Online Appendix

Stock return predictability: A factor-augmented predictive regression system with shrinkage method ¹

Saburo Ohno

Amundi Japan Ltd.

saburo.ohno@jp.amundi.com

and

Tomohiro Ando

Graduate School of Business, Keio University

andoh@kbs.keio.ac.jp

¹ The authors would like to the editor and three anonymous referees for many excellent suggestions which have substantially improved the paper. We are also grateful to them for pointing out several references. We would also like to thank Ruey Tsay for helpful advice and constructive comments. The research of TA is partially supported by the Seimeikai Foundation, Japan

Table 1: Summary of the mean squared error (MSE), mean absolute error (MAE) and correct sign prediction measure (SIGN). We regress y_{t+h} on $\hat{y}_{t+h}$ using a simple linear regression: $y_{t+h} = \alpha + \beta\hat{y}_{t+h} + \varepsilon_{t+h}$. The table also includes the estimated β coefficients, t -value, p -value and adjusted R^2 scores. Prediction periods; Panel A: from December 30, 1999 to October 31, 2012. Panel B: from December 30, 1999 to September 12, 2008. Panel C: from September 16, 2008 to October 31, 2012. Returns are calculated as from next day's opening price (OP) to closing price (CP), i.e., (CP-OP)/OP.

Panel A	h	MSE	MAE	SIGN	$\hat{\beta}$	t -value	p -value	Adjusted R^2
Daily Models	1 day	0.0001	0.0083	0.5130	0.3258	9.4759	0.0000	0.0274
	2 days	0.0004	0.0139	0.5371	0.1619	4.2401	0.0000	0.0054
	3 days	0.0007	0.0192	0.5133	0.0841	2.9261	0.0035	0.0024
	5 days	0.0021	0.0319	0.5342	0.0615	4.3739	0.0000	0.0057
	10 days	0.0053	0.0511	0.5219	0.0555	4.9366	0.0000	0.0074
	20 days	0.0100	0.0694	0.5437	0.0726	6.5383	0.0000	0.0131
Monthly Models	1 month	0.0030	0.0438	0.4975	-0.1715	-0.8295	0.4078	-0.0015
	2 month	0.0115	0.0824	0.4975	-0.0569	-0.7089	0.4792	-0.0025
	3 month	0.0279	0.1230	0.5074	-0.0502	-0.8803	0.3797	-0.0011
	6 month	0.1053	0.2607	0.4236	-0.1288	-2.8942	0.0042	0.0352
	12 month	0.1711	0.3498	0.3218	-0.3154	-5.1045	0.0000	0.1108

Panel B	h	MSE	MAE	SIGN	$\hat{\beta}$	t -value	p -value	Adjusted R^2
Daily Models	1 day	0.0001	0.0081	0.5084	0.2883	5.0578	0.0000	0.0113
	2 days	0.0003	0.0136	0.5294	0.1788	3.2981	0.0010	0.0046
	3 days	0.0006	0.0186	0.5163	0.1485	4.1339	0.0000	0.0075
	5 days	0.0018	0.0300	0.5308	0.0673	3.7223	0.0002	0.0060
	10 days	0.0048	0.0489	0.5285	0.0416	3.0128	0.0026	0.0038
	20 days	0.0073	0.0615	0.5261	0.0588	3.6144	0.0003	0.0056
Monthly Models	1 month	0.0026	0.0420	0.4706	-0.4610	-1.5646	0.1192	0.0094
	2 month	0.0106	0.0805	0.5098	-0.0318	-0.3360	0.7372	-0.0059
	3 month	0.0217	0.1130	0.4967	-0.0817	-0.9757	0.3304	-0.0003
	6 month	0.1051	0.2688	0.3856	-0.1924	-3.6070	0.0004	0.0732
	12 month	0.1895	0.3694	0.2961	-0.3395	-5.0759	0.0000	0.1409

Panel C	h	MSE	MAE	SIGN	$\hat{\beta}$	t -value	p -value	Adjusted R^2
Daily Models	1 day	0.0002	0.0087	0.5227	0.3440	7.4428	0.0000	0.0511
	2 days	0.0005	0.0148	0.5534	0.1490	2.5962	0.0095	0.0056
	3 days	0.0009	0.0205	0.5069	-0.0010	-0.0197	0.9843	-0.0010
	5 days	0.0029	0.0361	0.5415	0.0537	2.3082	0.0211	0.0043
	10 days	0.0062	0.0557	0.5079	0.0742	3.7339	0.0002	0.0126
	20 days	0.0158	0.0862	0.5810	0.0824	5.1418	0.0000	0.0245
Monthly Models	1 month	0.0041	0.0493	0.5800	-0.0069	-0.0190	0.9849	-0.0208
	2 month	0.0143	0.0880	0.4600	-0.1200	-0.7365	0.4623	-0.0094
	3 month	0.0469	0.1538	0.5400	-0.0148	-0.1666	0.8679	-0.0202
	6 month	0.1059	0.2359	0.5400	0.0321	0.3839	0.7014	-0.0177
	12 month	0.1149	0.2905	0.4000	-0.2953	-1.9055	0.0582	0.0510

Table 2: List of the top-10 variables in Daily Model, conditional mean prediction ($h = 2$), in order of the number and average times selected and the longest period for which the variable was continuously selected. Prediction periods; Panel A: from December 30, 1999 to October 31, 2012 for Daily Models; from December, 1995 to October, 2012 for Monthly Models. Panel B: from December 30, 1999 to September 12, 2008, for Daily Models; from December, 1995 to August, 2008 for Monthly Models. Panel C: from September 16, 2008 to October 31, 2012 for Daily Models; from September.

Panel A						
Rank	Selected times		Average terms		Longest terms	
	Variable	times	Variable	days	Variable	days
1	W_6	1413	W_6	14.57	W_6	230
2	W_{41}	883	W_{41}	10.27	W_5	81
3	W_7	441	W_1	9.09	W_1	71
4	W_1	427	W_7	7.35	W_7	64
5	W_{17}	398	W_2	6.48	W_2	63
6	W_2	324	W_{29}	6.44	W_{17}	62
7	W_{29}	322	W_5	6.42	W_9	61
8	W_{30}	319	W_{39}	6.05	W_{26}	60
9	W_{13}	300	W_9	5.82	W_{33}	52
10	W_9	285	W_{17}	5.45	W_{13}	47
Panel B						
Rank	Selected times		Average terms		Longest terms	
	Variable	times	Variable	days	Variable	days
1	W_6	833	W_6	14.36	W_6	230
2	W_{41}	539	W_7	10.10	W_5	81
3	W_7	303	W_{41}	9.29	W_7	64
4	W_2	229	W_5	9.27	W_2	63
5	W_{29}	226	W_1	8.80	W_{17}	62
6	W_{17}	225	W_{39}	7.56	W_{26}	60
7	W_{30}	222	W_{29}	7.53	W_1	60
8	W_1	220	W_{17}	7.50	W_{33}	52
9	W_5	204	W_2	6.94	W_9	49
10	W_{33}	193	W_{33}	6.23	W_{13}	47
Panel C						
Rank	Selected times		Average terms		Longest terms	
	Variable	times	Variable	days	Variable	days
1	W_6	580	W_6	14.87	W_6	138
2	W_{41}	344	W_{41}	12.29	W_1	71
3	W_1	207	W_1	9.41	W_{17}	61
4	W_{17}	173	W_{18}	6.05	W_9	61
5	W_9	142	W_2	5.59	W_{15}	44
6	f_4	138	W_{34}	5.56	W_{34}	42
7	W_7	138	W_9	5.46	W_{44}	38
8	f_6	132	W_{24}	5.21	W_2	37
9	W_{18}	127	W_{15}	5.18	W_{41}	36
10	W_{24}	125	W_{44}	5.06	W_{24}	32

Table 3: List of the top-10 variables in Daily Model, conditional mean prediction ($h = 20$), in order of the number and average times selected and the longest period for which the variable was continuously selected. Prediction periods; Panel A: from December 30, 1999 to October 31, 2012 for Daily Models; from December, 1995 to October, 2012 for Monthly Models. Panel B: from December 30, 1999 to September 12, 2008, for Daily Models; from December, 1995 to August, 2008 for Monthly Models. Panel C: from September 16, 2008 to October 31, 2012 for Daily Models; from September.

Panel A

Rank	Selected times		Average terms		Longest terms	
	Variable	times	Variable	days	Variable	days
1	W_{41}	1349	W_{41}	8.19	W_{15}	102
2	W_4	1189	W_4	5.36	W_2	87
3	W_{12}	1147	W_{19}	5.32	W_{30}	86
4	W_5	1100	W_7	5.27	W_{32}	73
5	W_{14}	1082	W_2	5.13	W_7	71
6	W_2	1078	W_{24}	4.94	W_5	70
7	W_{19}	1069	W_5	4.82	W_1	68
8	W_{32}	1066	W_{20}	4.82	W_9	60
9	W_1	1053	W_{14}	4.79	W_{24}	59
10	W_{30}	1044	W_{32}	4.78	W_{41}	54

Panel B

Rank	Selected times		Average terms		Longest terms	
	Variable	times	Variable	days	Variable	days
1	W_4	885	W_{41}	7.69	W_{15}	102
2	W_{41}	799	W_2	6.47	W_2	87
3	W_2	776	W_4	5.94	W_{30}	86
4	W_{14}	757	W_7	5.72	W_{32}	73
5	W_{12}	736	W_{19}	5.57	W_7	71
6	W_5	717	W_{14}	5.49	W_5	70
7	W_{19}	707	W_{15}	5.12	W_1	68
8	W_{30}	698	W_5	5.12	W_9	60
9	W_1	657	W_{30}	4.66	W_{14}	54
10	W_7	629	W_{16}	4.57	W_{16}	53

Panel C

Rank	Selected times		Average terms		Longest terms	
	Variable	times	Variable	days	Variable	days
1	W_{41}	550	W_{41}	9.02	W_{24}	59
2	W_{32}	451	W_{24}	5.85	W_{41}	54
3	W_{12}	411	W_{20}	5.70	W_{27}	50
4	W_1	396	W_{32}	5.13	W_{22}	47
5	W_{18}	388	W_{16}	4.97	W_{14}	45
6	W_{17}	385	W_{22}	4.94	W_4	43
7	W_5	383	W_{19}	4.81	W_7	43
8	W_{15}	364	W_{33}	4.65	W_{16}	42
9	W_{19}	362	W_{18}	4.62	W_{33}	42
10	f_4	362	W_7	4.60	W_{42}	41

Table 4: List of the top-10 variables in Monthly Model, conditional mean prediction ($h = 12$), in order of the number and average times selected and the longest period for which the variable was continuously selected. Prediction periods: Panel A: from December 30, 1999 to October 31, 2012. Panel B: from December 30, 1999 to September 12, 2008. Panel C: from September 16, 2008 to October 31, 2012.

Panel A						
Rank	Variable	Selected times	Variable	Average terms	Variable	Longest terms
		times		days		days
1	W_1	130	W_{55}	9.13	W_1	55
2	W_{38}	124	W_1	6.19	W_{38}	53
3	W_{42}	102	W_6	5.27	W_6	46
4	W_{58}	92	W_{40}	5.13	W_{40}	38
5	W_{17}	85	W_5	4.77	W_{55}	30
6	W_5	83	W_{38}	4.43	W_2	30
7	W_{23}	83	W_2	4.24	W_{11}	28
8	W_6	82	W_{43}	4.13	W_{26}	27
9	W_{26}	81	W_{27}	4.13	W_{42}	26
10	W_{40}	77	W_{23}	4.10	W_{27}	23
Panel B						
Rank	Variable	Selected times	Variable	Average terms	Variable	Longest terms
		times		days		days
1	W_{38}	94	W_{55}	9.0	W_{38}	53
2	W_1	81	W_{40}	7.88	W_6	46
3	W_{42}	80	W_6	6.33	W_{40}	38
4	W_6	79	W_{43}	5.18	W_{55}	30
5	W_{26}	74	W_{26}	4.63	W_2	30
6	W_{17}	71	W_{27}	4.55	W_{11}	28
7	W_{58}	67	W_{38}	4.48	W_{26}	27
8	W_2	65	W_{17}	4.44	W_{27}	23
9	W_{39}	64	W_2	4.33	W_{17}	23
10	W_{40}	63	W_1	4.05	W_5	21
Panel C						
Rank	Variable	Selected times	Variable	Average terms	Variable	Longest terms
		times		days		days
1	W_1	49	W_{90}	20.0	W_1	47
2	W_5	44	W_{18}	20.0	W_{90}	20
3	W_{23}	34	W_7	10.0	W_5	20
4	W_{36}	32	W_5	10.0	W_{18}	20
5	W_{38}	30	W_{89}	9.50	W_{14}	20
6	W_{45}	27	W_{55}	9.50	W_{107}	19
7	W_{107}	27	W_{20}	9.0	W_7	18
8	W_{58}	25	W_{101}	9.0	W_{23}	18
9	W_{48}	24	W_{36}	8.0	W_{89}	17
10	W_{14}	23	W_{23}	6.80	W_{76}	17

Figure 1: Time series plots of the estimated β coefficients for the 20-days-ahead prediction model (Daily Model ($h = 20$)). The horizontal axis represents the number of days since December 30, 1999. Plots are of MILLIQ market factor, as the sum of W_{34} and W_{37} (market capitalization-weighted average of modified ILLIQ, 25-days average base and 1-day base, respectively), MILLIQ SMB factor, as the sum of W_{35} and W_{38} (small-minus-big, 25-days average base and 1-day base, respectively), and of MILLIQ HML factor, as the sum of W_{36} and W_{39} (high-minus-low, 25-days average base and 1-day base, respectively).

MILLIQ Market factor

MILLIQ SMB factor

MILLIQ HML factor

Figure 2: Time series plots of the estimated β coefficients for the 12-month-ahead prediction model (Monthly Model ($h = 12$)). The horizontal axis represents the number of months since December, 1995. Plots are of MILLIQ market factor, as the sum of W_{55} , W_{58} , W_{61} , W_{64} , and W_{67} (market capitalization-weighted average of modified ILLIQ, 1-, 3-, 6- and 12-month differences, respectively), MILLIQ SMB factor, as the sum of W_{56} , W_{59} , W_{62} , W_{65} , and W_{68} (small-minus-big, 1-, 3-, 6- and 12-month differences, respectively), and of MILLIQ HML factor, as the sum of W_{57} , W_{60} , W_{63} , W_{66} , and W_{69} (high-minus-low, 1-, 3-, 6- and 12-month differences, respectively).

MILLIQ Market factor

MILLIQ SMB factor

MILLIQ HML factor

Figure 3: Time series plots of the estimated β coefficients for the 12-month-ahead prediction model (Monthly Model ($h = 12$)). The horizontal axis represents the number of months since December, 1995. Plots are of ICC market factor, as the sum of W_{40} , W_{43} , W_{46} , W_{49} , and W_{52} (market capitalization-weighted average of implied cost of capital, 1-, 3-, 6- and 12-month differences, respectively), ICC SMB factor, as the sum of W_{41} , W_{44} , W_{47} , W_{50} , and W_{53} (small-minus-big, 1-, 3-, 6- and 12-month differences, respectively), and of ICC HML factor, as the sum of W_{42} , W_{45} , W_{48} , W_{51} , and W_{54} (high-minus-low, 1-, 3-, 6- and 12-month differences, respectively).

ICC Market factor

ICC SMB factor

ICC HML factor