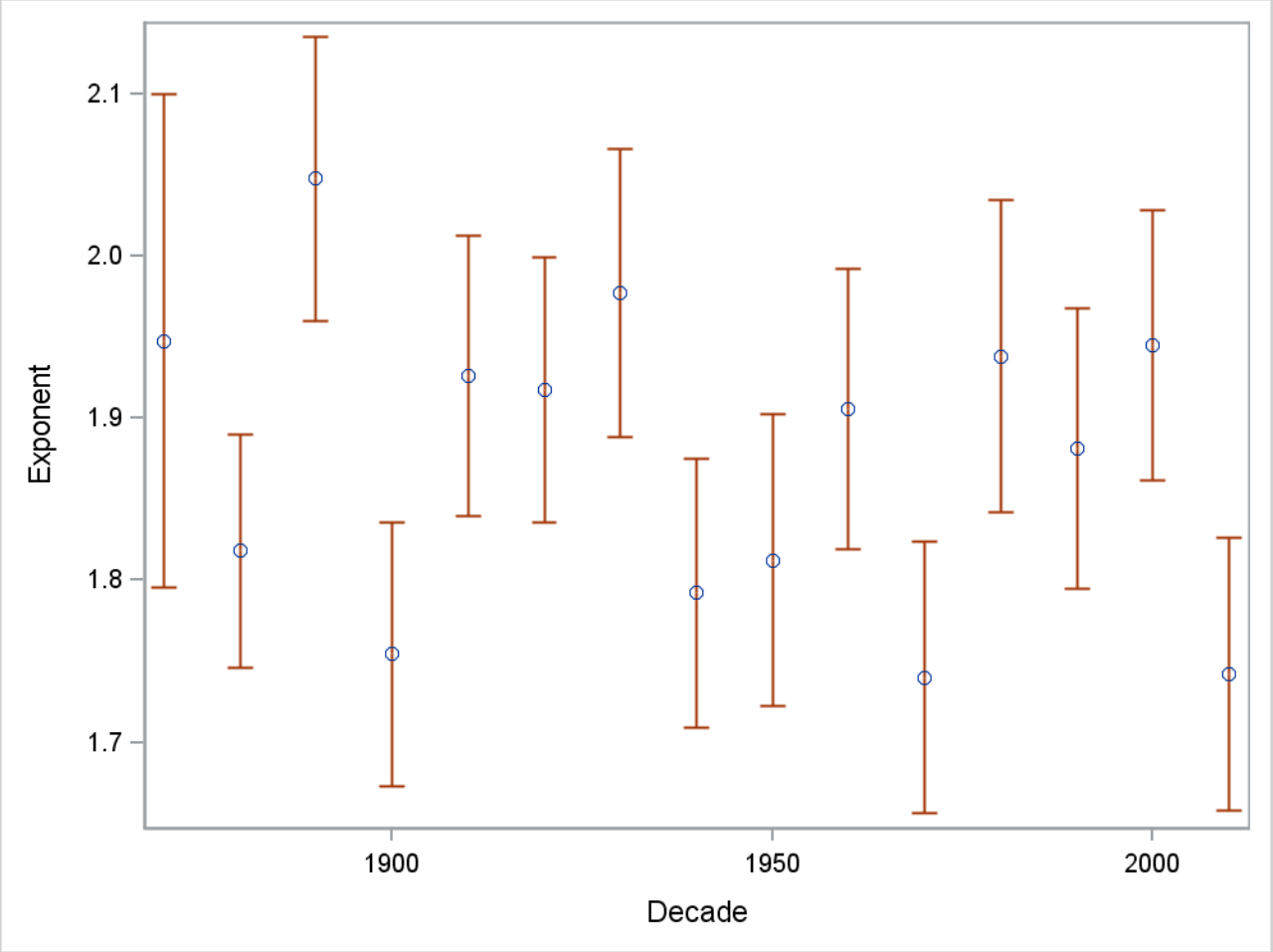
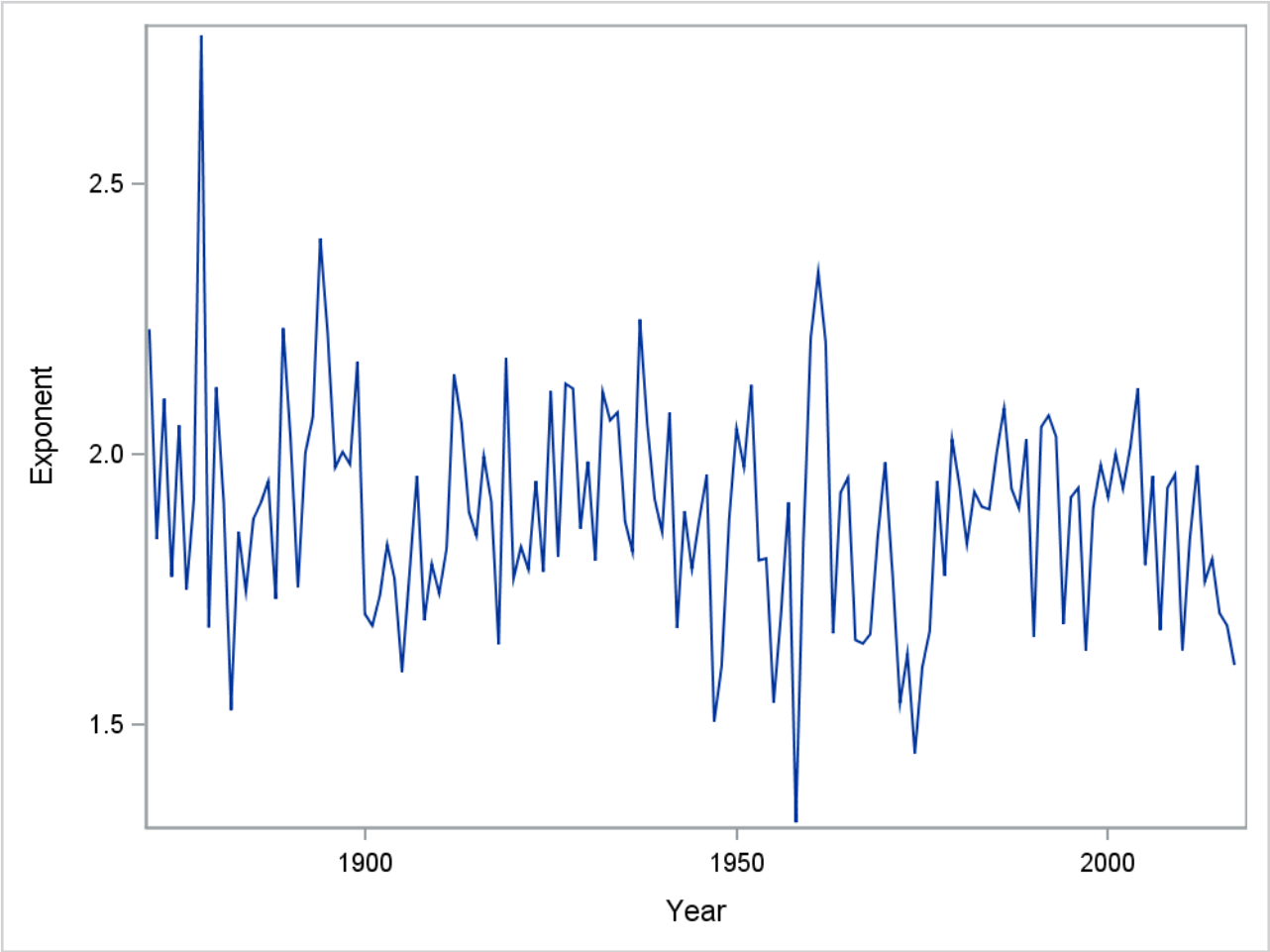
Estimates of Pythogorean Exponent by Decade

Obs	decade	Model	Dependent	Variable	DF	Estimate	StdErr	tValue	Probt	LowerCL	UpperCL
1	1870	MODEL1	log_wl	log_rra	1	1.94769	0.07626	25.54	<.0001	1.79578	2.09960
2	1880	MODEL1	log_wl	log_rra	1	1.81809	0.03650	49.81	<.0001	1.74600	1.89018
3	1890	MODEL1	log_wl	log_rra	1	2.04762	0.04431	46.21	<.0001	1.96000	2.13524
4	1900	MODEL1	log_wl	log_rra	1	1.75445	0.04108	42.71	<.0001	1.67329	1.83560
5	1910	MODEL1	log_wl	log_rra	1	1.92607	0.04399	43.79	<.0001	1.83925	2.01288
6	1920	MODEL1	log_wl	log_rra	1	1.91741	0.04154	46.15	<.0001	1.83536	1.99946
7	1930	MODEL1	log_wl	log_rra	1	1.97738	0.04498	43.96	<.0001	1.88854	2.06622
8	1940	MODEL1	log_wl	log_rra	1	1.79214	0.04186	42.81	<.0001	1.70946	1.87481
9	1950	MODEL1	log_wl	log_rra	1	1.81236	0.04572	39.64	<.0001	1.72206	1.90267
10	1960	MODEL1	log_wl	log_rra	1	1.90554	0.04396	43.34	<.0001	1.81884	1.99224
11	1970	MODEL1	log_wl	log_rra	1	1.73991	0.04245	40.99	<.0001	1.65630	1.82351
12	1980	MODEL1	log_wl	log_rra	1	1.93812	0.04885	39.68	<.0001	1.84193	2.03430
13	1990	MODEL1	log_wl	log_rra	1	1.88133	0.04398	42.78	<.0001	1.79475	1.96790
14	2000	MODEL1	log_wl	log_rra	1	1.94508	0.04248	45.79	<.0001	1.86148	2.02869
15	2010	MODEL1	log_wl	log_rra	1	1.74223	0.04283	40.68	<.0001	1.65785	1.82661



Error bars indicate the 95% CI for the exponent estimation. Notice no particular pattern throughout.

Estimates of Pythagorean Exponent by Year



The REG Procedure
Model: MODEL1
Dependent Variable: Estimate Parameter Estimate

Number of Observations Read	147
Number of Observations Used	147

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.17392	0.17392	4.44	0.0368
Error	145	5.67803	0.03916		
Corrected Total	146	5.85195			

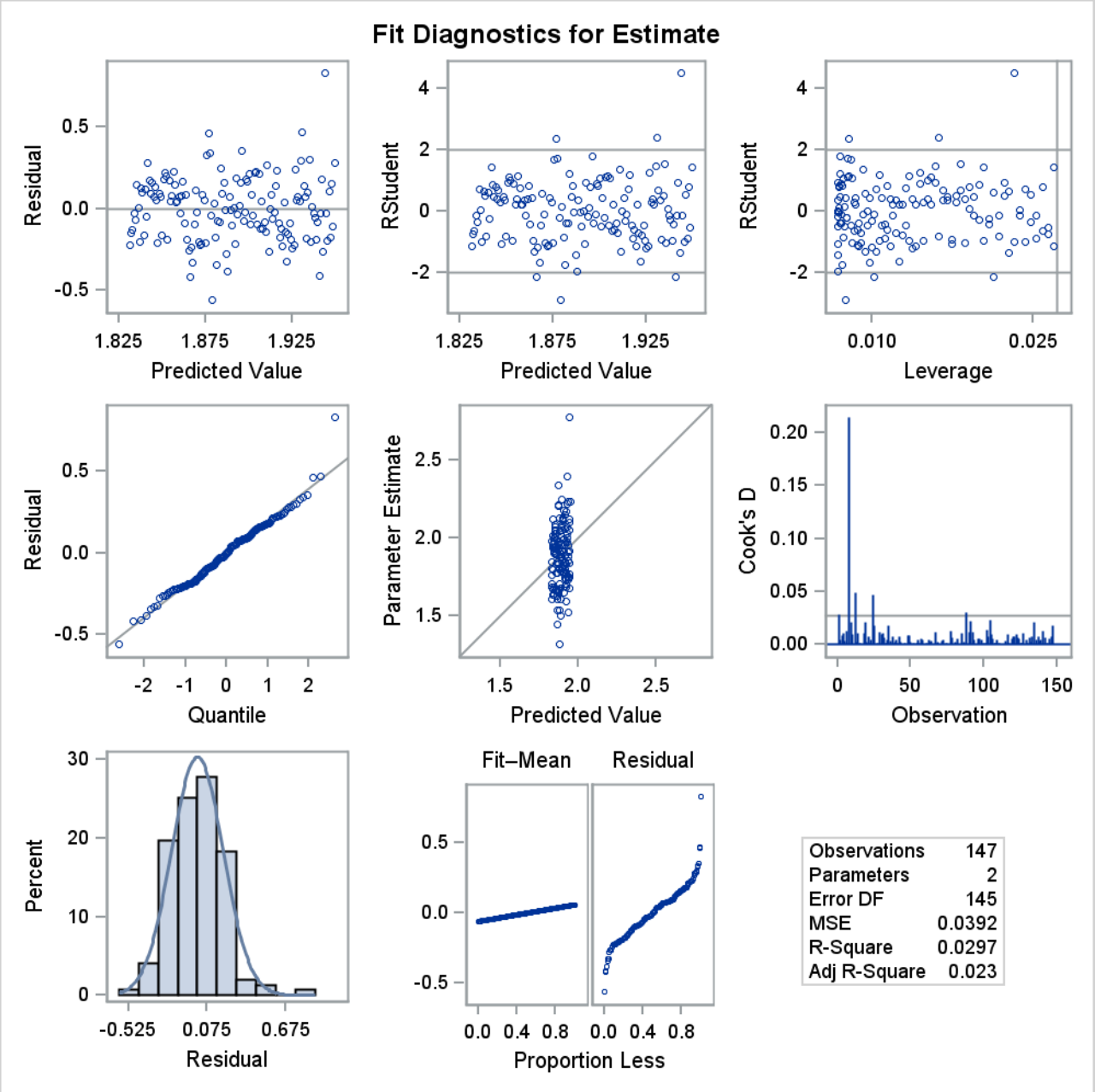
Root MSE	0.19789	R-Square	0.0297
Dependent Mean	1.89059	Adj R-Sq	0.0230
Coeff Var	10.46686		

Linear Regression of Pythagorean Exponent as a Function of Year

The REG Procedure
Model: MODEL1
Dependent Variable: Estimate Parameter Estimate

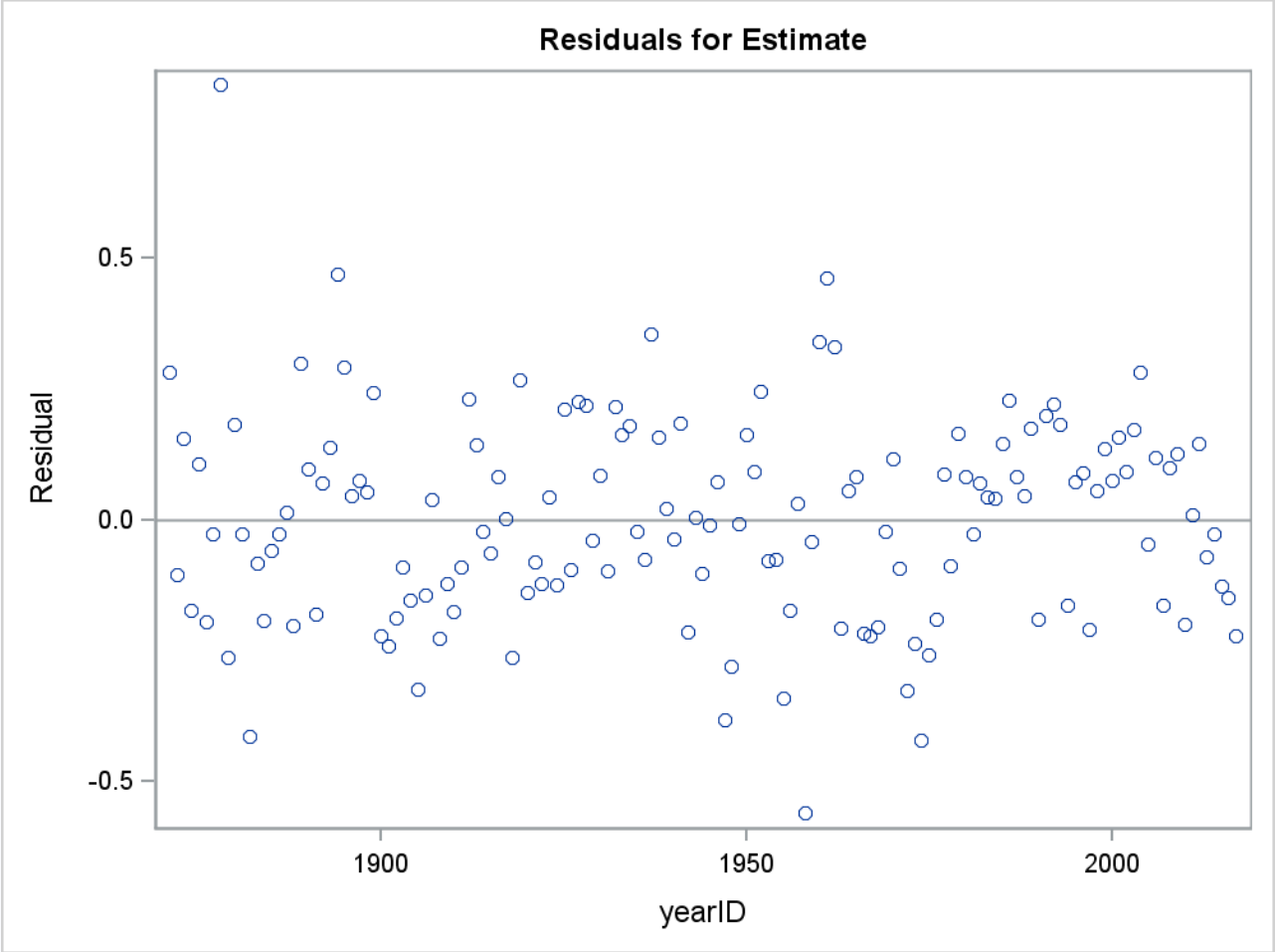
Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	95% Confidence Limits	
Intercept	Intercept	1	3.46639	0.74789	4.63	<.0001	1.98821	4.94457
yearID		1	-0.00081059	0.00038463	-2.11	0.0368	-0.00157	-0.00005039



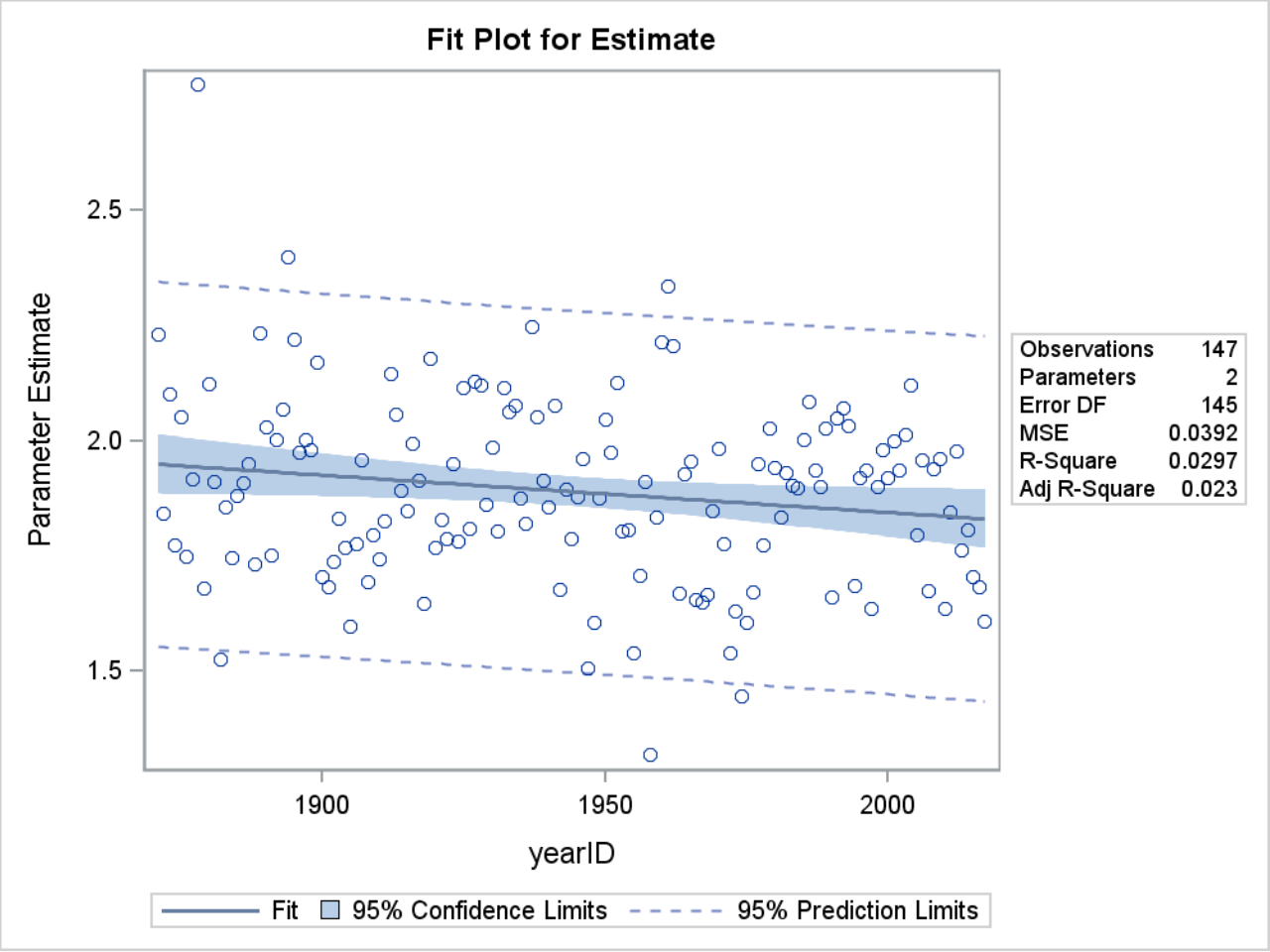
Linear Regression of Pythagorean Exponent as a Function of Year

The REG Procedure
Model: MODEL1
Dependent Variable: Estimate Parameter Estimate



Linear Regression of Pythagorean Exponent as a Function of Year

The REG Procedure
Model: MODEL1
Dependent Variable: Estimate Parameter Estimate



The REG Procedure
Model: MODEL1
Dependent Variable: Estimate Parameter Estimate

Number of Observations Read	146
Number of Observations Used	146

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.09331	0.09331	2.70	0.1024
Error	144	4.97273	0.03453		
Corrected Total	145	5.06604			

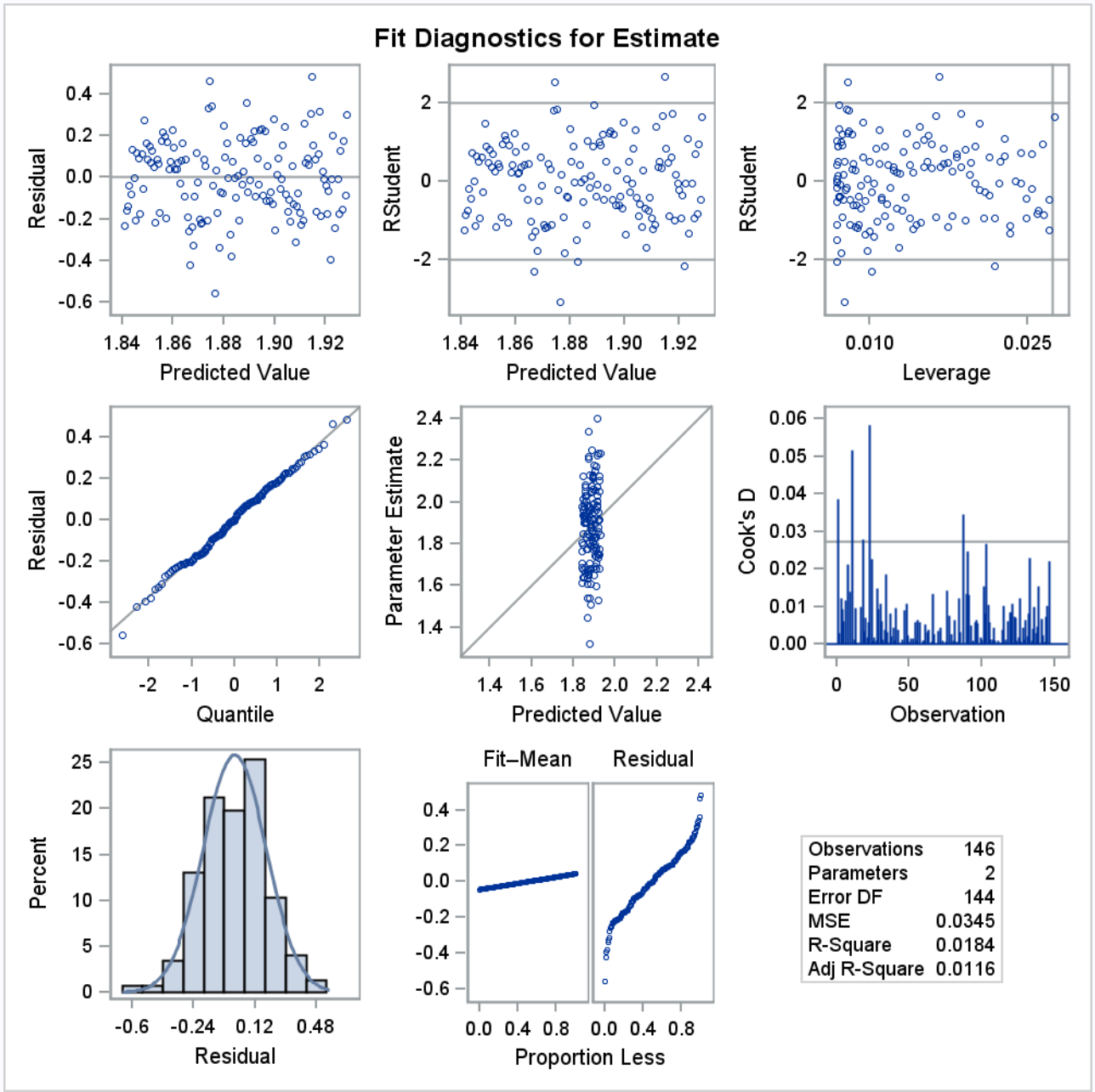
Root MSE	0.18583	R-Square	0.0184
Dependent Mean	1.88454	Adj R-Sq	0.0116
Coeff Var	9.86075		

Linear Regression of Pythagorean Exponent as a Function of Year (without 1878 outlier)

The REG Procedure
Model: MODEL1
Dependent Variable: Estimate Parameter Estimate

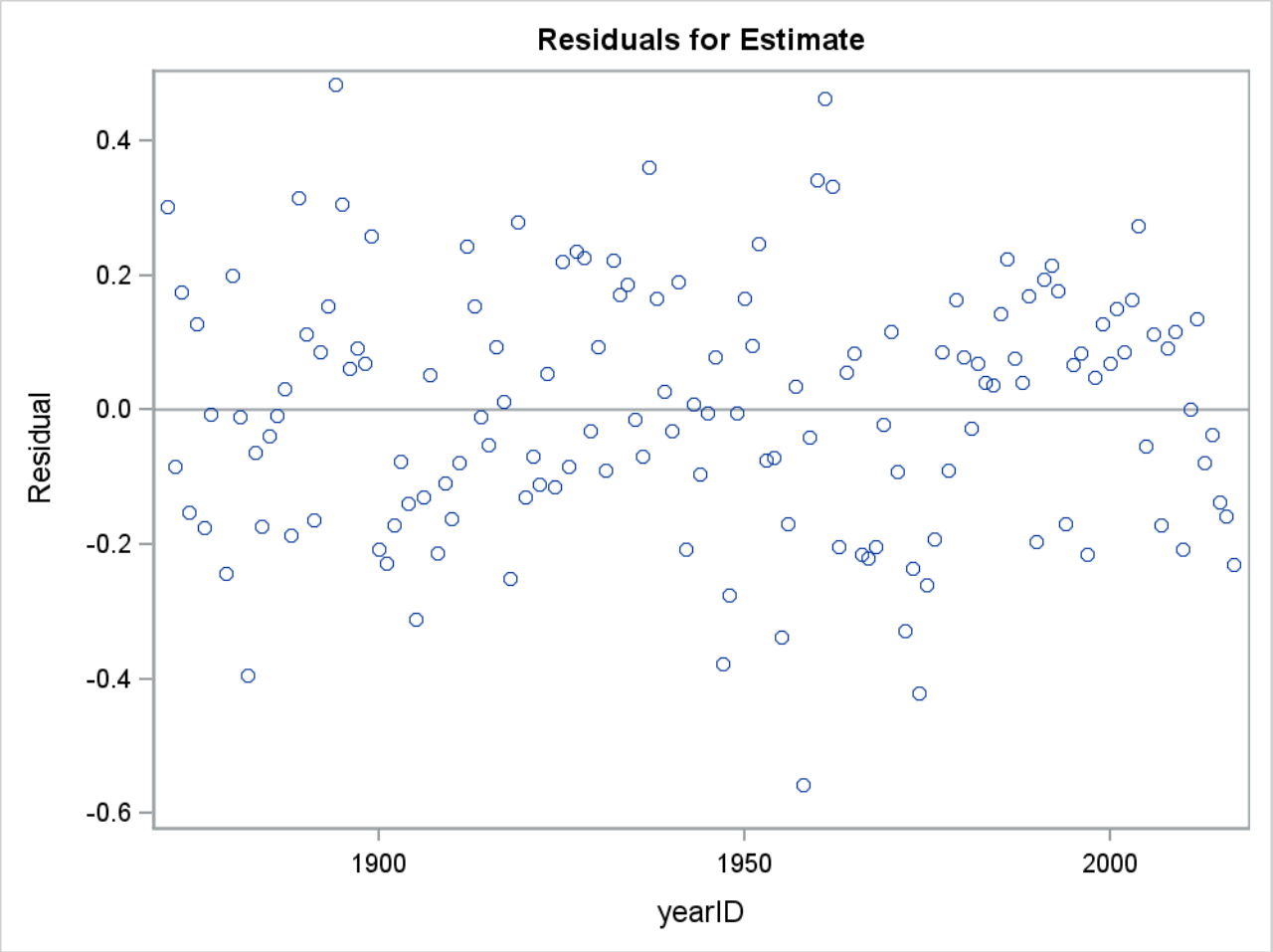
Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	95% Confidence Limits	
Intercept	Intercept	1	3.04871	0.70838	4.30	<.0001	1.64854	4.44889
yearID		1	-0.00059871	0.00036422	-1.64	0.1024	-0.00132	0.00012120



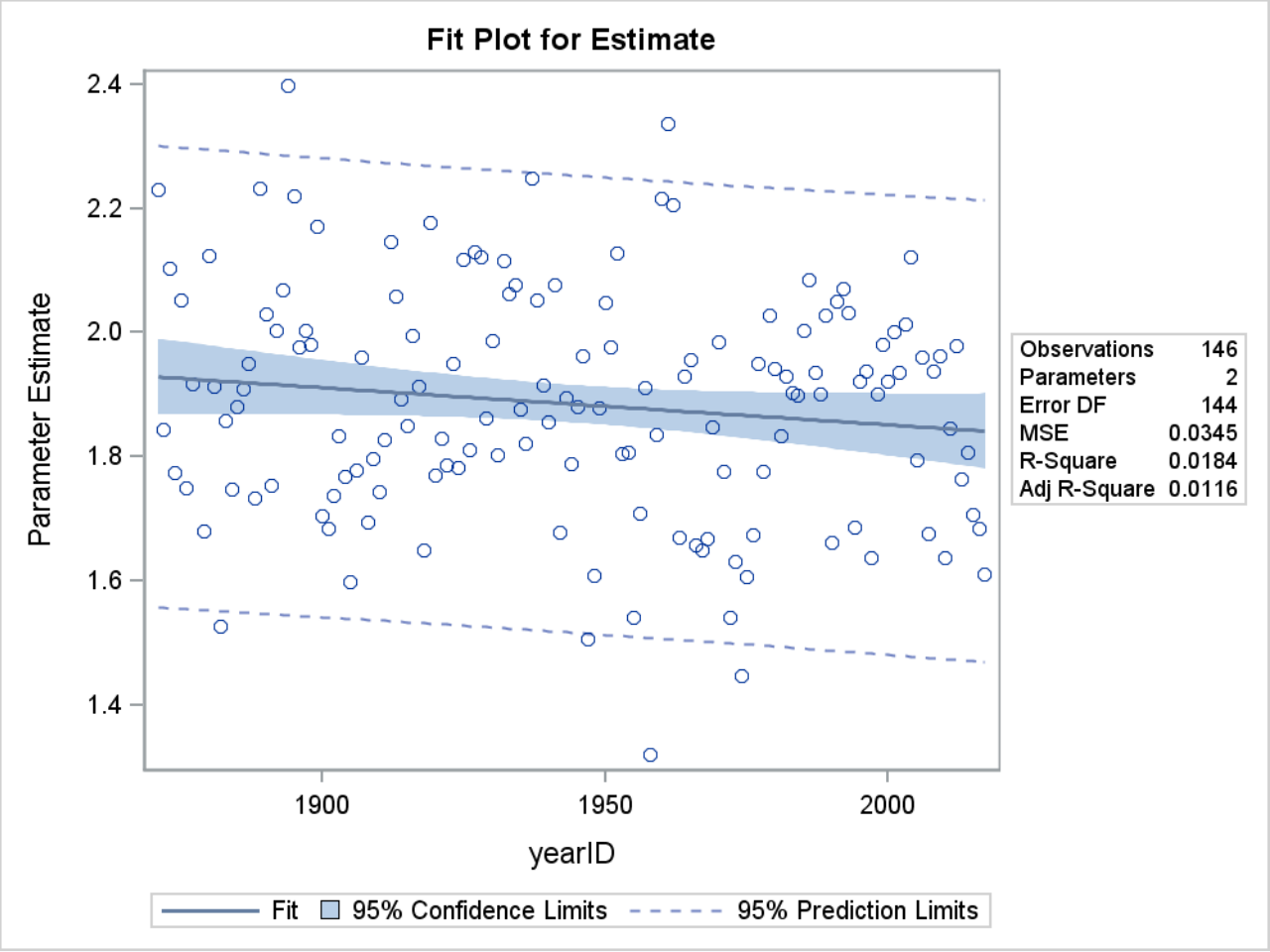
Linear Regression of Pythagorean Exponent as a Function of Year (without 1878 outlier)

The REG Procedure
Model: MODEL1
Dependent Variable: Estimate Parameter Estimate



Linear Regression of Pythagorean Exponent as a Function of Year (without 1878 outlier)

The REG Procedure
Model: MODEL1
Dependent Variable: Estimate Parameter Estimate



When the 1878 outlier is removed, the slope is no longer significant. Safe to assume variation in exponent is due to error alone.

The REG Procedure
Model: MODEL1
Dependent Variable: log_wl

Number of Observations Read	2865
Number of Observations Used	2863
Number of Observations with Missing Values	2

Note: No intercept in model. R-Square is redefined.

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	79.01866	79.01866	27702.6	<.0001
Error	2862	8.16355	0.00285		
Uncorrected Total	2863	87.18221			

Root MSE	0.05341	**R-Square**	0.9064
Dependent Mean	-0.00363	**Adj R-Sq**	0.9063
Coeff Var	-1470.81390		

Linear Regression to get Overall Exponent

The REG Procedure
Model: MODEL1
Dependent Variable: log_wl

Parameter Estimates							
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	95% Confidence Limits	
log_rra	1	1.87752	0.01128	166.44	<.0001	1.85540	1.89964

Make note of the differences in the overall exponent estimation from the linear regression method and the following nonlinear method, which iteratively determines the exponent as it belongs in the formula.

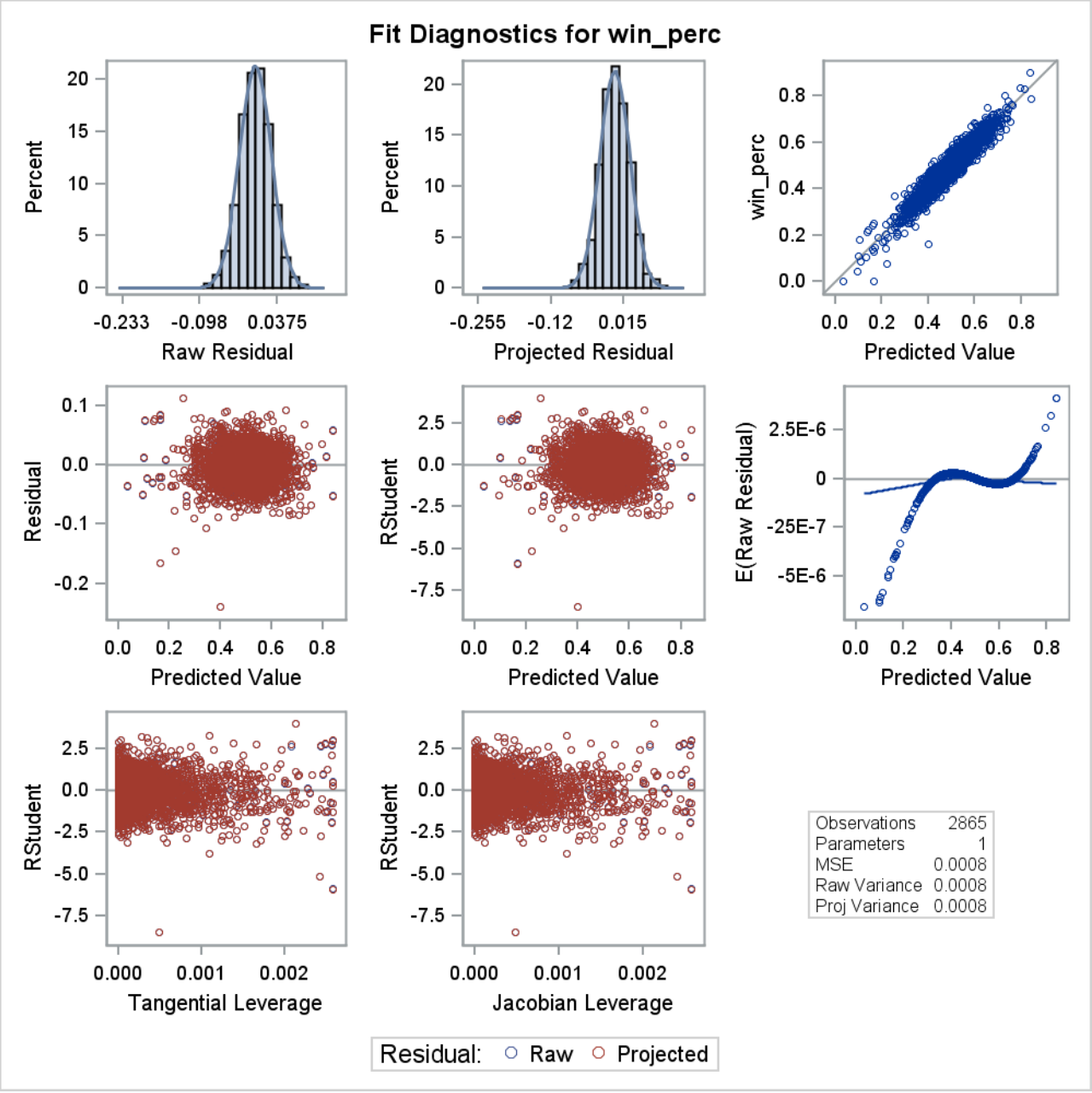
The NLIN Procedure
Dependent Variable win_perc
Method: Gauss-Newton

Iterative Phase		
Iter	x	Sum of Squares
0	2.0000	2.3651
1	1.8670	2.2699
2	1.8681	2.2699

NOTE: Convergence criterion met.

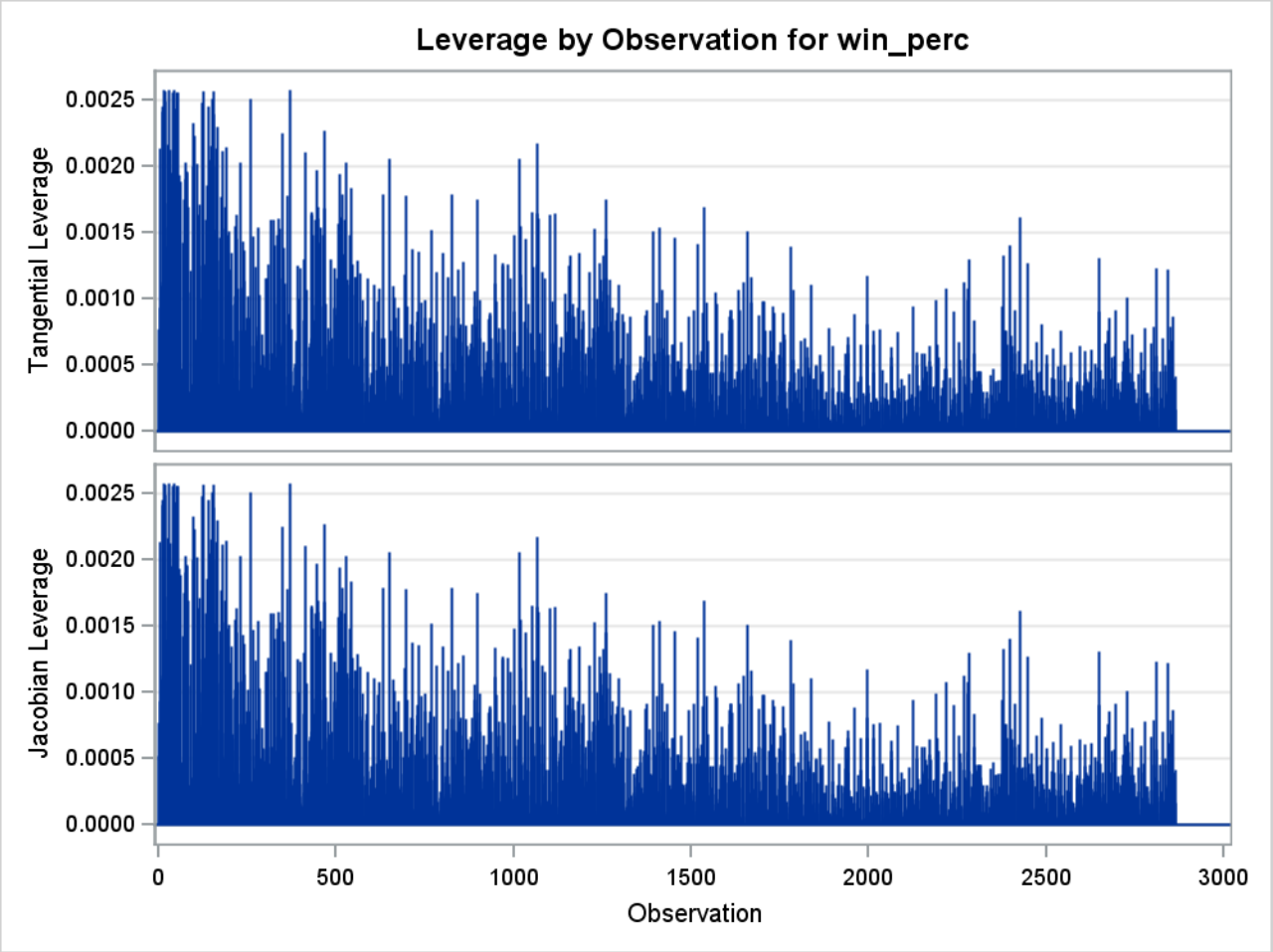
Estimating Pythagorean Exponent with Nonlinear Method

The NLIN Procedure
Dependent Variable win_perc
Method: Gauss-Newton



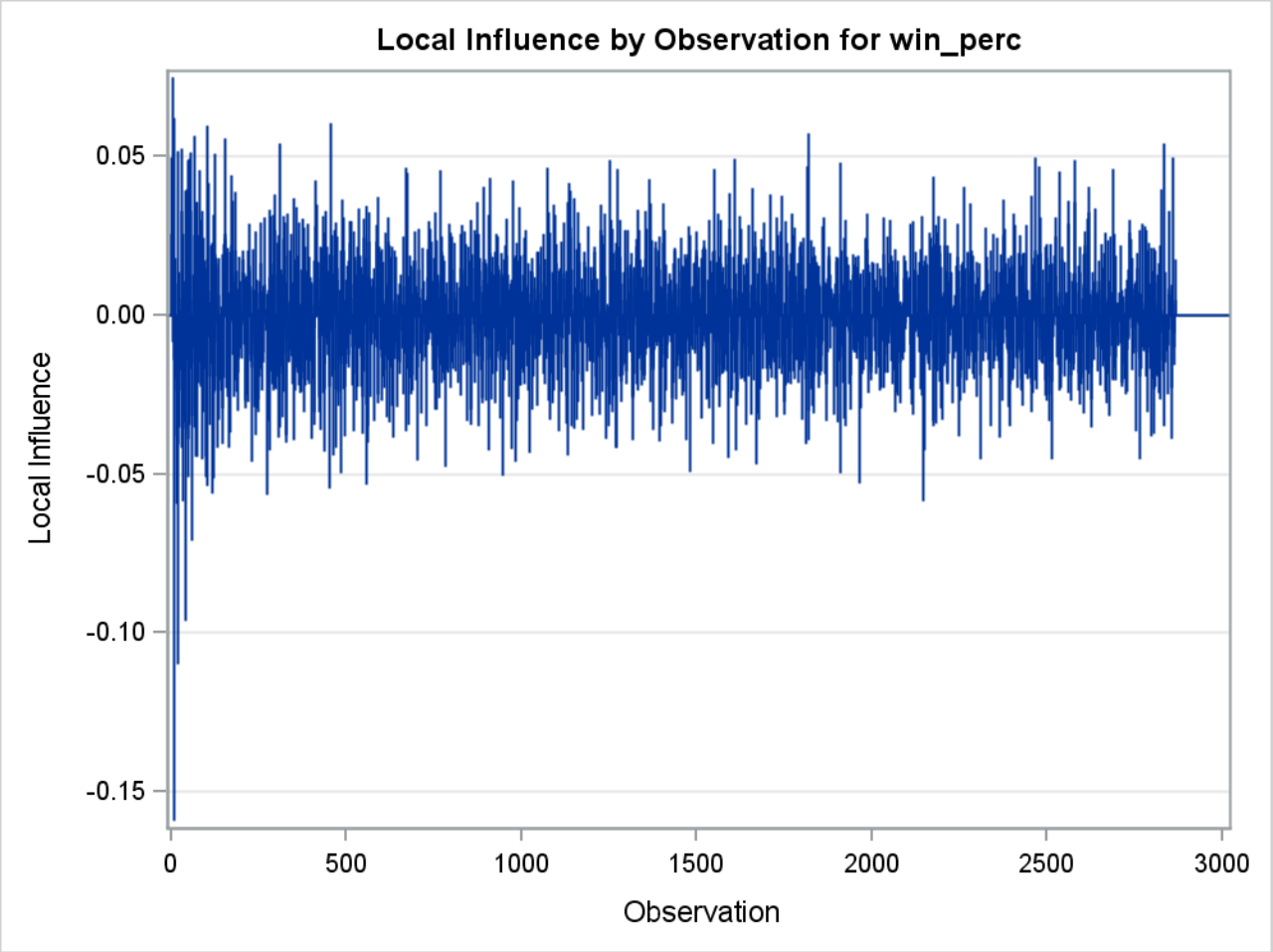
Estimating Pythagorean Exponent with Nonlinear Method

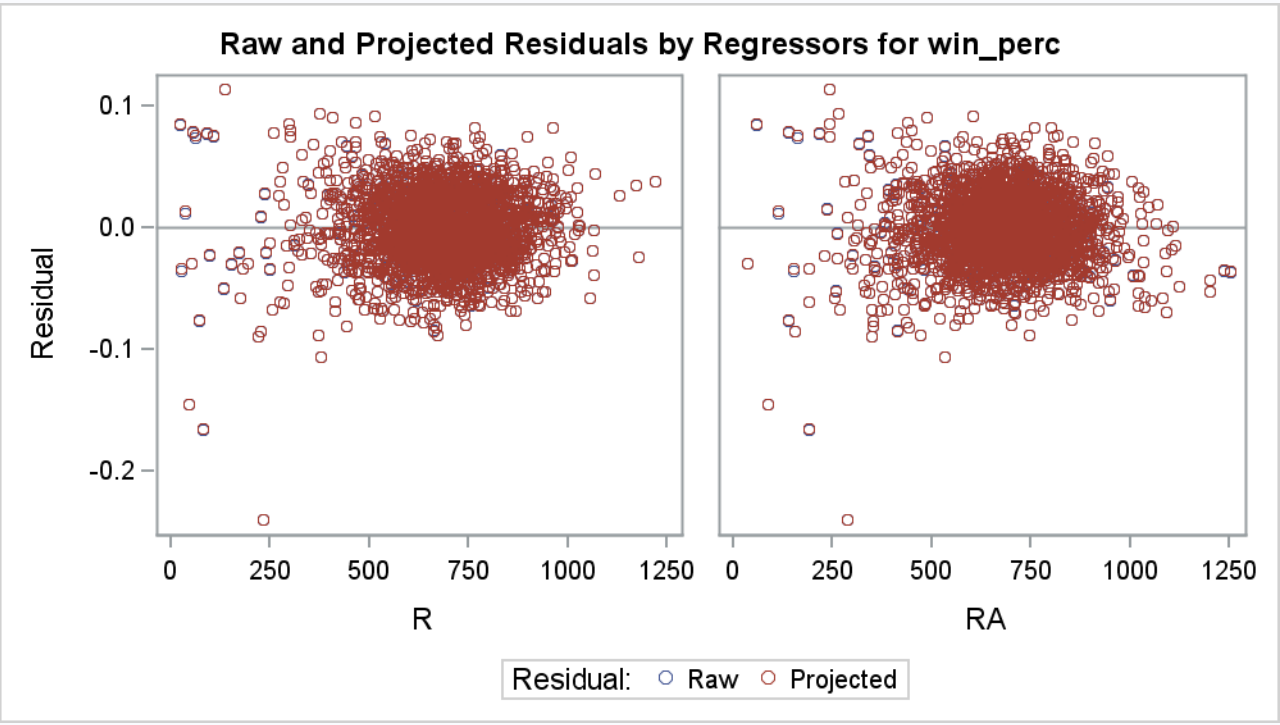
The NLIN Procedure
Dependent Variable win_perc
Method: Gauss-Newton



Estimating Pythagorean Exponent with Nonlinear Method

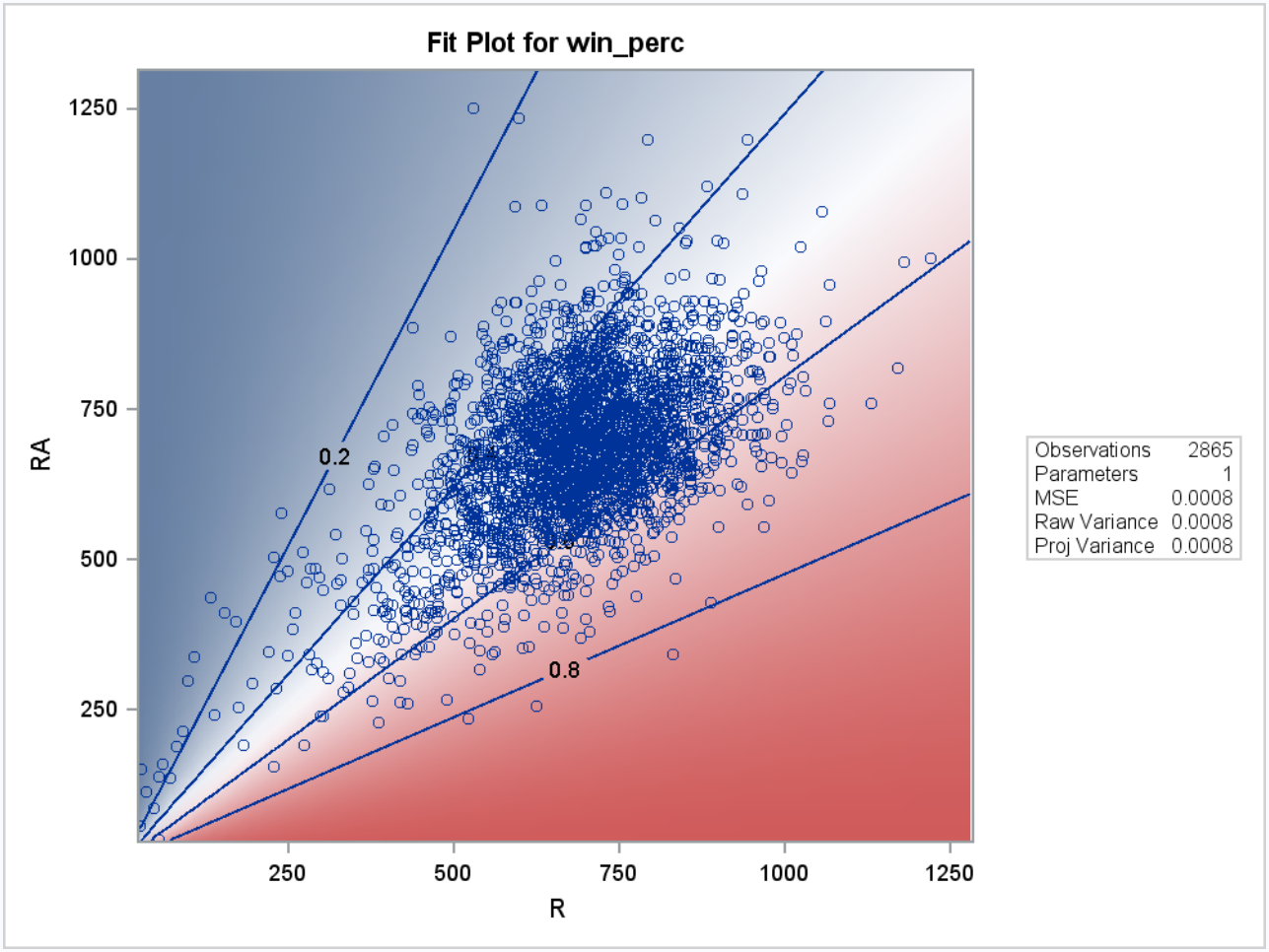
The NLIN Procedure
Dependent Variable win_perc
Method: Gauss-Newton





Estimating Pythagorean Exponent with Nonlinear Method

The NLIN Procedure
Dependent Variable win_perc
Method: Gauss-Newton



Estimation Summary	
Method	Gauss-Newton
Iterations	2
R	6.073E-7
PPC(x)	2.077E-7
RPC(x)	0.000588
Object	2.955E-6
Objective	2.269891
Observations Read	2865
Observations Used	2865
Observations Missing	0

Note: An intercept was not specified for this model.

Source	DF	Sum of Squares	Mean Square	F Value	Approx Pr > F
Model	1	734.9	734.9	927306	<.0001
Error	2864	2.2699	0.000793		
Uncorrected Total	2865	737.2			

Estimating Pythagorean Exponent with Nonlinear Method

The NLIN Procedure

Parameter	Estimate	Approx Std Error	Approximate 95% Confidence Limits		Skewness
x	1.8681	0.0119	1.8447	1.8915	0.00431

Approximate Correlation Matrix	
	x
x	1.0000000

The NLIN Procedure
Dependent Variable win_perc
Method: Gauss-Newton

Iterative Phase			
Iter	x	y	Sum of Squares
0	2.0000	2.0000	2.3651
1	1.8670	1.8674	2.2690
2	1.8681	1.8684	2.2690

NOTE: Convergence criterion met.

Estimation Summary	
Method	Gauss-Newton
Iterations	2
R	3.348E-6
PPC(y)	4.427E-7
RPC(y)	0.000567
Object	2.881E-6
Objective	2.26896
Observations Read	2865
Observations Used	2865
Observations Missing	0

Note: An intercept was not specified for this model.

Source	DF	Sum of Squares	Mean Square	F Value	Approx Pr > F
Model	2	734.9	367.5	463682	<.0001
Error	2863	2.2690	0.000793		
Uncorrected Total	2865	737.2			

Estimating Different Exponents Within the Formula

The NLIN Procedure

Parameter	Estimate	Approx Std Error	Approximate 95% Confidence Limits		Skewness
x	1.8681	0.0119	1.8447	1.8915	0.00430
y	1.8684	0.0119	1.8450	1.8919	0.00432

Approximate Correlation Matrix		
	x	y
x	1.0000000	0.9996106
y	0.9996106	1.0000000

The REG Procedure
Model: MODEL1
Dependent Variable: win_perc

Number of Observations Read	30
Number of Observations Used	30

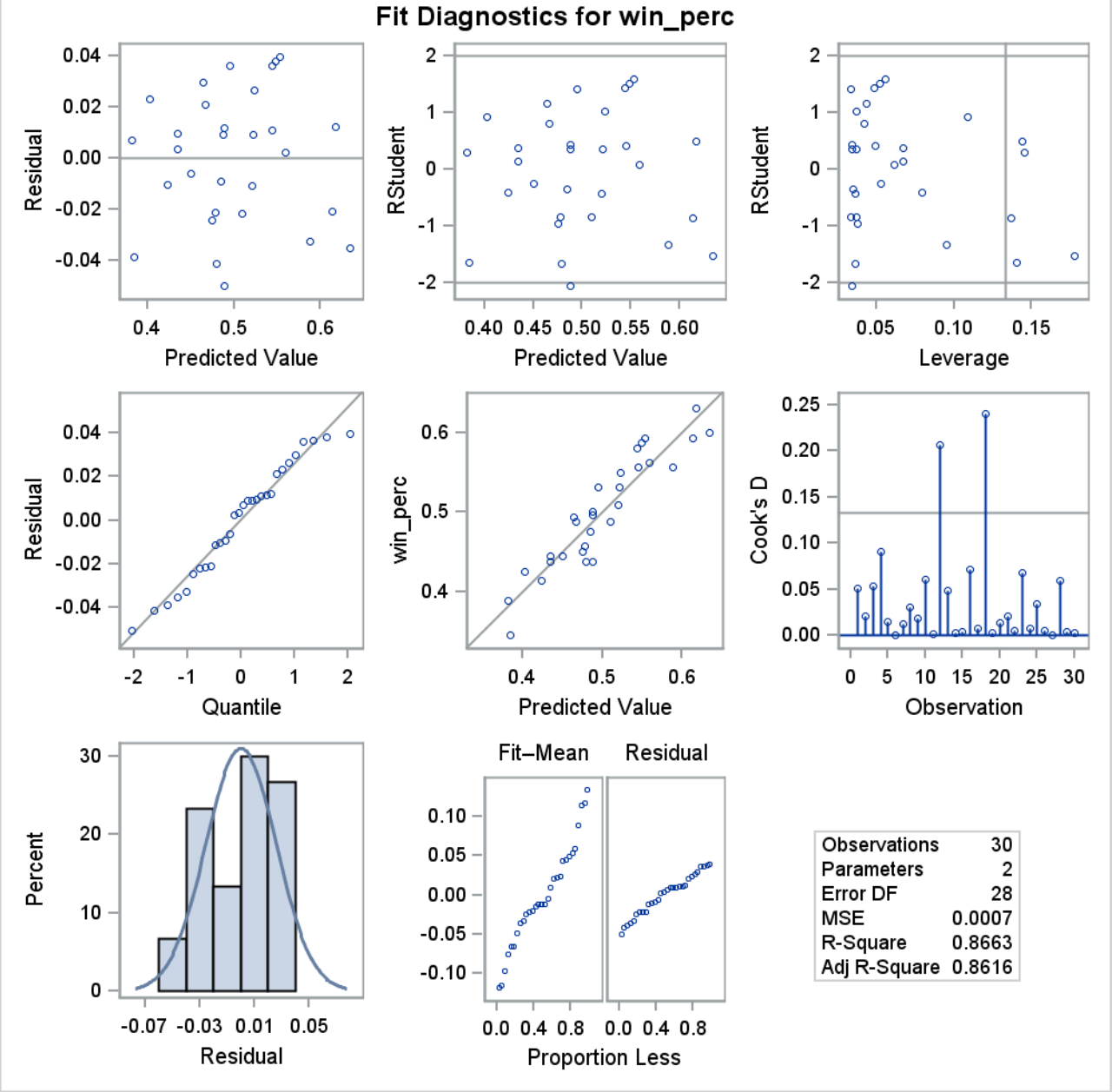
Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.12477	0.12477	181.50	<.0001
Error	28	0.01925	0.00068744		
Corrected Total	29	0.14402			

Root MSE	0.02622	R-Square	0.8663
Dependent Mean	0.50000	Adj R-Sq	0.8616
Coeff Var	5.24379		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	0.50000	0.00479	104.45	<.0001
run_diff	1	0.00063902	0.00004743	13.47	<.0001

Run differential vs Winning Percent for 2011 (like in paper)

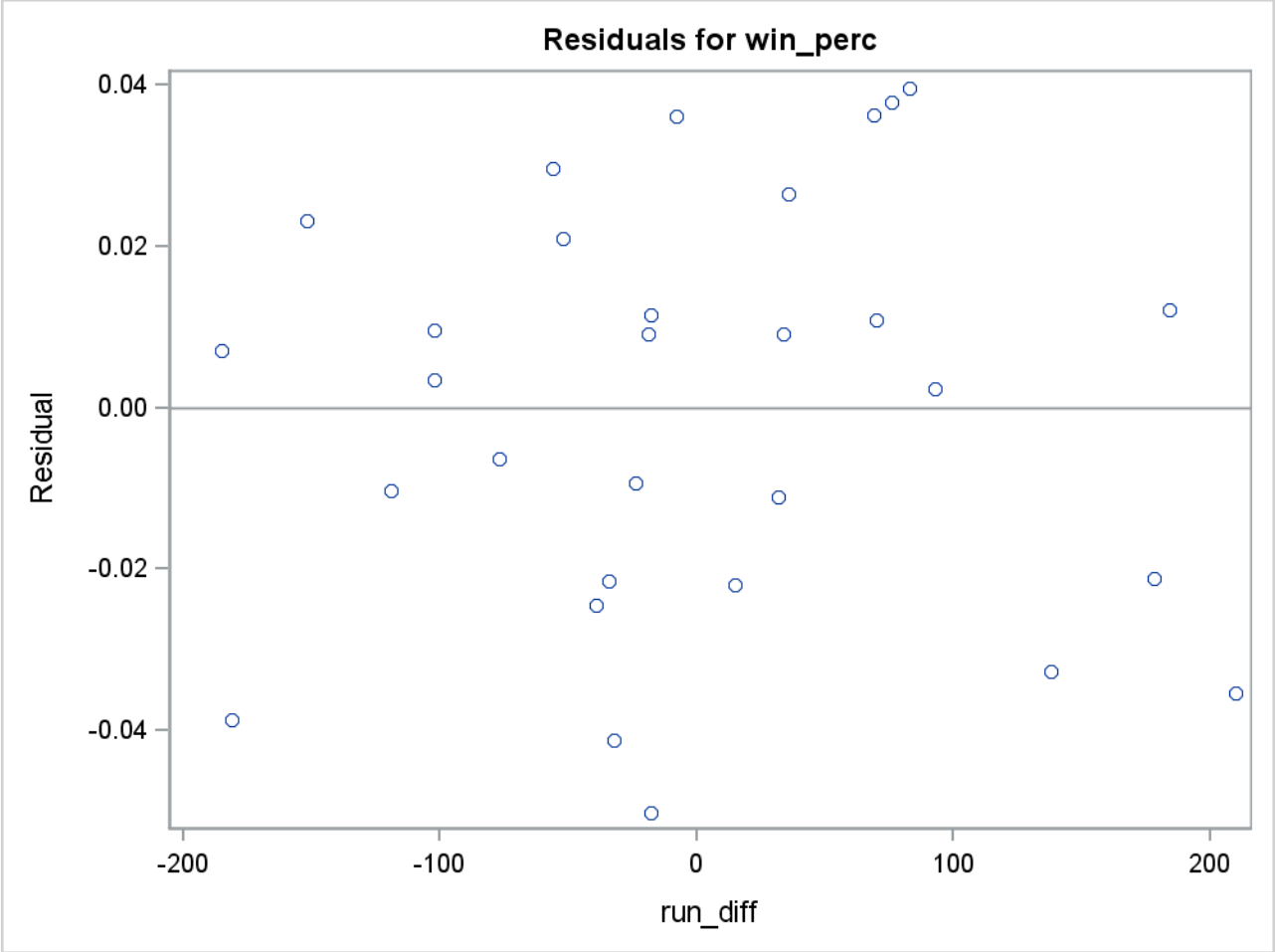
The REG Procedure
Model: MODEL1
Dependent Variable: win_perc



Observations	30
Parameters	2
Error DF	28
MSE	0.0007
R-Square	0.8663
Adj R-Square	0.8616

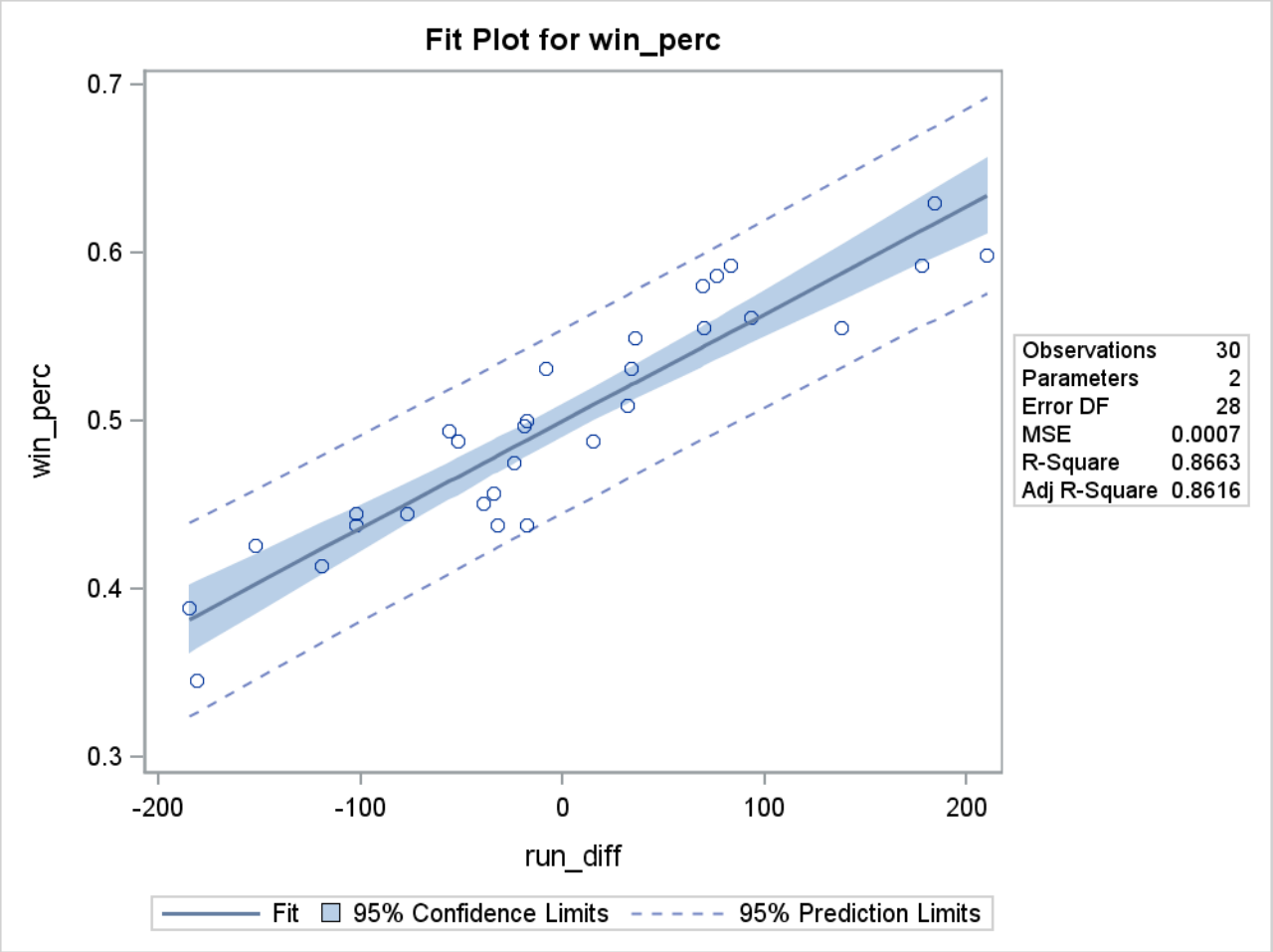
Run differential vs Winning Percent for 2011 (like in paper)

The REG Procedure
Model: MODEL1
Dependent Variable: win_perc



Run differential vs Winning Percent for 2011 (like in paper)

The REG Procedure
Model: MODEL1
Dependent Variable: win_perc



The REG Procedure
Model: MODEL1
Dependent Variable: win_perc

Number of Observations Read	1324
Number of Observations Used	1324

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	5.46850	5.46850	8325.44	<.0001
Error	1322	0.86835	0.00065684		
Corrected Total	1323	6.33685			

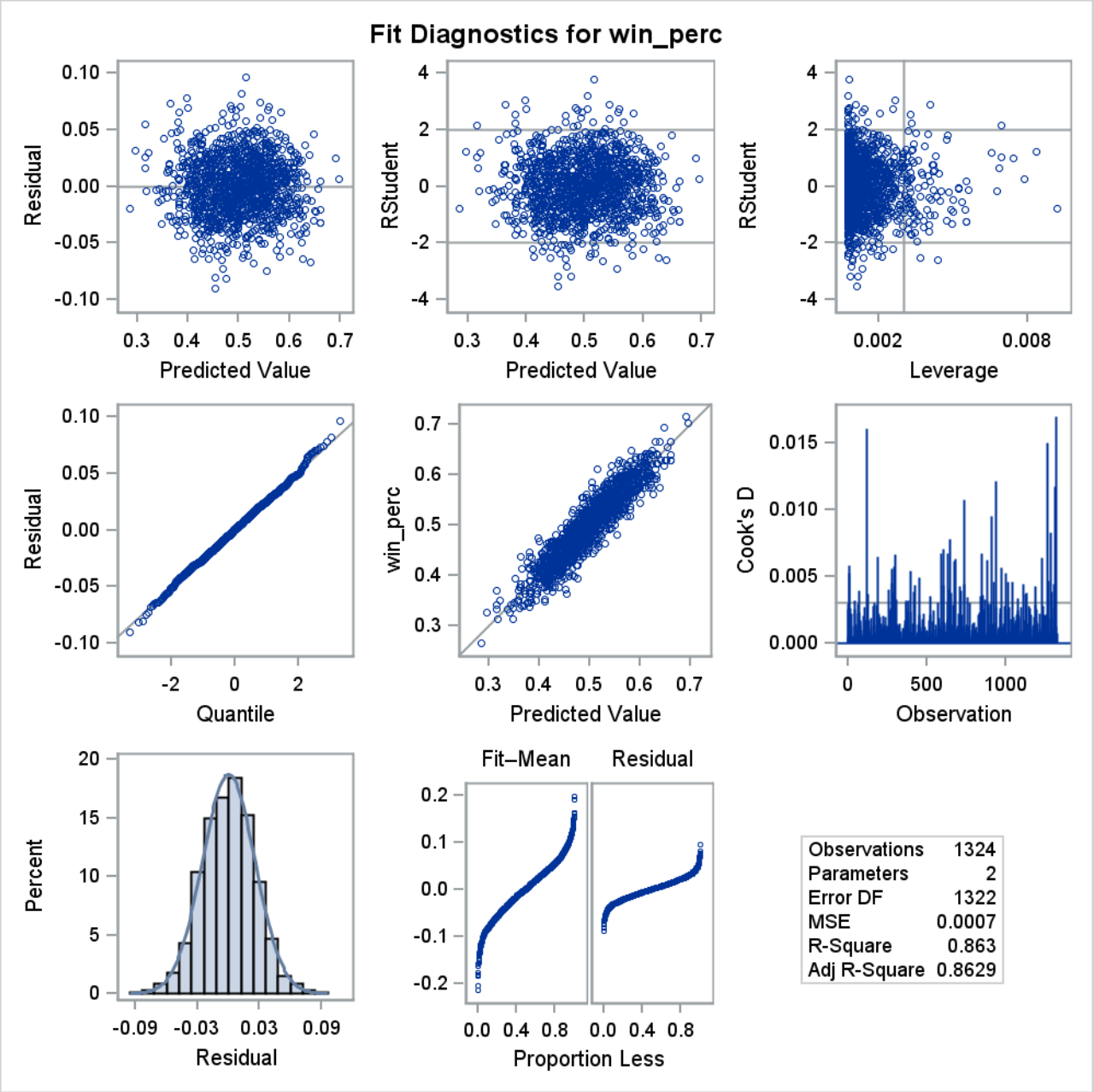
Root MSE	0.02563	R-Square	0.8630
Dependent Mean	0.49998	Adj R-Sq	0.8629
Coeff Var	5.12594		

Overall Linear Regression for Run Differential vs Winning Percent after 1970

The REG Procedure
Model: MODEL1
Dependent Variable: win_perc

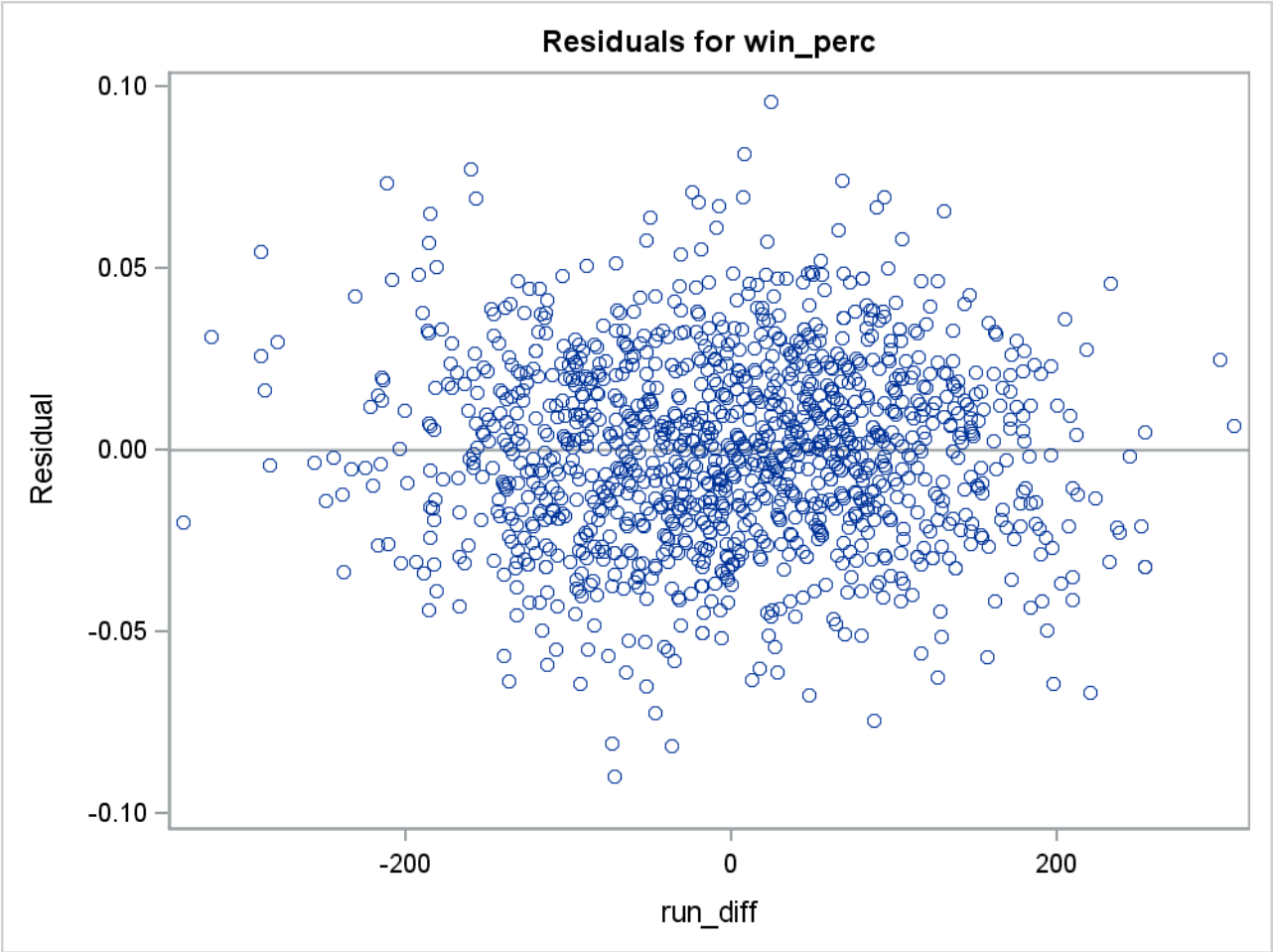
Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	0.49998	0.00070435	709.86	<.0001
run_diff	1	0.00063741	0.00000699	91.24	<.0001



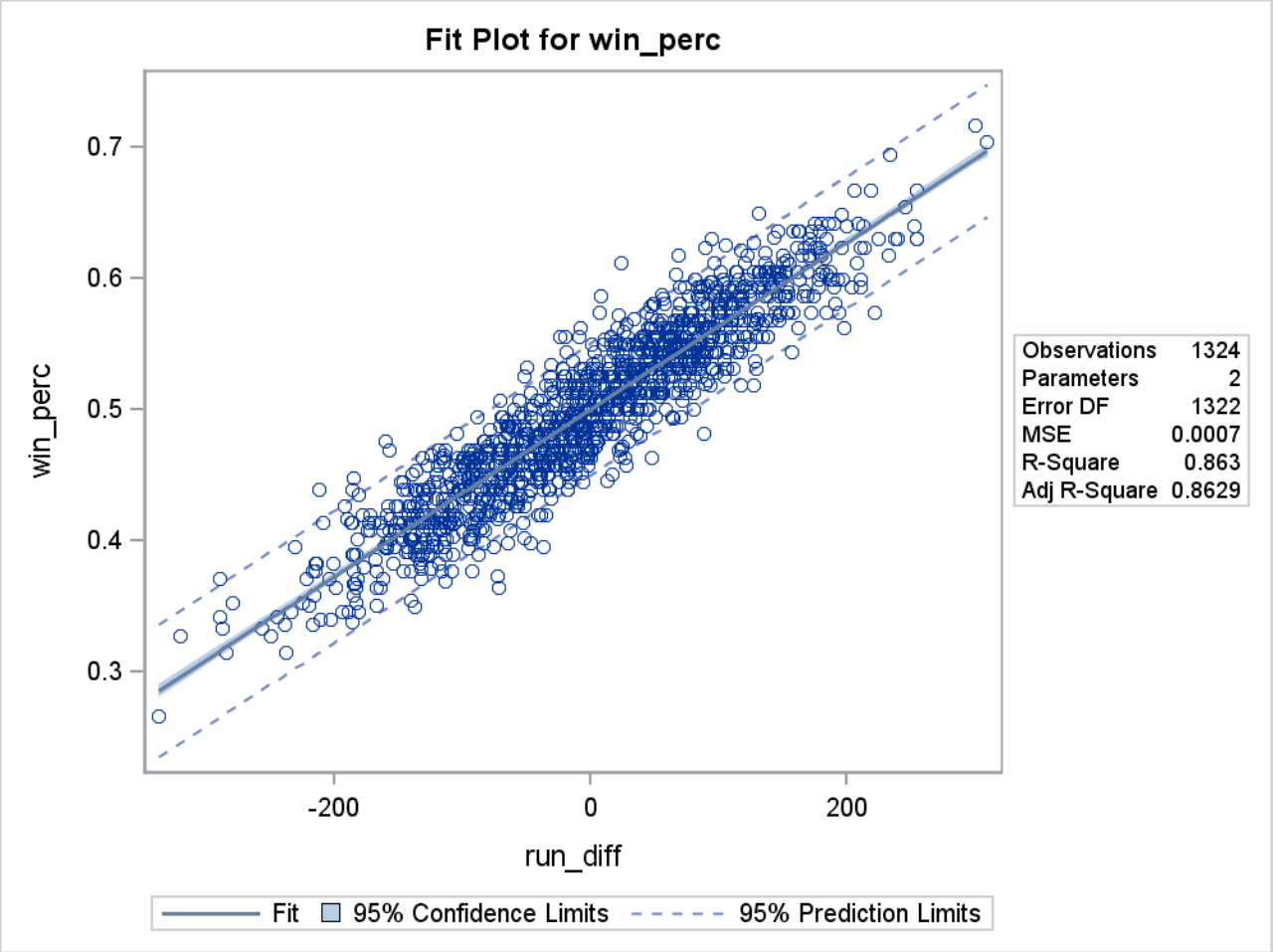
Overall Linear Regression for Run Differential vs Winning Percent after 1970

The REG Procedure
Model: MODEL1
Dependent Variable: win_perc



Overall Linear Regression for Run Differential vs Winning Percent after 1970

The REG Procedure
Model: MODEL1
Dependent Variable: win_perc



The GLMSELECT Procedure

Data Set	WORK.TRAIN
Validation Data Set	WORK.VALID
Dependent Variable	win_perc
Selection Method	Backward
Select Criterion	SBC
Stop Criterion	SBC
Choose Criterion	Cross Validation
Cross Validation Method	Random
Cross Validation Fold	5
Effect Hierarchy Enforced	None
Random Number Seed	925767001

Observation Profile for Analysis Data	
Number of Observations Read	885
Number of Observations Used	885
Number of Observations Used for Training	885

Allvars Model Winning Percent: Backward Selection

The GLMSELECT Procedure

Observation Profile for Validation Data	
Number of Observations Read	439
Number of Observations Used	439

Dimensions	
Number of Effects	29
Number of Parameters	29

The GLMSELECT Procedure

Backward Selection Summary

Step	Effect Removed	Number Effects In	SBC	ASE	Validation ASE	CV PRESS
0		29	-6802.5847	0.0004	0.0004	0.3656
1	E	28	-6809.3104	0.0004	0.0004	0.3638
2	SO	27	-6815.8975	0.0004	0.0004	0.3624
3	SF	26	-6822.2268	0.0004	0.0004	0.3617
4	DP	25	-6828.1221	0.0004	0.0004	0.3615
5	slg	24	-6833.8111	0.0004	0.0004*	0.3605
6	HR	23	-6840.1387	0.0004	0.0004	0.3602
7	HA	22	-6844.5146	0.0004	0.0004	0.3606
8	ER	21	-6848.2075	0.0004	0.0004	0.3602
9	HRA	20	-6851.8894	0.0004	0.0004	0.3596*
10	BBA	19	-6855.8960	0.0004	0.0004	0.3596
11	SOA	18	-6859.7030	0.0004	0.0004	0.3604
12	_3B	17	-6861.1339*	0.0004	0.0004	0.3616

* Optimal Value of Criterion

Selection stopped at a local minimum of the SBC criterion.

Stop Details

Candidate For	Effect	Candidate SBC		Compare SBC
Removal	ERA	-6856.6681	>	-6861.1339

The GLMSELECT Procedure
Selected Model

The selected model, based on Cross Validation, is the model at Step 9.

Effects:	Intercept R AB _1B _2B _3B BB SB CS RA ERA CG SHO SV IPouts BBA SOA FP obp HBP

Allvars Model Winning Percent: Backward Selection

The GLMSELECT Procedure
Selected Model

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value
Model	19	4.03640	0.21244	557.60
Error	865	0.32956	0.00038099	
Corrected Total	884	4.36595		

Root MSE	0.01952
Dependent Mean	0.49869
R-Square	0.9245
Adj R-Sq	0.9229
AIC	-6060.60118
AICC	-6059.53049
SBC	-6851.88942
ASE (Train)	0.00037238
ASE (Validate)	0.00040668
CV PRESS	0.35959

Parameter Estimates

Parameter	DF	Estimate	Standard Error	t Value
Intercept	1	-1.514531	0.304428	-4.98
R	1	0.000470	0.000026403	17.78
AB	1	-0.000214	0.000023873	-8.98
_1B	1	-0.000124	0.000032251	-3.85
_2B	1	-0.000143	0.000041287	-3.47
_3B	1	-0.000151	0.000079747	-1.89
BB	1	-0.000299	0.000031760	-9.41
SB	1	0.000097492	0.000024261	4.02
CS	1	-0.000447	0.000069392	-6.44
RA	1	-0.000185	0.000059809	-3.10
ERA	1	-0.031870	0.009514	-3.35
CG	1	0.000848	0.000089166	9.51
SHO	1	0.000798	0.000227	3.52
SV	1	0.002019	0.000115	17.53
IPouts	1	0.000284	0.000028779	9.88
BBA	1	-0.000021297	0.000012912	-1.65
SOA	1	0.000012066	0.000006178	1.95
FP	1	1.299498	0.324295	4.01
obp	1	2.641651	0.291073	9.08
HBP	1	-0.000333	0.000061204	-5.44

Allvars Model Winning Percent: Forward Selection

The GLMSELECT Procedure

Data Set	WORK.TRAIN
Validation Data Set	WORK.VALID
Dependent Variable	win_perc
Selection Method	Forward
Select Criterion	SBC
Stop Criterion	SBC
Choose Criterion	Cross Validation
Cross Validation Method	Random
Cross Validation Fold	5
Effect Hierarchy Enforced	None
Random Number Seed	925860000

Observation Profile for Analysis Data	
Number of Observations Read	885
Number of Observations Used	885
Number of Observations Used for Training	885

Observation Profile for Validation Data	
Number of Observations Read	439
Number of Observations Used	439

Dimensions	
Number of Effects	29
Number of Parameters	29

Allvars Model Winning Percent: Forward Selection

The GLMSELECT Procedure

Forward Selection Summary

Step	Effect Entered	Number Effects In	SBC	ASE	Validation ASE	CV PRESS
0	Intercept	1	-4694.1138	0.0049	0.0045	4.3794
1	obp	2	-4993.5492	0.0035	0.0037	3.1204
2	ERA	3	-5848.8860	0.0013	0.0014	1.1731
3	slg	4	-6257.3032	0.0008	0.0009	0.7381
4	SV	5	-6419.6172	0.0007	0.0007	0.6115
5	_2B	6	-6465.0205	0.0006	0.0007	0.5760
6	CG	7	-6491.6314	0.0006	0.0007	0.5582
7	E	8	-6571.8770	0.0006	0.0006	0.5092
8	R	9	-6596.3344	0.0005	0.0006	0.4938
9	AB	10	-6760.6269	0.0004	0.0005	0.4088
10	IPouts	11	-6778.1677	0.0004	0.0005	0.3977
11	BB	12	-6810.5500	0.0004	0.0005	0.3815
12	CS	13	-6832.4168	0.0004	0.0004	0.3701
13	SB	14	-6842.5920	0.0004	0.0004	0.3642
14	SHO	15	-6849.4157	0.0004	0.0004	0.3597
15	HBP	16	-6856.9481	0.0004	0.0004	0.3553
16	RA	17	-6860.8607*	0.0004	0.0004*	0.3525*

* Optimal Value of Criterion

Selection stopped at a local minimum of the SBC criterion.

Stop Details

Candidate For	Effect	Candidate SBC		Compare SBC
Entry	DP	-6857.5135	>	-6860.8607

The GLMSELECT Procedure
Selected Model

The selected model, based on Cross Validation, is the model at Step 16.

Effects:	Intercept R AB _2B BB SB CS RA ERA CG SHO SV IPouts E obp slg HBP

Allvars Model Winning Percent: Forward Selection

The GLMSELECT Procedure
Selected Model

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value
Model	16	4.03213	0.25201	655.27
Error	868	0.33382	0.00038459	
Corrected Total	884	4.36595		

Root MSE	0.01961
Dependent Mean	0.49869
R-Square	0.9235
Adj R-Sq	0.9221
AIC	-6055.21572
AICC	-6054.42588
SBC	-6860.86073
ASE (Train)	0.00037720
ASE (Validate)	0.00041804
CV PRESS	0.35248

Parameter Estimates

Parameter	DF	Estimate	Standard Error	t Value
Intercept	1	-0.021258	0.068056	-0.31
R	1	0.000454	0.000029495	15.41
AB	1	-0.000258	0.000020792	-12.40
_2B	1	-0.000063710	0.000032228	-1.98
BB	1	-0.000187	0.000023031	-8.14
SB	1	0.000091802	0.000024155	3.80
CS	1	-0.000476	0.000069032	-6.89
RA	1	-0.000196	0.000060418	-3.25
ERA	1	-0.032293	0.009651	-3.35
CG	1	0.000763	0.000081134	9.41
SHO	1	0.000845	0.000226	3.73
SV	1	0.001983	0.000114	17.37
IPouts	1	0.000301	0.000028302	10.64
E	1	-0.000218	0.000051454	-4.23
obp	1	1.640779	0.157566	10.41
slg	1	0.278775	0.066827	4.17
HBP	1	-0.000198	0.000055210	-3.58

The GLMSELECT Procedure

Data Set	WORK.TRAIN
Validation Data Set	WORK.VALID
Dependent Variable	win_perc
Selection Method	Stepwise

Allvars Model Winning Percent: Stepwise Selection

The GLMSELECT Procedure

Select Criterion	SBC
Stop Criterion	SBC
Choose Criterion	Cross Validation
Cross Validation Method	Random
Cross Validation Fold	5
Effect Hierarchy Enforced	None
Random Number Seed	925954001

Observation Profile for Analysis Data	
Number of Observations Read	885
Number of Observations Used	885
Number of Observations Used for Training	885

Observation Profile for Validation Data	
Number of Observations Read	439
Number of Observations Used	439

Dimensions	
Number of Effects	29
Number of Parameters	29

The GLMSELECT Procedure

Stepwise Selection Summary

Step	Effect Entered	Effect Removed	Number Effects In	SBC	ASE	Validation ASE	CV PRESS
0	Intercept		1	-4694.1138	0.0049	0.0045	4.3757
1	obp		2	-4993.5492	0.0035	0.0037	3.0992
2	ERA		3	-5848.8860	0.0013	0.0014	1.1721
3	slg		4	-6257.3032	0.0008	0.0009	0.7373
4	SV		5	-6419.6172	0.0007	0.0007	0.6095
5	_2B		6	-6465.0205	0.0006	0.0007	0.5764
6	CG		7	-6491.6314	0.0006	0.0007	0.5561
7	E		8	-6571.8770	0.0006	0.0006	0.5079
8	R		9	-6596.3344	0.0005	0.0006	0.4905
9	AB		10	-6760.6269	0.0004	0.0005	0.4040
10		_2B	9	-6765.4636	0.0004	0.0005	0.4042
11	IPouts		10	-6783.9959	0.0004	0.0005	0.3934
12	BB		11	-6815.2043	0.0004	0.0005	0.3795
13	CS		12	-6835.9900	0.0004	0.0004	0.3688
14	SB		13	-6844.7079	0.0004	0.0004	0.3629
15	HBP		14	-6852.3727	0.0004	0.0004	0.3580
16	SHO		15	-6859.9182	0.0004	0.0004	0.3523
17	RA		16	-6863.6708*	0.0004	0.0004*	0.3490*

* Optimal Value of Criterion

Allvars Model Winning Percent: Stepwise Selection

The GLMSELECT Procedure

Selection stopped at a local minimum of the SBC criterion.

Stop Details				
Candidate For	Effect	Candidate SBC		Compare SBC
Entry	HR	-6863.1285	>	-6863.6708
Removal	RA	-6859.9182	>	-6863.6708

The GLMSELECT Procedure
Selected Model

The selected model, based on Cross Validation, is the model at Step 17.

Effects:	Intercept R AB BB SB CS RA ERA CG SHO SV IPouts E obp slg HBP

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Value
Model	15	4.03063	0.26871	696.36
Error	869	0.33533	0.00038588	
Corrected Total	884	4.36595		

Root MSE	0.01964
Dependent Mean	0.49869
R-Square	0.9232
Adj R-Sq	0.9219
AIC	-6053.24018
AICC	-6052.53430
SBC	-6863.67078
ASE (Train)	0.00037890
ASE (Validate)	0.00041800
CV PRESS	0.34902

Allvars Model Winning Percent: Stepwise Selection

The GLMSELECT Procedure
Selected Model

Parameter Estimates

Parameter	DF	Estimate	Standard Error	t Value
Intercept	1	0.003206	0.067034	0.05
R	1	0.000458	0.000029484	15.54
AB	1	-0.000263	0.000020627	-12.77
BB	1	-0.000185	0.000023048	-8.05
SB	1	0.000087780	0.000024109	3.64
CS	1	-0.000468	0.000069027	-6.77
RA	1	-0.000195	0.000060517	-3.23
ERA	1	-0.032278	0.009667	-3.34
CG	1	0.000816	0.000076782	10.63
SHO	1	0.000815	0.000226	3.60
SV	1	0.002010	0.000114	17.70
IPouts	1	0.000303	0.000028331	10.70
E	1	-0.000211	0.000051428	-4.10
obp	1	1.602490	0.156632	10.23
slg	1	0.247110	0.064987	3.80
HBP	1	-0.000212	0.000054827	-3.87

The GLMSELECT Procedure

Data Set	WORK.TRAIN
Validation Data Set	WORK.VALID
Dependent Variable	win_perc
Selection Method	None
Random Number Seed	926032000

Observation Profile for Analysis Data	
Number of Observations Read	885
Number of Observations Used	885
Number of Observations Used for Training	885

Observation Profile for Validation Data	
Number of Observations Read	439
Number of Observations Used	439

Dimensions	
Number of Effects	29
Number of Parameters	29

Least Squares Regression

The GLMSELECT Procedure

Least Squares Summary

Step	Effect Entered	Number Effects In	Adjusted R-Square	AIC	BIC	SBC	ASE	Validation ASE	CV PRESS	Pr > F
0	Intercept	1	0.0000	-3811.8994	-4700.4492	-4694.1138	0.0049	0.0045	4.3730*	1.0000
1	R	2	0.2260	-4037.6669	-4927.8898	-4915.0957	0.0038	0.0038	.	<.0001
2	AB	3	0.2939	-4117.8555	-5009.7863	-4990.4988	0.0035	0.0034	.	<.0001
3	_1B	4	0.3118	-4139.5775	-5033.2554	-5007.4351	0.0034	0.0035	.	<.0001
4	_2B	5	0.3177	-4146.2607	-5041.6991	-5009.3328	0.0034	0.0034	.	0.0033
5	_3B	6	0.3173	-4144.7074	-5041.9181	-5002.9939	0.0033	0.0034	.	0.5055
6	HR	7	0.3210	-4148.5200	-5047.4906	-5002.0209	0.0033	0.0033	.	0.0164
7	BB	8	0.3273	-4155.7452	-5056.4641	-5004.4605	0.0033	0.0034	.	0.0025
8	SO	9	0.3280	-4155.6877	-5058.1704	-4999.6174	0.0033	0.0033	.	0.1657
9	SB	10	0.3406	-4171.4636	-5075.6597	-5010.6077	0.0032	0.0034	.	<.0001
10	CS	11	0.3577	-4193.7560	-5099.6324	-5028.1145	0.0031	0.0033	.	<.0001
11	RA	12	0.8675	-5589.7874	-6485.0001	-6419.3603	0.0006	0.0006	.	<.0001
12	ER	13	0.8674	-5588.1767	-6484.2051	-6412.9641	0.0006	0.0006	.	0.5358
13	ERA	14	0.8740	-5632.6002	-6528.5128	-6452.6019	0.0006	0.0006	.	<.0001
14	CG	15	0.8741	-5631.6800	-6528.3277	-6446.8962	0.0006	0.0006	.	0.3030
15	SHO	16	0.8755	-5641.2105	-6538.3426	-6451.6411	0.0006	0.0006	.	0.0008
16	SV	17	0.9113	-5940.2863	-6829.2392	-6745.9313	0.0004	0.0005	.	<.0001
17	IPouts	18	0.9191	-6020.0596	-6906.1360	-6820.9190*	0.0004	0.0004	.	<.0001
18	HA	19	0.9190	-6018.3224	-6904.4475	-6814.3962	0.0004	0.0004	.	0.6122
19	HRA	20	0.9190	-6016.9712	-6903.1270	-6808.2595	0.0004	0.0004	.	0.4260
20	BBA	21	0.9196	-6022.7318	-6908.5851	-6809.2344	0.0004	0.0004	.	0.0059
21	SOA	22	0.9198	-6024.0238	-6909.7568	-6805.7409	0.0004	0.0004	.	0.0733
22	E	23	0.9206	-6032.2095	-6917.4517	-6809.1410	0.0004	0.0004	.	0.0016
23	DP	24	0.9206	-6030.9128	-6916.1276	-6803.0587	0.0004	0.0004	.	0.4083
24	FP	25	0.9208	-6032.5156	-6917.5363	-6799.8759	0.0004	0.0004	.	0.0614
25	obp	26	0.9216	-6040.0732	-6924.5305	-6802.6479	0.0004	0.0004	.	0.0023
26	slg	27	0.9227	-6051.6700	-6935.2628	-6809.4591	0.0004	0.0004	.	0.0003
27	HBP	28	0.9231*	-6055.9949*	-6939.1242*	-6808.9984	0.0004	0.0004	.	0.0134
28	SF	29	0.9231	-6054.3668	-6937.4041	-6802.5847	0.0004	0.0004*	.	0.5488

* Optimal Value of Criterion

Least Squares Regression

The GLMSELECT Procedure
Least Squares Model (No Selection)

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	28	4.04075	0.14431	379.86	<.0001
Error	856	0.32520	0.00037991		
Corrected Total	884	4.36595			

Root MSE	0.01949
Dependent Mean	0.49869
R-Square	0.9255
Adj R-Sq	0.9231
AIC	-6054.36676
AICC	-6052.18878
BIC	-6937.40411
C(p)	29.00000
SBC	-6802.58472
ASE (Train)	0.00036746
ASE (Validate)	0.00040310
CV PRESS	.

Parameter Estimates					
Parameter	DF	Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-1.749789	1.649249	-1.06	0.2890
R	1	0.000480	0.000031845	15.09	<.0001
AB	1	-0.000201	0.000042433	-4.74	<.0001
_1B	1	-0.000150	0.000111	-1.35	0.1770
_2B	1	-0.000304	0.000133	-2.29	0.0222
_3B	1	-0.000449	0.000224	-2.00	0.0453
HR	1	-0.000407	0.000312	-1.30	0.1929
BB	1	-0.000239	0.000117	-2.05	0.0409
SO	1	0.000003398	0.000007732	0.44	0.6604
SB	1	0.000092731	0.000024928	3.72	0.0002
CS	1	-0.000458	0.000069732	-6.57	<.0001
RA	1	-0.000224	0.000071724	-3.13	0.0018
ER	1	0.000311	0.000140	2.22	0.0267
ERA	1	-0.065179	0.019380	-3.36	0.0008
CG	1	0.000861	0.000093224	9.23	<.0001
SHO	1	0.000742	0.000230	3.23	0.0013
SV	1	0.002031	0.000116	17.47	<.0001
IPouts	1	0.000263	0.000035353	7.45	<.0001
HA	1	-0.000027869	0.000021193	-1.32	0.1889
HRA	1	-0.000098413	0.000044950	-2.19	0.0288
BBA	1	-0.000037834	0.000017021	-2.22	0.0265

Least Squares Regression

The GLMSELECT Procedure
Least Squares Model (No Selection)

Parameter Estimates					
Parameter	DF	Estimate	Standard Error	t Value	Pr > |t|
SOA	1	0.000008950	0.000007801	1.15	0.2515
E	1	0.000065363	0.000271	0.24	0.8098
DP	1	-0.000046460	0.000045909	-1.01	0.3118
FP	1	1.597554	1.700963	0.94	0.3479
obp	1	2.046551	1.046988	1.95	0.0509
slg	1	0.689840	0.589739	1.17	0.2424
HBP	1	-0.000255	0.000129	-1.98	0.0480
SF	1	-0.000066788	0.000111	-0.60	0.5488

The GLMSELECT Procedure

Data Set	WORK.TRAIN
Validation Data Set	WORK.VALID
Dependent Variable	win_perc
Selection Method	ELASTICNET
Stop Criterion	SBC
Choose Criterion	Cross Validation
Cross Validation Method	Random
Cross Validation Fold	5
Effect Hierarchy Enforced	None
Random Number Seed	926142000

Observation Profile for Analysis Data	
Number of Observations Read	885
Number of Observations Used	885
Number of Observations Used for Training	885

Observation Profile for Validation Data	
Number of Observations Read	439
Number of Observations Used	439

Dimensions	
Number of Effects	29
Number of Parameters	29

Elastic Net

The GLMSELECT Procedure

Elastic Net Selection Summary

Step	Effect Entered	Effect Removed	Number Effects In	Adjusted R-Square	AIC	BIC	SBC	ASE	Validation ASE	CV PRESS	Pr > F
0	Intercept		1	0.0000	-3811.8994	-4699.9163	-4694.1138	0.0049	0.0045	4.3934	1.0000
1	obp		2	0.0313	-3839.0120	-4728.6385	-4716.4408	0.0048	0.0044	3.1047	<.0001
2	ERA		3	0.0445	-3850.1699	-4741.4120	-4722.8131	0.0047	0.0043	1.1788	0.0003
3	SV		4	0.1201	-3922.1363	-4814.8348	-4789.9940	0.0043	0.0040	0.8654	<.0001
4	RA		5	0.2825	-4101.7203	-4995.4295	-4964.7923	0.0035	0.0033	0.8239	<.0001
5	R		6	0.4968	-4414.7440	-5308.4233	-5273.0304	0.0025	0.0023	0.4863	<.0001
6	slg		7	0.5075	-4432.6924	-5327.5449	-5286.1933	0.0024	0.0023	0.4808	<.0001
7	ER		8	0.6557	-4748.5384	-5641.8725	-5597.2537	0.0017	0.0016	0.4808	<.0001
8	SHO		9	0.7104	-4900.6714	-5793.0719	-5744.6011	0.0014	0.0014	0.4713	<.0001
9	BBA		10	0.7968	-5213.2911	-6101.0308	-6052.4353	0.0010	0.0010	0.4709	<.0001
10	HR		11	0.8160	-5299.9239	-6185.8969	-6134.2824	0.0009	0.0009	0.4681	<.0001
11	SF		12	0.8197	-5317.0258	-6202.5836	-6146.5987	0.0009	0.0009	0.4672	<.0001
12	BB		13	0.8209	-5322.2737	-6207.6685	-6147.0611	0.0009	0.0009	0.4657	0.0075
13	HA		14	0.8368	-5403.0658	-6285.7995	-6223.0676	0.0008	0.0008	0.4664	<.0001
14	CG		15	0.8393	-5415.9328	-6297.9607	-6231.1490	0.0008	0.0008	0.4137	0.0001
15	FP		16	0.8582	-5525.6958	-6402.8901	-6336.1264	0.0007	0.0007	0.4129	<.0001
16	SB		17	0.8717	-5613.1493	-6485.5025	-6418.7943	0.0006	0.0006	0.4092	<.0001
17	HRA		18	0.8726	-5618.3700	-6489.5725	-6419.2294	0.0006	0.0006	0.4094	0.0078
18	E		19	0.8737	-5625.4723	-6495.3675	-6421.5462	0.0006	0.0006	0.4098	0.0028
19	SO		20	0.8788	-5660.9429	-6527.6993	-6452.2312	0.0006	0.0006	0.4127	<.0001
20	_1B		21	0.8888*	-5735.8734*	-6596.3846*	-6522.3760*	0.0005	0.0006*	0.4019*	<.0001

* Optimal Value of Criterion

Selection stopped at a local minimum of the SBC criterion.

Stop Details

Candidate For	Effect	Candidate SBC		Compare SBC
Entry	SOA	-6520.7249	>	-6522.3760

The GLMSELECT Procedure
Selected Model

The selected model, based on Cross Validation, is the model at Step 20.

Effects:	Intercept R _1B HR BB SO SB RA ER ERA CG SHO SV HA HRA BBA E FP obp slg SF

Elastic Net

The GLMSELECT Procedure
Selected Model

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Value
Model	20	3.89138	0.19457	354.23
Error	864	0.47457	0.00054927	
Corrected Total	884	4.36595		

Root MSE	0.02344
Dependent Mean	0.49869
R-Square	0.8913
Adj R-Sq	0.8888
AIC	-5735.87338
AICC	-5734.69936
BIC	-6596.38456
C(p)	-289.95191
SBC	-6522.37603
ASE (Train)	0.00053624
ASE (Validate)	0.00055045
CV PRESS	0.40187

Parameter Estimates		
Parameter	DF	Estimate
Intercept	1	-0.670306
R	1	0.000164
_1B	1	0.000019137
HR	1	0.000184
BB	1	0.000051491
SO	1	-0.000008837
SB	1	0.000064818
RA	1	-0.000120
ER	1	-0.000109
ERA	1	-0.021245
CG	1	0.000427
SHO	1	0.001193
SV	1	0.001643
HA	1	-0.000051006
HRA	1	-0.000072281
BBA	1	-0.000067565
E	1	-0.000076982
FP	1	0.844449
obp	1	0.890647
slg	1	0.341494
SF	1	0.000310

Lasso

The GLMSELECT Procedure

Data Set	WORK.TRAIN
Validation Data Set	WORK.VALID
Dependent Variable	win_perc
Selection Method	LASSO
Stop Criterion	SBC
Choose Criterion	Cross Validation
Cross Validation Method	Random
Cross Validation Fold	5
Effect Hierarchy Enforced	None
Random Number Seed	926329000

Observation Profile for Analysis Data	
Number of Observations Read	885
Number of Observations Used	885
Number of Observations Used for Training	885

Observation Profile for Validation Data	
Number of Observations Read	439
Number of Observations Used	439

Dimensions	
Number of Effects	29
Number of Parameters	29

Lasso

The GLMSELECT Procedure

LASSO Selection Summary

Step	Effect Entered	Effect Removed	Number Effects In	Adjusted R-Square	AIC	BIC	SBC	ASE	Validation ASE	CV PRESS	Pr > F
0	Intercept		1	0.0000	-3811.8994	-4700.4492	-4694.1138	0.0049	0.0045	4.3718	1.0000
1	obp		2	0.0366	-3843.8857	-4734.2583	-4721.3145	0.0047	0.0044	3.0950	<.0001
2	ERA		3	0.0706	-3874.6741	-4766.8574	-4747.3173	0.0046	0.0042	1.1793	<.0001
3	SV		4	0.3326	-4166.7972	-5060.4346	-5034.6549	0.0033	0.0031	0.8575	<.0001
4	RA		5	0.4940	-4410.8352	-5305.7415	-5273.9073	0.0025	0.0024	0.8189	<.0001
5	R		6	0.8007	-5234.4389	-6127.5229	-6092.7253	0.0010	0.0009	0.4870	<.0001
6	slg		7	0.8327	-5388.2312	-6281.3189	-6241.7321	0.0008	0.0008	0.4832	<.0001
7	SHO		8	0.8735	-5634.8938	-6526.3717	-6483.6091	0.0006	0.0006	0.4754	<.0001
8	CG		9	0.8883	-5743.5643	-6634.2300	-6587.4940	0.0005	0.0005	0.4204	<.0001
9	BBA		10	0.8922	-5774.4900	-6665.2066	-6613.6341	0.0005	0.0005	0.4210	<.0001
10	HR		11	0.9014	-5852.5759	-6742.2778	-6686.9344	0.0005	0.0005	0.4192	<.0001
11	SB		12	0.9033*	-5868.2864*	-6758.0115*	-6697.8593*	0.0005	0.0005*	0.4165*	<.0001

* Optimal Value of Criterion

Selection stopped at a local minimum of the SBC criterion.

Stop Details

Candidate For	Effect	Candidate SBC		Compare SBC
Entry	DP	-6695.1469	>	-6697.8593

The GLMSELECT Procedure
Selected Model

The selected model, based on Cross Validation, is the model at Step 11.

Effects:	Intercept R HR SB RA ERA CG SHO SV BBA obp slg

Lasso

The GLMSELECT Procedure
Selected Model

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Value
Model	11	3.94894	0.35899	751.53
Error	873	0.41702	0.00047769	
Corrected Total	884	4.36595		

Root MSE	0.02186
Dependent Mean	0.49869
R-Square	0.9045
Adj R-Sq	0.9033
AIC	-5868.28639
AICC	-5867.86848
BIC	-6758.01146
C(p)	236.67410
SBC	-6697.85934
ASE (Train)	0.00047121
ASE (Validate)	0.00047320
CV PRESS	0.41654

Parameter Estimates		
Parameter	DF	Estimate
Intercept	1	0.171946
R	1	0.000361
HR	1	0.000075248
SB	1	0.000010768
RA	1	-0.000484
ERA	1	-0.000727
CG	1	0.000562
SHO	1	0.000823
SV	1	0.001916
BBA	1	-0.000011736
obp	1	0.793595
slg	1	0.153596

The PLS Procedure

Data Set	WORK.DEVELOP_SAMPLE2
Factor Extraction Method	Principal Components Regression
Number of Response Variables	1
Number of Predictor Parameters	28
Missing Value Handling	Exclude
Maximum Number of Factors	15
Validation Method	Test Set Validation (WORK.VALID)
Validation Testing Criterion	Prob PRESS > 0.1

PCR

The PLS Procedure

Number of Random Permutations	1000
Random Permutation Seed	12345

Number of Observations Read	2865
Number of Observations Used	885

Test Set Validation for the Number of Extracted Factors

Number of Extracted Factors	Root Mean PRESS	Prob > PRESS
0	0.953962	<.0001
1	0.955636	<.0001
2	0.940363	<.0001
3	0.524577	<.0001
4	0.410427	<.0001
5	0.399583	<.0001
6	0.391556	<.0001
7	0.384864	<.0001
8	0.37002	<.0001
9	0.35181	0.0060
10	0.353019	<.0001
11	0.350752	0.0050
12	0.344678	0.0290
13	0.344109	0.0350
14	0.339352	0.1610
15	0.337295	1.0000

Minimum root mean PRESS	0.3373
Minimizing number of factors	15
Smallest number of factors with p > 0.1	14

The PLS Procedure

Percent Variation Accounted for by Principal Components

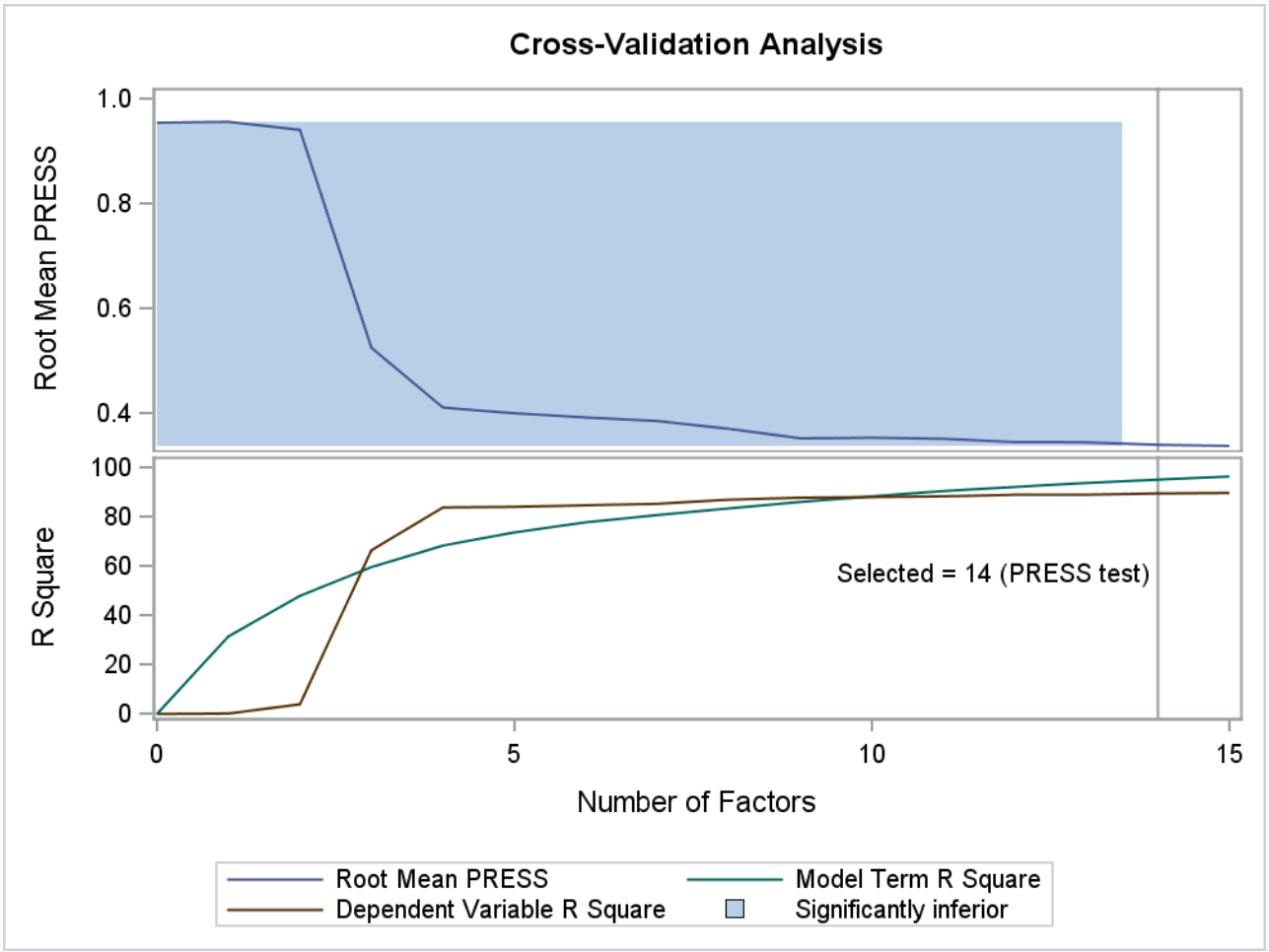
Number of Extracted Factors	Model Effects		Dependent Variables	
	Current	Total	Current	Total
1	31.4537	31.4537	0.1878	0.1878
2	16.4327	47.8865	3.7563	3.9441
3	11.6589	59.5453	62.4431	66.3873
4	8.7302	68.2755	17.3430	83.7303
5	5.3159	73.5914	0.2528	83.9831

PCR

The PLS Procedure

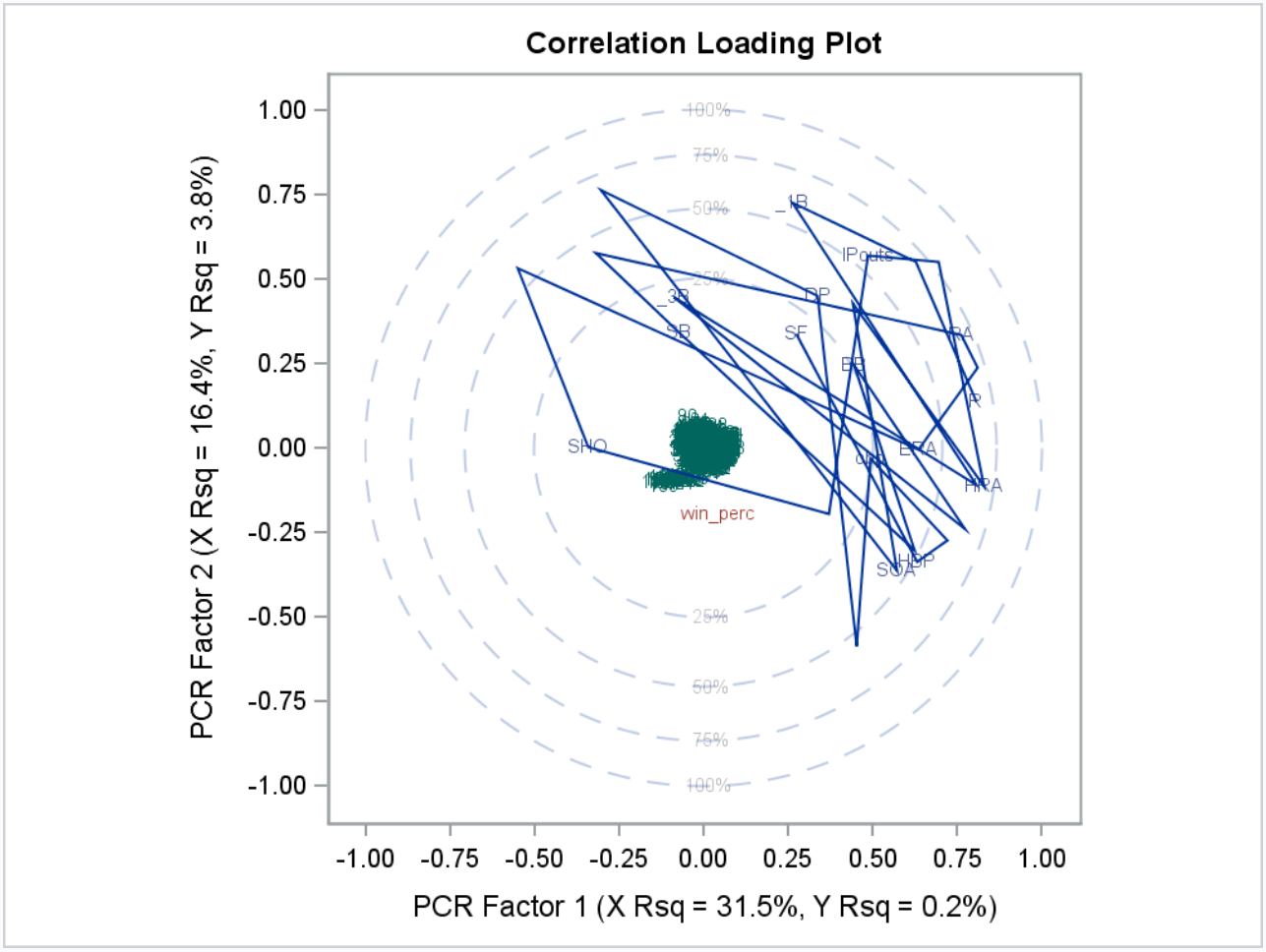
Percent Variation Accounted for by Principal Components

Number of Extracted Factors	Model Effects		Dependent Variables	
	Current	Total	Current	Total
6	4.1231	77.7145	0.6426	84.6256
7	2.9313	80.6458	0.5822	85.2078
8	2.7073	83.3532	1.6818	86.8897
9	2.6010	85.9542	0.8216	87.7112
10	2.2718	88.2260	0.2345	87.9457
11	2.1552	90.3812	0.2931	88.2388
12	1.6739	92.0551	0.6456	88.8844
13	1.5810	93.6361	0.0049	88.8893
14	1.3702	95.0063	0.5209	89.4101



PCR

The PLS Procedure



Parameter Estimates for Centered and Scaled Data	
	win_perc
Intercept	0.0000000000
R	0.1901681596
AB	0.0206909287
_1B	0.0630307486
_2B	0.0105195393
_3B	0.0154338676
HR	0.1994654543
BB	0.0741094953
SO	-.0681196464
SB	0.0309990031
CS	0.0215811923
RA	-.1832528337
ER	-.1794565826
ERA	-.1969207367
CG	0.0712101789
SHO	0.0306122558
SV	0.2046619794

PCR

The PLS Procedure

Parameter Estimates for Centered and Scaled Data	
	win_perc
lPouts	0.0173102189
HA	-.1408440596
HRA	-.0736512579
BBA	-.0771183675
SOA	0.0184500364
E	-.0333878514
DP	0.0133189151
FP	0.0436050733
obp	0.1748420582
slg	0.2070006124
HBP	-.0040281764
SF	0.0352061404

Parameter Estimates	
	win_perc
Intercept	-.6898752955
R	0.0001369134
AB	0.0000041814
_1B	0.0000511216
_2B	0.0000180431
_3B	0.0001119703
HR	0.0003445401
BB	0.0000707199
SO	-.0000255316
SB	0.0000548105
CS	0.0000909119
RA	-.0001298983
ER	-.0001323895
ERA	-.0241958867
CG	0.0003204875
SHO	0.0005208474
SV	0.0015596182
lPouts	0.0000045163
HA	-.0000809487
HRA	-.0001478242
BBA	-.0000748358
SOA	0.0000068261
E	-.0000959044
DP	0.0000465824
FP	0.8438368877
obp	0.8493613045

PCR

The PLS Procedure

Parameter Estimates	
	win_perc
slg	0.4519476041
HBP	-.0000168170
SF	0.0002753330

The PLS Procedure

Data Set	WORK.PCR_PRED
Factor Extraction Method	Partial Least Squares
PLS Algorithm	NIPALS
Number of Response Variables	1
Number of Predictor Parameters	28
Missing Value Handling	Exclude
Maximum Number of Factors	15
Validation Method	Test Set Validation (WORK.VALID)
Validation Testing Criterion	Prob PRESS > 0.1
Number of Random Permutations	1000
Random Permutation Seed	12345

Number of Observations Read	2865
Number of Observations Used	885

Test Set Validation for the Number of Extracted Factors		
Number of Extracted Factors	Root Mean PRESS	Prob > PRESS
0	0.953962	<.0001
1	0.395315	<.0001
2	0.360367	<.0001
3	0.356221	<.0001
4	0.342408	<.0001
5	0.326701	<.0001
6	0.324776	<.0001
7	0.317728	<.0001
8	0.314587	<.0001
9	0.304988	<.0001
10	0.301718	<.0001
11	0.295701	0.0010
12	0.293626	0.0050
13	0.293413	0.0050
14	0.292428	0.0030
15	0.284774	1.0000

PLS

The PLS Procedure

Minimum root mean PRESS	0.2848
Minimizing number of factors	15
Smallest number of factors with p > 0.1	15

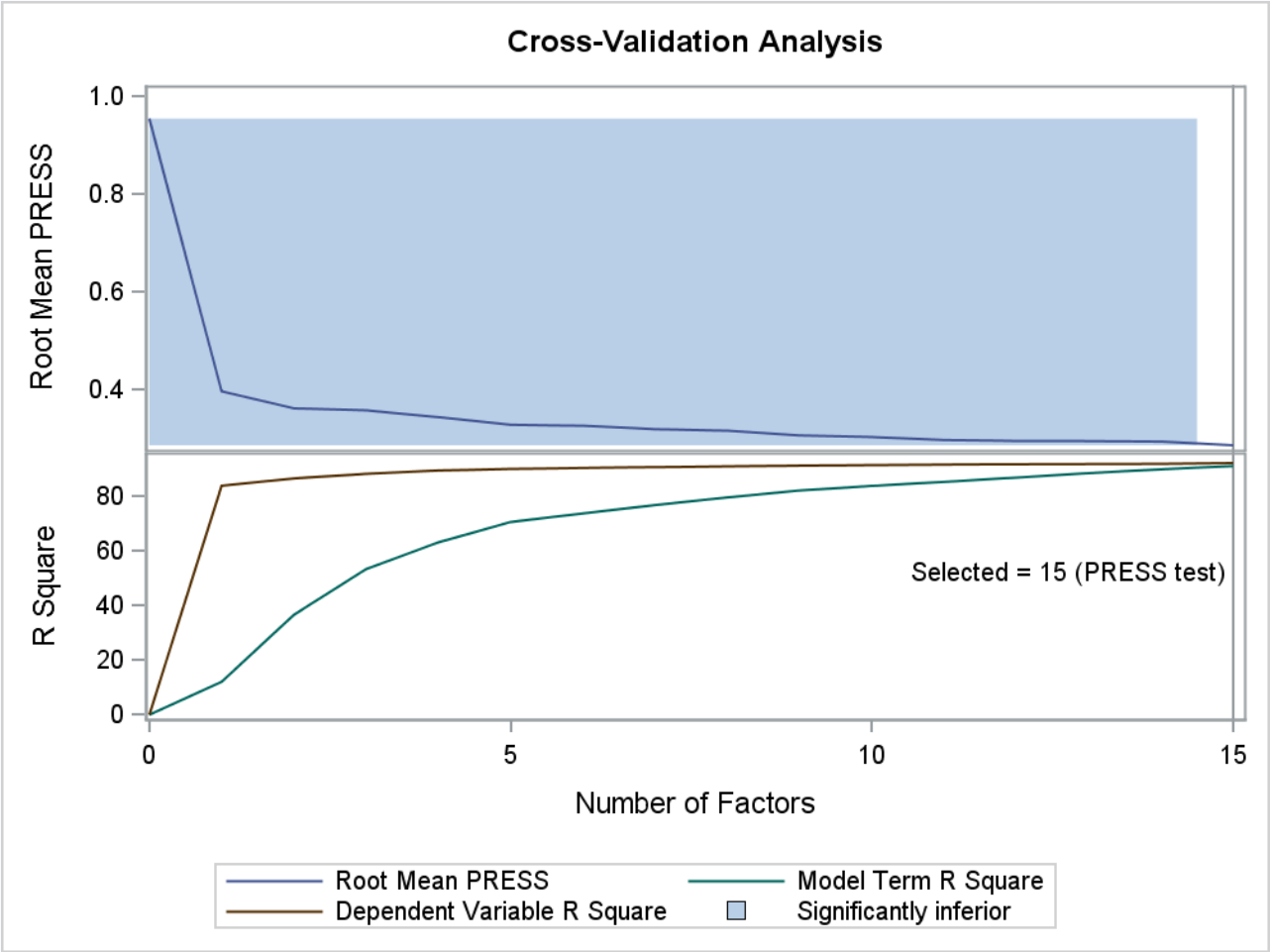
The PLS Procedure

Percent Variation Accounted for by Partial Least Squares Factors

Number of Extracted Factors	Model Effects		Dependent Variables	
	Current	Total	Current	Total
1	12.0440	12.0440	83.9048	83.9048
2	24.5593	36.6033	2.7148	86.6196
3	16.7705	53.3739	1.6480	88.2676
4	9.8004	63.1742	1.2275	89.4951
5	7.4453	70.6195	0.5806	90.0758
6	3.1318	73.7513	0.3600	90.4357
7	3.0646	76.8159	0.2883	90.7240
8	2.7894	79.6053	0.3049	91.0289
9	2.5600	82.1653	0.2488	91.2776
10	1.6701	83.8354	0.2027	91.4803
11	1.4933	85.3287	0.1919	91.6723
12	1.5761	86.9048	0.1184	91.7907
13	1.6267	88.5315	0.0704	91.8611
14	1.3947	89.9262	0.1061	91.9672
15	1.2245	91.1507	0.2823	92.2496

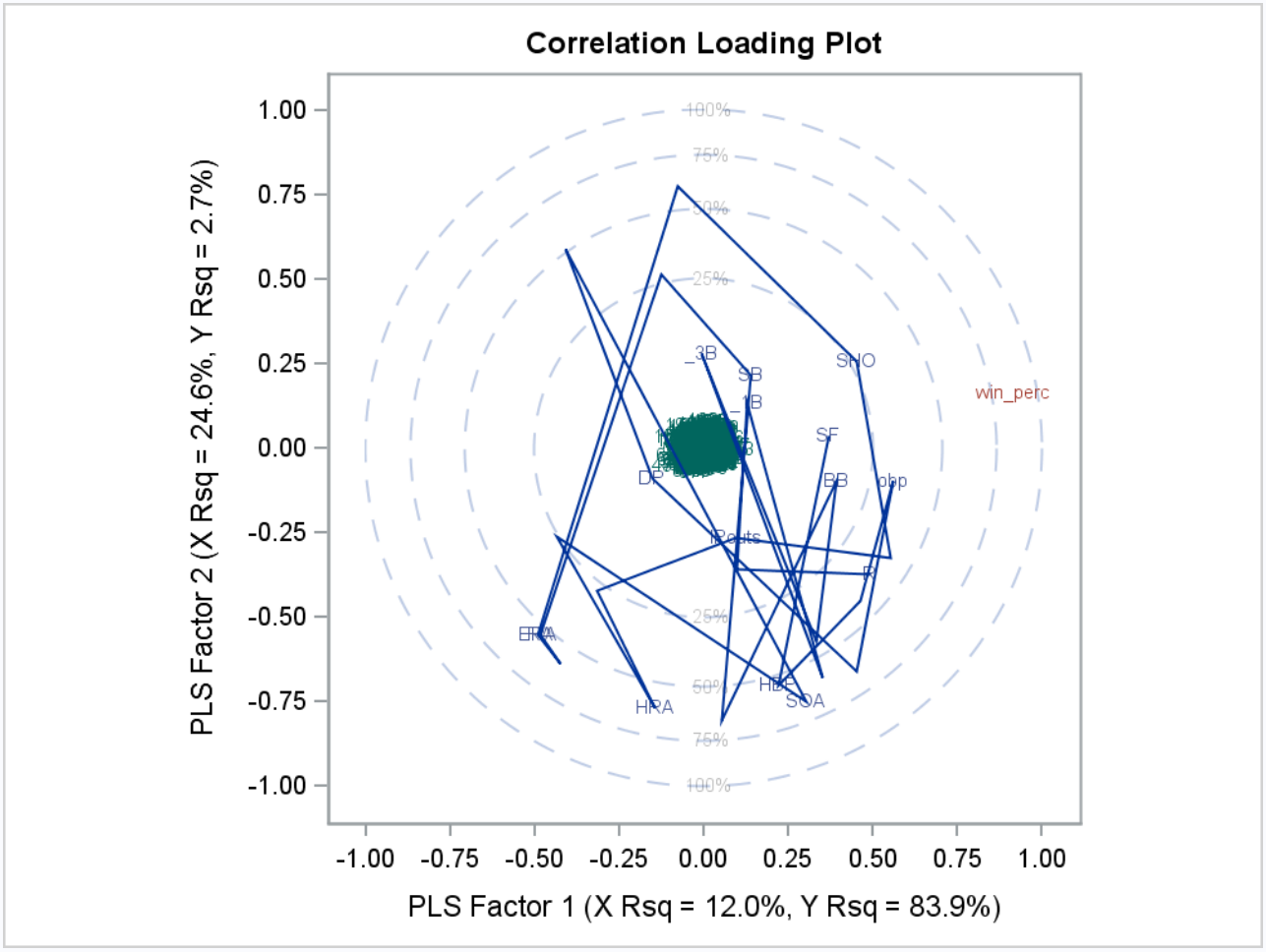
PLS

The PLS Procedure



PLS

The PLS Procedure



Parameter Estimates for Centered and Scaled Data	
	win_perc
Intercept	0.0000000000
R	0.7029364356
AB	-.5933462724
_1B	-.2502325922
_2B	-.1632718185
_3B	-.0501221114
HR	-.1983282454
BB	-.2398216411
SO	-.0114898855
SB	0.0688651253
CS	-.0937997775
RA	-.2801832124
ER	0.0754365328
ERA	-.3490763827
CG	0.1954892317
SHO	0.0514728845
SV	0.2513348786

PLS

The PLS Procedure

Parameter Estimates for Centered and Scaled Data	
	win_perc
lPouts	0.6501136611
HA	-.0169857390
HRA	-.0048337053
BBA	-.0038972311
SOA	0.0456186535
E	0.1518743332
DP	-.0029536970
FP	0.2150821817
obp	0.4231548642
slg	0.2005168544
HBP	-.0515919505
SF	-.0074077940

Parameter Estimates	
	win_perc
Intercept	-4.253459358
R	0.000506086
AB	-0.000119910
_1B	-0.000202953
_2B	-0.000280044
_3B	-0.000363628
HR	-0.000342576
BB	-0.000228853
SO	-0.000004306
SB	0.000121763
CS	-0.000395137
RA	-0.000198607
ER	0.000055651
ERA	-0.042891433
CG	0.000879816
SHO	0.000875777
SV	0.001915287
lPouts	0.000169617
HA	-0.000009762
HRA	-0.000009702
BBA	-0.000003782
SOA	0.000016878
E	0.000436249
DP	-0.000010330
FP	4.162228501
obp	2.055634503

PLS

The PLS Procedure

Parameter Estimates	
	win_perc
slg	0.437791516
HBP	-0.000215388
SF	-0.000057933

The MEANS Procedure

Variable	N	Mean	Std Dev	Minimum	Maximum
SerrorPCR	439	0.000568759	0.000837452	1.5777076E-9	0.0082880
SerrorPLS	439	0.000400524	0.000608476	1.8514471E-9	0.0062473