Replicating Gu, Kelly, Xiu (2020, RFS)

Seongdeok Ko

June 8, 2021

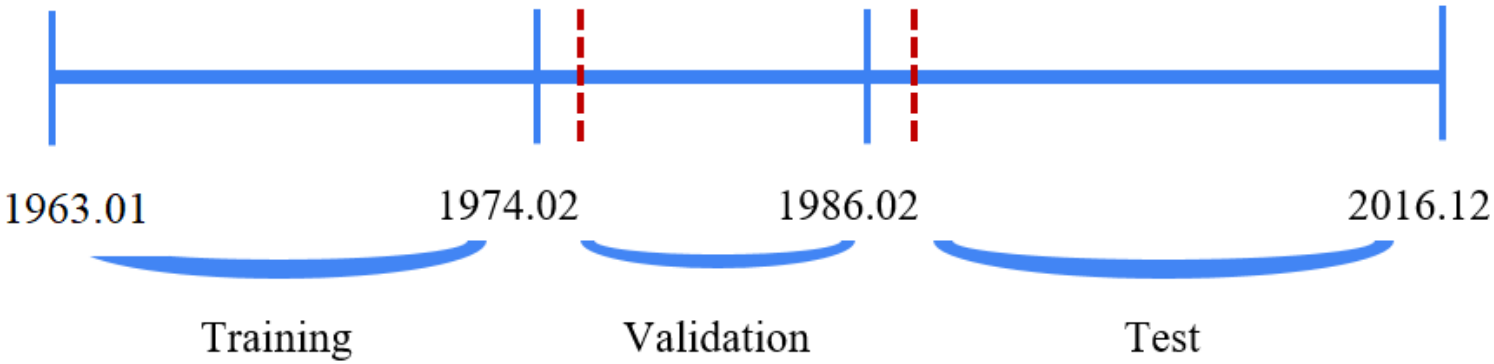
1. **Main Result of Target Paper**
2. **Replication Methodology and Data**
3. **Replication Results**

Main Results

- There do exist statistically significant Jensen's alphas for many portfolio based on ML predictions.
- The results shows that the six factors are not mean-variance efficient in the economy.
- By the existence of statistically significant abnormal return we can reject the joint hypothesis that six-factor model is correct and the market is efficient.

Replication Methodology and Data

- Estimate $\mathbb{E}_t[r_{i,t+1}]$ with various ML techniques including PCR, PLS, Elastic Net, GLM, Random Forest,Neural Networks
- Using expanding windows for estimation : re-estimate parameters every year
- Y variable : Individual excess returns $R_{i,t} - R_{f,t}$ from CRSP for all firms listed in the NYSE, AMEX and NASDAQ (1963 $\sim$ 2016)



Replication Methodology and Data

- Because of many missing variables in the early years I dropped sample from 1957 to 1962.
- X variables
  - 94 firm characteristics from Green et al.(2017)
  - 8 macroeconomic variables from Welch and Goyal (2008)
  - 74 industry indicator variables.
- In the original paper, total number of covariates is 920.
- Because of memory limit, I used only $94 + 74 = 168$ variables.

Main Results

Drawdowns, turnover, and risk-adjusted performance of machine learning portfolios

	OLS-3 +H	PLS	PCR	ENet +H	GLM +H	RF	GBRT +H	NN1	NN2	NN3	NN4	NN5
Drawdowns and turnover (value weighted)												
Max DD(%)	69.60	41.13	42.17	60.71	37.09	52.27	48.75	61.60	55.29	30.84	51.78	57.52
Max 1M loss(%)	24.72	27.40	18.38	27.40	15.61	26.21	21.83	18.59	37.02	30.84	33.03	38.95
Turnover(%)	58.20	110.87	125.86	151.59	145.26	133.87	143.53	121.02	122.46	123.50	126.81	125.37
Drawdowns and turnover (equally weighted)												
Max DD(%)	84.74	32.35	31.39	33.70	21.01	46.42	37.19	18.25	25.81	17.34	14.72	21.78
Max 1M loss(%)	37.94	32.35	22.33	32.35	15.74	34.63	22.34	12.79	25.81	12.50	9.01	21.78
Turnover(%)	57.24	104.47	118.07	142.78	137.97	120.29	134.24	112.35	112.43	113.76	114.17	114.34
Risk-adjusted performance (value weighted)												
Mean ret.	0.94	1.02	1.22	0.60	1.06	1.62	0.99	1.81	1.92	2.12	2.26	1.97
FF5+Mom α	0.39	0.24	0.62	−0.23	0.38	1.20	0.66	1.20	1.33	1.52	1.76	1.43
$t(\alpha)$	2.76	1.09	2.89	−0.89	1.68	3.95	3.11	4.68	4.74	4.92	6.00	4.71
R^2	78.60	34.95	39.11	28.04	30.78	13.43	20.68	27.67	25.81	20.84	20.47	18.23
IR	0.54	0.21	0.57	−0.17	0.33	0.77	0.61	0.92	0.93	0.96	1.18	0.92
Risk-adjusted performance (equally weighted)												
Mean ret.	1.34	2.08	2.45	2.11	2.31	2.38	2.14	2.91	3.31	3.27	3.33	3.09
FF5+Mom α	0.83	1.40	1.95	1.32	1.79	1.88	1.87	2.60	3.07	3.02	3.08	2.78
$t(\alpha)$	6.64	5.90	9.92	4.77	8.09	6.66	8.19	10.51	11.66	11.70	12.28	10.68
R^2	84.26	26.27	40.50	20.89	21.25	19.91	11.19	13.98	10.60	9.63	11.57	14.54
IR	1.30	1.15	1.94	0.93	1.58	1.30	1.60	2.06	2.28	2.29	2.40	2.09

Replication Results

Table: Risk-Adjusted Performance of ML portfolio - Value Weighted

	PLS	**PCR**	**ENET**	**GLM**	**RF**	**NN3**
Mean Return	0.35	0.52	0.56	0.61	0.49	-0.04
FF6 α	0.14	0.32	0.65	0.71	0.72	-0.077
t(α)	0.56	1.10	2.25	2.80	1.77	-0.64
R^2	5.02	7.04	3.43	1.92	3.13	2.45
SR	0.26	0.40	0.37	0.42	0.24	-0.08

Table: Risk-Adjusted Performance of ML portfolio - Equal Weighted

	PLS	**PCR**	**ENET**	**GLM**	**RF**	**NN3**
Mean Return	1.48	1.49	1.47	1.24	1.81	0.18
FF6 α	1.23	1.06	1.43	1.23	1.70	0.19
t(α)	5.31	3.55	4.62	4.26	4.08	2.17
R^2	4.19	12.59	8.84	6.37	4.16	1.83
SR	1.16	0.95	0.91	0.77	0.88	0.51

Replication Results

- For PLS, PCR and RF model abnormal return was about half of the target paper.
- Unlike target paper, both elastic net and GLM shows significant abnormal return.
- Using limited number of X variables Random Forest and Neural network model give us insignificant abnormal return.
- Unlike the target paper, neural network does not shows great performances.
- When Equal Weighted scheme is implemented, the performance get enhanced.

Appendix : Replicating Table 7

Table: Performance of ML Portfolio - (1)

	PLS			PCR			E-net		
	Avg	SD	SR	Avg	SD	SR	Avg	SD	SR
Low	0.53	6.70	0.27	0.54	4.62	0.40	0.46	4.96	0.32
2	0.46	5.69	0.28	0.74	4.73	0.54	0.62	4.41	0.49
3	0.68	4.97	0.47	0.85	4.65	0.63	0.66	4.11	0.56
4	0.64	4.81	0.46	0.74	4.59	0.56	0.67	4.33	0.53
5	0.71	4.87	0.50	0.76	4.64	0.56	0.73	4.66	0.54
6	0.79	4.75	0.57	0.82	4.88	0.58	0.70	4.69	0.52
7	0.73	4.54	0.55	0.86	5.07	0.58	0.55	5.30	0.36
8	0.78	4.57	0.59	0.89	5.40	0.57	0.80	5.86	0.47
9	0.83	4.59	0.62	0.90	5.74	0.54	1.14	6.70	0.57
10	0.88	4.71	0.65	1.06	6.74	0.55	1.03	8.07	0.44

Appendix : Replicating Table 7

Table: Performance of ML Portfolio - (2)

	GLM			RF			NN3		
	Avg	SD	SR	Avg	SD	SR	Avg	SD	SR
Low	0.41	5.43	0.26	0.47	6.27	0.26	0.66	5.01	0.46
2	0.57	4.82	0.41	0.62	4.79	0.45	0.80	4.91	0.56
3	0.63	4.29	0.51	0.65	4.67	0.48	0.67	4.77	0.49
4	0.66	4.36	0.52	0.79	4.34	0.63	0.66	4.58	0.50
5	0.52	4.68	0.39	0.78	4.31	0.62	0.65	4.56	0.49
6	0.66	4.42	0.52	0.73	4.35	0.58	0.57	4.43	0.45
7	0.57	4,73	0.41	0.68	4.77	0.49	0.67	4.33	0.53
8	0.66	5.03	0.45	0.61	5.69	0.37	0.70	4.55	0.53
9	0.81	6.12	0.46	0.51	7.10	0.24	0.72	4.53	0.55
10	1.02	7.56	0.47	0.97	9.36	0.36	0.61	4.50	0.47

Thank You