STAT 222 - Feature Selection

AUTHOR
Kenneth Chen

Recovery Stage

Stepwise Selection - Low Return

The table below contains the following regression information for the following stepwise selections:

1. Backward AIC
2. Backward BIC
3. Forward AIC
4. Forward BIC

[1] "R^2 for Full Model: 0.773" [1] "Adjusted R^2 for Full Model: 0.756"

	Dependent variable:			
	`R1−RF`			
	(1)	(2)	(3)	(4)
CMA	0.233^{***}	0.216^{***}	0.216^{***}	0.216^{***}
	(0.061)	(0.054)	(0.054)	(0.054)
CRD	-0.043^{*}			
	(0.025)			
LIQ	-0.154^{***}	-0.157^{***}	-0.157^{***}	-0.157^{***}
	(0.028)	(0.023)	(0.023)	(0.023)
ME	0.135^{***}	0.155^{***}	0.155^{***}	0.155^{***}
	(0.041)	(0.040)	(0.040)	(0.040)
`Mkt−RF`	0.579^{***}	0.570^{***}	0.570^{***}	0.570^{***}
	(0.028)	(0.027)	(0.027)	(0.027)
RMW	-0.118^{**}			
	(0.056)			
ROE	0.084^{*}			
	(0.050)			
Constant	0.259^{***}	0.235^{***}	0.235^{***}	0.235^{***}
	(0.052)	(0.050)	(0.050)	(0.050)
Observations	175	175	175	175
R^2	0.772	0.764	0.764	0.764
Adjusted R^2	0.762	0.759	0.759	0.759

Residual Std. Error	0.628 (df = 167)	0.633 (df = 170)	0.633 (df = 170)	0.633 (df = 170)
F Statistic	80.676*** (df = 7; 167)	137.864*** (df = 4; 170)	137.864*** (df = 4; 170)	137.864*** (df = 4; 170)
Note:				*p<0.1;* ***p<0.05;*** p<0.01

[1] "Selected Factors: CMA, LIQ, ME, `Mkt−RF`"

Stepwise Selection - Medium Return

The table below contains the following regression information for the following stepwise selections:

1. Backward AIC
2. Backward BIC
3. Forward AIC
4. Forward BIC

[1] "R^2 for Full Model: 0.742" [1] "Adjusted R^2 for Full Model: 0.723"

	Dependent variable:			
	`R2−RF`			
	(1)	(2)	(3)	(4)
CMA	0.189***	0.189***		
	(0.062)	(0.062)		
LIQ	-0.132***	-0.132***	-0.098***	-0.098***
	(0.026)	(0.026)	(0.026)	(0.026)
ME	0.171***	0.171***		
	(0.045)	(0.045)		
PC2			-0.080***	-0.080***
			(0.026)	(0.026)
`Mkt−RF`	0.591***	0.591***	0.619***	0.619***
	(0.031)	(0.031)	(0.036)	(0.036)
PC3			0.132***	0.132***
			(0.029)	(0.029)
Constant	0.484***	0.484***	0.504***	0.504***
	(0.058)	(0.058)	(0.059)	(0.059)
Observations	175	175	175	175
R^2	0.739	0.739	0.733	0.733
Adjusted R^2	0.732	0.732	0.726	0.726
Residual Std. Error (df = 170)	0.726	0.726	0.735	0.735
F Statistic (df = 4; 170)	120.089***	120.089***	116.454***	116.454***

Note:	$p<0.1$; $p<0.05$; $p<0.01$

[1] "Selected Factors: CMA, LIQ, ME, `Mkt-RF`"

Stepwise Selection - High Return

The table below contains the following regression information for the following stepwise selections:

1. Backward AIC
2. Backward BIC
3. Forward AIC
4. Forward BIC

[1] "R^2 for Full Model: 0.694" [1] "Adjusted R^2 for Full Model: 0.671"

	Dependent variable:			
	`R3-RF`			
	(1)	(2)	(3)	(4)
CMA	0.162^{**}			
	(0.076)			
LIQ	-0.103^{***}	-0.115^{***}	-0.125^{**}	
	(0.033)	(0.032)	(0.053)	
ME	0.195^{***}	0.240^{***}		
	(0.056)	(0.053)		
MOM			-0.069	
			(0.048)	
`Mkt-RF`	0.620^{***}	0.603^{***}	0.607^{***}	0.605^{***}
	(0.039)	(0.038)	(0.052)	(0.039)
PC3			0.166^{***}	0.141^{***}
			(0.041)	(0.036)
PC2			-0.017	-0.083^{***}
			(0.051)	(0.032)
Constant	0.709^{***}	0.709^{***}	0.775^{***}	0.715^{***}
	(0.071)	(0.072)	(0.077)	(0.072)
Observations	175	175	175	175
R^2	0.689	0.681	0.686	0.674
Adjusted R^2	0.682	0.676	0.676	0.668
Residual Std. Error	0.898 (df = 170)	0.907 (df = 171)	0.906 (df = 169)	0.917 (df = 171)
F Statistic	94.299^{***} (df = 4; 170)	121.741^{***} (df = 3; 171)	73.691^{***} (df = 5; 169)	117.909^{***} (df = 3; 171)

Note:	*p<0.1;* ***p<0.05;*** p<0.01

[1] “Selected Factors: LIQ, ME, `Mkt-RF` ”

Expansion Stage

Stepwise Selection - Low Return

The table below contains the following regression information for the following stepwise selections:

1. Backward AIC
2. Backward BIC
3. Forward AIC
4. Forward BIC

[1] “R^2 for Full Model: 0.675” [1] “Adjusted R^2 for Full Model: 0.669”

	Dependent variable:			
	`R1-RF`			
	(1)	(2)	(3)	(4)
CRD	-0.053**			
	(0.025)			
HML	0.288***	0.324***	0.214***	0.282***
	(0.037)	(0.034)	(0.045)	(0.035)
MOM	-0.062**	-0.083***		
	(0.027)	(0.025)		
PC6			0.346***	0.315***
			(0.048)	(0.046)
PC7			-0.127***	-0.137***
			(0.040)	(0.040)
ROE			-0.268***	-0.272***
			(0.056)	(0.057)
`Mkt-RF`	0.803***	0.795***	0.690***	0.684***
	(0.023)	(0.023)	(0.024)	(0.024)
RMW	0.316***	0.309***	0.162***	0.160***
	(0.041)	(0.041)	(0.054)	(0.054)
SMB	0.283***	0.285***		
	(0.035)	(0.035)		
CMA			0.152**	
			(0.064)	

Constant	0.132***	0.119***	0.194***	0.195***
	(0.040)	(0.040)	(0.040)	(0.040)
Observations	599	599	599	599
R^2	0.672	0.670	0.672	0.669
Adjusted R^2	0.669	0.667	0.668	0.666
Residual Std. Error	0.943 (df = 592)	0.946 (df = 593)	0.944 (df = 591)	0.947 (df = 592)
F Statistic	202.516*** (df = 6; 592)	240.674*** (df = 5; 593)	173.266*** (df = 7; 591)	199.625*** (df = 6; 592)
Note:				*p<0.1;* ***p<0.05;*** p<0.01

[1] "Selected Factors: HML, MOM, `Mkt-RF`, RMW, SMB"

Stepwise Selection - Medium Return

The table below contains the following regression information for the following stepwise selections:

1. Backward AIC
2. Backward BIC
3. Forward AIC
4. Forward BIC

[1] "R^2 for Full Model: 0.575" [1] "Adjusted R^2 for Full Model: 0.566"

	Dependent variable:			
	`R2-RF`			
	(1)	(2)	(3)	(4)
HML	0.315***	0.328***	0.207***	0.290***
	(0.047)	(0.044)	(0.060)	(0.045)
LIQ	0.074**			
	(0.033)			
PC6			0.554***	0.375***
			(0.097)	(0.061)
`Mkt-RF`	0.826***	0.851***	0.693***	0.754***
	(0.034)	(0.031)	(0.039)	(0.031)
RMW	0.398***	0.275***		
	(0.076)	(0.056)		
ROE	-0.131*		-0.515***	-0.216***
	(0.073)		(0.146)	(0.060)
SMB	0.255***	0.293***		
	(0.049)	(0.047)		

PC5			0.180***	
			(0.063)	
PC2			0.099*	
			(0.057)	
Constant	0.316***	0.320***	0.462***	0.407***
	(0.054)	(0.054)	(0.058)	(0.054)
Observations	599	599	599	599
R^2	0.573	0.567	0.571	0.564
Adjusted R^2	0.568	0.564	0.566	0.561
Residual Std. Error	1.280 (df = 592)	1.287 (df = 594)	1.283 (df = 592)	1.291 (df = 594)
F Statistic	132.302*** (df = 6; 592)	194.374*** (df = 4; 594)	131.097*** (df = 6; 592)	192.112*** (df = 4; 594)
Note:				*p<0.1;* ***p<0.05;*** p<0.01

[1] "Selected Factors: HML, `Mkt-RF`, RMW, SMB"

Stepwise Selection - High Return

The table below contains the following regression information for the following stepwise selections:

1. Backward AIC
2. Backward BIC
3. Forward AIC
4. Forward BIC

[1] "R^2 for Full Model: 0.46" [1] "Adjusted R^2 for Full Model: 0.448"

	Dependent variable:			
	`R3-RF`			
	(1)	(2)	(3)	(4)
HML	0.351***	0.351***	0.361***	0.351***
	(0.064)	(0.064)	(0.064)	(0.064)
LIQ	0.190***	0.190***	0.166***	0.185***
	(0.047)	(0.047)	(0.054)	(0.040)
PC6			-0.171	0.216***
			(0.160)	(0.072)
`Mkt-RF`	0.862***	0.862***	0.942***	0.796***
	(0.048)	(0.048)	(0.073)	(0.046)
RMW	0.282***	0.282***	0.473**	
	(0.083)	(0.083)	(0.184)	

PC7			0.159	
			(0.109)	
SMB	0.231***	0.231***	0.391***	
	(0.070)	(0.070)	(0.136)	
Constant	0.497***	0.497***	0.463***	0.546***
	(0.077)	(0.077)	(0.083)	(0.077)
Observations	599	599	599	599
R^2	0.455	0.455	0.457	0.449
Adjusted R^2	0.451	0.451	0.451	0.446
Residual Std. Error	1.838 (df = 593)	1.838 (df = 593)	1.838 (df = 591)	1.847 (df = 594)
F Statistic	99.112*** (df = 5; 593)	99.112*** (df = 5; 593)	71.147*** (df = 7; 591)	121.112*** (df = 4; 594)
Note:				*p<0.1;* ***p<0.05;*** p<0.01

[1] “Selected Factors: `Mkt−RF`, HML, LIQ, PC6”

Downturn Stage

Stepwise Selection - Low Return

The table below contains the following regression information for the following stepwise selections:

1. Backward AIC
2. Backward BIC
3. Forward AIC
4. Forward BIC

[1] “R^2 for Full Model: 0.6” [1] “Adjusted R^2 for Full Model: 0.593”

	Dependent variable:			
	`R1−RF`			
	(1)	(2)	(3)	(4)
CRD	-0.105***	-0.135***	-0.110***	-0.110***
	(0.027)	(0.024)	(0.026)	(0.026)
IA	0.120**		0.139***	0.139***
	(0.050)		(0.049)	(0.049)
LIQ	-0.078***	-0.084***		
	(0.027)	(0.027)		
ME	-0.189			

	(0.120)			
Mkt−RF	0.668***	0.647***	0.616***	0.616***
	(0.025)	(0.023)	(0.022)	(0.022)
RMW	0.220***	0.218***	0.205***	0.205***
	(0.048)	(0.048)	(0.045)	(0.045)
SMB	0.286**	0.098***		
	(0.122)	(0.037)		
PC5			0.121***	0.121***
			(0.033)	(0.033)
Constant	0.152***	0.166***	0.152***	0.152***
	(0.042)	(0.042)	(0.042)	(0.042)
Observations	669	669	669	669
R^2	0.598	0.594	0.597	0.597
Adjusted R^2	0.594	0.591	0.593	0.593
Residual Std. Error	1.057 (df = 661)	1.060 (df = 663)	1.057 (df = 663)	1.057 (df = 663)
F Statistic	140.461*** (df = 7; 661)	193.878*** (df = 5; 663)	196.037*** (df = 5; 663)	196.037*** (df = 5; 663)
Note:				*p<0.1;* ***p<0.05;*** p<0.01

[1] "Selected Factors: `Mkt−RF`, CRD, RMW, PC5, IA"

Stepwise Selection - Medium Return

The table below contains the following regression information for the following stepwise selections:

1. Backward AIC
2. Backward BIC
3. Forward AIC
4. Forward BIC

[1] "R^2 for Full Model: 0.447" [1] "Adjusted R^2 for Full Model: 0.437"

	Dependent variable:			
	R2−RF			
	(1)	(2)	(3)	(4)
CMA	-0.265			
	(0.172)			
CRD	-0.091**	-0.118***	-0.067*	
	(0.039)	(0.034)	(0.039)	
IA	0.313**			

	(0.150)			
LIQ	-0.088**			
	(0.039)			
Mkt−RF	0.694***	0.660***	0.780***	0.844***
	(0.034)	(0.030)	(0.060)	(0.054)
PC1			-0.142**	-0.240***
			(0.059)	(0.044)
RMW	0.243***	0.253***	0.176**	
	(0.070)	(0.064)	(0.078)	
SMB	0.114**		0.166**	0.217***
	(0.054)		(0.067)	(0.062)
Constant	0.346***	0.356***	0.344***	0.333***
	(0.061)	(0.060)	(0.060)	(0.060)
Observations	669	669	669	669
R^2	0.446	0.437	0.443	0.437
Adjusted R^2	0.440	0.434	0.439	0.435
Residual Std. Error	1.521 (df = 661)	1.529 (df = 665)	1.523 (df = 663)	1.529 (df = 665)
F Statistic	76.008*** (df = 7; 661)	171.747*** (df = 3; 665)	105.335*** (df = 5; 663)	172.145*** (df = 3; 665)
Note:				*p<0.1;* ***p<0.05;*** p<0.01

[1] "Selected Factors: `Mkt−RF`, PC1, SMB"

Stepwise Selection - High Return

The table below contains the following regression information for the following stepwise selections:

1. Backward AIC
2. Backward BIC
3. Forward AIC
4. Forward BIC

[1] "R^2 for Full Model: 0.32" [1] "Adjusted R^2 for Full Model: 0.308"

	Dependent variable:			
	R3−RF			
	(1)	(2)	(3)	(4)
CMA	-0.451*			
	(0.250)			
IA	0.494**			

	(0.225)			
`Mkt−RF`	0.737***	0.698***	0.757***	0.698***
	(0.047)	(0.041)	(0.046)	(0.041)
RMW	0.264***		0.234**	
	(0.101)		(0.098)	
SMB	0.140*			
	(0.078)			
PC3			-0.098*	
			(0.056)	
Constant	0.541***	0.554***	0.532***	0.554***
	(0.090)	(0.089)	(0.089)	(0.089)
Observations	669	669	669	669
R^2	0.316	0.304	0.311	0.304
Adjusted R^2	0.310	0.303	0.308	0.303
Residual Std. Error	2.287 (df = 663)	2.300 (df = 667)	2.291 (df = 665)	2.300 (df = 667)
F Statistic	61.130*** (df = 5; 663)	290.925*** (df = 1; 667)	100.214*** (df = 3; 665)	290.925*** (df = 1; 667)
Note:				*p<0.1;* ***p<0.05;*** p<0.01

[1] "Selected Factors: `Mkt−RF`"

Depression Stage

Stepwise Selection - Low Return

The table below contains the following regression information for the following stepwise selections:

1. Backward AIC
2. Backward BIC
3. Forward AIC
4. Forward BIC

[1] "R^2 for Full Model: 0.054" [1] "Adjusted R^2 for Full Model: 0.013"

	Dependent variable:			
	`R1−RF`			
	(1)	(2)	(3)	(4)
`Mkt−RF`	2.079***	1.641***	1.641***	1.641***
	(0.699)	(0.582)	(0.582)	(0.582)

RMW	-4.574**			
	(2.142)			
ROE	4.447*			
	(2.331)			
Constant	2.675	2.614	2.614	2.614
	(2.069)	(2.040)	(2.040)	(2.040)
Observations	288	288	288	288
R^2	0.043	0.027	0.027	0.027
Adjusted R^2	0.033	0.024	0.024	0.024
Residual Std. Error	34.442 (df = 284)	34.613 (df = 286)	34.613 (df = 286)	34.613 (df = 286)
F Statistic	4.291*** (df = 3; 284)	7.949*** (df = 1; 286)	7.949*** (df = 1; 286)	7.949*** (df = 1; 286)
Note:				*p<0.1;* ***p<0.05;*** p<0.01

[1] "Selected Factors: `Mkt-RF`, RMW, ROE"

Stepwise Selection - Medium Return

The table below contains the following regression information for the following stepwise selections:

1. Backward AIC
2. Backward BIC
3. Forward AIC
4. Forward BIC

[1] "R^2 for Full Model: 0.054" [1] "Adjusted R^2 for Full Model: 0.012"

	Dependent variable:			
	`R2-RF`			
	(1)	(2)	(3)	(4)
`Mkt-RF`	2.088***	1.620***	1.620***	1.620***
	(0.705)	(0.587)	(0.587)	(0.587)
RMW	-4.612**			
	(2.160)			
ROE	4.595*			
	(2.350)			
Constant	3.017	2.981	2.981	2.981
	(2.086)	(2.058)	(2.058)	(2.058)
Observations	288	288	288	288
R^2	0.043	0.026	0.026	0.026
Adjusted R^2	0.032	0.023	0.023	0.023
Residual Std. Error	34.726 (df = 284)	34.903 (df = 286)	34.903 (df = 286)	34.903 (df = 286)

F Statistic	4.208*** (df = 3; 284)	7.623*** (df = 1; 286)	7.623*** (df = 1; 286)	7.623*** (df = 1; 286)
Note:				*p<0.1;* ***p<0.05;*** p<0.01

[1] "Selected Factors: `Mkt−RF`, RMW, ROE"

Stepwise Selection - High Return

The table below contains the following regression information for the following stepwise selections:

1. Backward AIC
2. Backward BIC
3. Forward AIC
4. Forward BIC

[1] "R^2 for Full Model: 0.043" [1] "Adjusted R^2 for Full Model: 0.002"

	Dependent variable:			
	`R3−RF`			
	(1)	(2)	(3)	(4)
`Mkt−RF`	2.894**		2.124**	
	(1.190)		(0.992)	
RMW	-8.114**			
	(3.649)			
ROE	7.853**			
	(3.970)			
Constant	4.974	4.636	4.858	4.636
	(3.523)	(3.498)	(3.478)	(3.498)
Observations	288	288	288	288
R^2	0.034	0.000	0.016	0.000
Adjusted R^2	0.023	0.000	0.012	0.000
Residual Std. Error	58.660 (df = 284)	59.358 (df = 287)	58.991 (df = 286)	59.358 (df = 287)
F Statistic	3.293** (df = 3; 284)		4.586** (df = 1; 286)	
Note:				*p<0.1;* ***p<0.05;*** p<0.01

[1] "Selected Factors: `Mkt−RF`, RMW, ROE"