92)	(4.75)	(4.89)	(5.65)
Net of FF7 and Backtested DAP_N	-0.07	0.21	0.45	0.78	0.81	0.87	0.98
	(-0.31)	(1.01)	(2.25)	(3.96)	(4.30)	(4.58)	(5.25)
Explained by Backtested DAP	0.89	0.86	0.87	0.91	0.8	0.73	0.73
Explained by Backtested DAP_N	1.02	0.93	1.06	0.9	0.87	0.79	0.81

4. The post-publication test

- To explain the NN returns at very long training period (240m), we further conduct two post-publication tests:

- Test 1: we augment the FF7 factors by a **Pre-publication Ratio,** the fraction of anomalies before publication at any given time.
- Test 2: we allow NN and DAPs to only trade on **published anomalies**.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Excluding 20% Microcap Stocks				All Stocks			
VARIABLES	nn1	nn1	nn1	nn1	nn1	nn1	nn1	nn1
Constant	1.930***	1.164***	0.786**	0.296	2.499***	2.037***	1.332***	1.126***
	(8.45)	(5.14)	(2.07)	(0.83)	(9.37)	(7.75)	(2.99)	(2.64)
pre_pub			2.283***	1.786***			2.329***	1.856***
			(3.73)	(3.15)			(3.25)	(2.69)
backtest DAP_N		0.444***		0.428***		0.414***		0.397***
		(8.95)		(8.68)		(6.78)		(6.51)
smb	-0.001	0.074	0.000	0.073	0.086	0.019	0.088	0.023
	(-0.02)	(0.97)	(0.00)	(0.96)	(0.89)	(0.20)	(0.92)	(0.25)
hml	0.315***	0.314***	0.294***	0.298***	0.136	-0.020	0.114	-0.030
	(3.17)	(3.44)	(3.00)	(3.29)	(1.17)	(-0.18)	(0.99)	(-0.27)
rmw	0.270**	0.017	0.278***	0.032	0.038	-0.154	0.046	-0.140
	(2.58)	(0.17)	(2.69)	(0.32)	(0.31)	(-1.29)	(0.38)	(-1.18)
cma	0.523***	0.058	0.525***	0.077	0.862***	0.438**	0.865***	0.458***
	(3.59)	(0.40)	(3.67)	(0.54)	(5.07)	(2.52)	(5.14)	(2.66)
mkt_rf	-0.094*	-0.052	-0.094*	-0.053	-0.053	-0.053	-0.053	-0.053
	(-1.70)	(-1.01)	(-1.73)	(-1.05)	(-0.81)	(-0.87)	(-0.82)	(-0.87)
mom	0.630***	0.548***	0.612***	0.536***	0.211***	0.252***	0.192***	0.235***
	(12.15)	(11.29)	(11.91)	(11.13)	(3.49)	(4.35)	(3.19)	(4.06)
st_rev	0.098	0.541***	0.087	0.516***	0.332***	0.629***	0.321***	0.608***
	(1.42)	(6.72)	(1.27)	(6.45)	(4.11)	(7.11)	(4.01)	(6.89)

Test 1:

Adding the **Pre-publication Ratio** absorbs NN1 alpha when excluding 20% microcap stocks—but not so when including these stocks.

	(1)	(2)	(3)	(4)	(5)
			Exclude Small Stocks (20%)		
Alpha	0.514***	0.300*	0.323*	0.221	0.234
	(2.61)	(1.78)	(1.91)	(1.12)	(1.19)
SMB	-0.151**	0.006	-0.063	-0.148**	-0.173**
	(-2.12)	(0.10)	(-1.02)	(-2.15)	(-2.52)
HML	0.453***	0.125	0.167**	0.421***	0.389***
	(5.30)	(1.62)	(2.16)	(5.07)	(4.68)
RMW	0.686***	0.230***	0.339***	0.623***	0.605***
	(7.59)	(2.71)	(4.11)	(7.08)	(6.86)
CMA	0.495***	0.242**	0.178	0.305**	0.324***
	(3.95)	(2.24)	(1.61)	(2.42)	(2.60)
MKT	-0.185***	-0.074*	-0.124***	-0.167***	-0.180***
	(-3.88)	(-1.78)	(-3.02)	(-3.62)	(-3.91)
MOM	0.639***	0.224***	0.272***	0.585***	0.574***
	(14.29)	(4.48)	(5.61)	(13.22)	(12.89)
SREV	0.391***	0.384***	0.397***	0.605***	0.629***
	(6.56)	(7.58)	(7.77)	(8.72)	(8.90)
Base DAP		0.525***			
		(12.81)			
Backtested DAP			0.544***		
			(12.34)		
DAP_N				0.183***	
				(5.54)	
BackTested DAP_N					0.216***
					(5.77)
Adj R2	0.593	0.707	0.701	0.621	0.623
Observations	432	432	432	432	432

Test 2:

Even at very long training period (240m), DAP can still absorb NN-generated alphas when:

1. excluding 20% microcap stocks;
2. NN and DAPs can only trade published anomalies.

	Whole Sample				
Alpha	0.884***	0.487***	0.385**	0.781***	0.787***
	(4.49)	(2.79)	(2.30)	(3.95)	(3.98)
SMB	0.277***	-0.068	0.006	0.227***	0.236***
	(3.89)	(-0.99)	(0.10)	(3.14)	(3.28)
HML	0.547***	0.238***	0.219***	0.481***	0.502***
	(6.38)	(3.01)	(2.92)	(5.50)	(5.81)
RMW	0.197**	0.142*	0.116	0.184**	0.171*
	(2.18)	(1.80)	(1.54)	(2.05)	(1.90)
CMA	0.559***	-0.001	-0.162	0.437***	0.425***
	(4.45)	(-0.01)	(-1.39)	(3.36)	(3.21)
MKT	-0.163***	-0.130***	-0.120***	-0.179***	-0.182***
	(-3.42)	(-3.13)	(-3.01)	(-3.77)	(-3.81)
MOM	0.061	0.160***	0.174***	0.090**	0.083*
	(1.35)	(4.01)	(4.57)	(2.00)	(1.85)
SREV	0.560***	0.404***	0.418***	0.628***	0.627***
	(9.39)	(7.53)	(8.24)	(9.99)	(9.92)
Base DAP		0.481***			
		(11.76)			
Backtested DAP			0.656***		
			(13.81)		
DAP_N				0.108***	
				(3.14)	
BackTested DAP_N					0.114***
					(2.99)
Adj R2	0.432	0.572	0.608	0.444	0.443
Observations	432	432	432	432	432

- However, including microcap stocks still allow NN to deliver FF7 and DAP-adjusted returns

- This result suggests three sources of NN performance:

1. A nonlinear strategy akin to DAPs
2. An aptitude for generating returns related to unpublished anomalies.
3. A loading on (spurious) small stock returns.

Conclusions

- Our analysis reveals three main sources of neutral networks:
  - DAP-type of dynamic alphas.
  - The loading on unpublished anomalies.
  - The trading of small stocks (could be spurious).
- Ultimately, theory-based dynamic trading strategies (DAPs) can help explain a large portion of the returns generated by machine learning.

Thank you very much!