

Michigan SAS Conference
June 7, 2017

**Lag Modeling of the
Baseball Trade Deadline**

David J. Corliss, PhD

Director, Peace-Work

OUTLINE



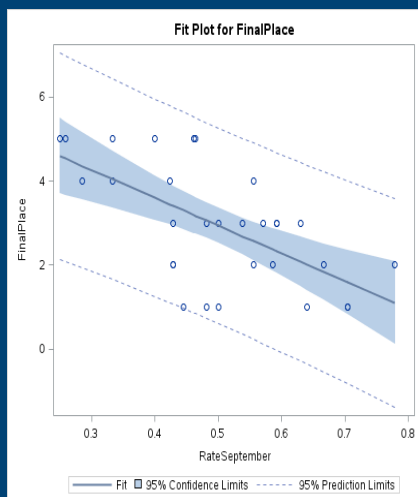
Baseball's Trade Deadline

Lag Models

Regressors by Month

Polynomial Distributed Lag

Summary: Trade Up?



Baseball's Trade Deadline

TRADERUMORS

Top 30 Trade Deadline Candidates For 2017

ESPN

TRADE

Five reasons why the 2017 MLB trade deadline will be insane Sports Illustrated



STATS STANDINGS

USA TODAY

SPORTS

LIFE

MONEY

TECH

TRAVEL

OPINION

67°

CROSSWORDS

WASHINGTON

MLB trade targets: Top 25 players

Lag Models – Variables

Pearson Correlation Coefficients, N = 30

Prob > |r| under H0: Rho=0

	RateApril	RateMay	RateJune	RateJuly	RateAugust	RateSeptember	RateSpring	FinalPlace
RateApril	1.00000	0.28698	0.15902	0.34627	0.66620	0.35075	0.42484	-0.62485
		0.1241	0.4013	0.0609	<.0001	0.0574	0.0193	
RateMay	0.28698	1.00000	-0.02411	0.12697	0.27896	0.34268	-0.11647	-0.49392
	0.1241		0.8994	0.5037	0.1355	0.0638	0.5399	
RateJune	0.15902	-0.02411	1.00000	0.25013	0.06874	0.43940	0.12328	-0.32988
	0.4013	0.8994		0.1825	0.7181	0.0151	0.5163	
RateJuly	0.34627	0.12697	0.25013	1.00000	0.36815	0.35926	0.07401	-0.60854
	0.0609	0.5037	0.1825		0.0453	0.0512	0.6975	
RateAugust	0.66620	0.27896	0.06874	0.36815	1.00000	0.17994	-0.01802	-0.68765
	<.0001	0.1355	0.7181	0.0453		0.3413	0.9247	
RateSeptember	0.35075	0.34268	0.43940	0.35926	0.17994	1.00000	0.14406	-0.62891
	0.0574	0.0638	0.0151	0.0512	0.3413		0.4476	
RateSpring	0.42484	-0.11647	0.12328	0.07401	-0.01802	0.14406	1.00000	-0.09659
	0.0193	0.5399	0.5163	0.6975	0.9247	0.4476		
FinalPlace	-0.62485	-0.49392	-0.32988	-0.60854	-0.68765	-0.62891	-0.09659	1.00000
	0.0002	0.0055	0.0750	0.0004	<.0001	0.0002	0.6116	

Regressors by Month: Just T-1 and T-2

The REG Procedure	
Model: MODEL2	
Dependent Variable: FinalPlace	

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	2	42.62505	21.31252	37.75	<.0001
Error	27	15.24162	0.56450		
Corrected Total	29	57.86667			

Root MSE	0.75134	R-Square	0.7366
Dependent Mean	2.93333	Adj R-Sq	0.7171
Coeff Var	25.61369		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	9.05341	0.71757	12.62	<.0001
RateSeptember	1	-5.50040	1.05785	-5.20	<.0001
RateAugust	1	-6.73155	1.13844	-5.91	<.0001

Regressors by Month: Multicoliniarity

The REG Procedure	
Model: MODEL4	
Dependent Variable: FinalPlace	

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	4	45.76889	11.44222	23.65	<.0001
Error	25	12.09778	0.48391		
Corrected Total	29	57.86667			

Root MSE	0.69564	R-Square	0.7909
Dependent Mean	2.93333	Adj R-Sq	0.7575
Coeff Var	23.71491		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	9.84333	0.92733	10.61	<.0001
RateSeptember	1	-4.48195	1.12232	-3.99	0.0005
RateAugust	1	-5.82807	1.11832	-5.21	<.0001
RateJuly	1	-2.76288	1.13085	-2.44	0.0220
RateJune	1	-0.74086	1.73885	-0.43	0.6737

Regressors by Month: All Months

The REG Procedure
 Model: MODEL5
 Dependent Variable: FinalPlace

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	7	48.12958	6.87565	15.53	<.0001
Error	22	9.73709	0.44260		
Corrected Total	29	57.86667			

Root MSE	0.66528	R-Square	0.8317
Dependent Mean	2.93333	Adj R-Sq	0.7782
Coeff Var	22.67994		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	11.11371	1.22090	9.10	<.0001
RateSeptember	1	-3.39979	1.18492	-2.87	0.0089
RateAugust	1	-5.07058	1.52682	-3.32	0.0031
RateJuly	1	-2.87419	1.08466	-2.65	0.0146
RateJune	1	-1.47438	1.69895	-0.87	0.3949
RateMay	1	-2.53199	1.13909	-2.22	0.0368
RateApril	1	-0.32416	2.00001	-0.16	0.8727
RateSpring	1	-0.67937	1.74593	-0.39	0.7009

Finite Distributed Lag: SAS / ETS

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	35.97258	8.99315	10.27	<.0001
Error	25	21.89408	0.87576		
Corrected Total	29	57.86667			

Root MSE	0.93582	R-Square	0.6216
Dependent Mean	2.93333	Adj R-Sq	0.5611
Coeff Var	31.90303		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	10.54399	1.38856	7.59	<.0001
RateJuly	1	-4.92295	1.42557	-3.45	0.0020
RateJune	1	-4.38429	2.21658	-1.98	0.0591
RateMay	1	-4.68814	1.40730	-3.33	0.0027
Salary	1	-5.76231E-9	3.994239E-9	-1.44	0.1615

Polynomial Distributed Lag: SAS / ETS

```
data work.TradeDeadline;
  set tsa.bb_win_rates_2013;
  X = RateJuly;
  x11 = RateJune;
  x12 = RateMay;
  x13 = RateApril;
  x14 = RateSpring;
run;

proc pdlreg
  data=baseball_time_series;
  model finalplace =
  win_rate(3,3) salary / covb;
run;
```

Polynomial Distributed Lag: SAS / ETS

The SAS System

The PDLREG Procedure

Ordinary Least Squares Estimates			
SSE	17.9268035	DFE	19
MSE	0.94352	Root MSE	0.97135
SBC	81.9457109	AIC	74.632456
MAE	0.72912854	AICC	79.2991226
MAPE	33.4210879	HQC	76.6608421
Durbin-Watson	2.6689	Total R-Square	0.6013

Parameter Estimates

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > t
Intercept	1	-6.4736	7.4093	-0.87	0.3932
X**0	1	8.3016	6.0082	1.38	0.1831
X**1	1	4.8543	1.9866	2.44	0.0245
X**2	1	-3.6729	1.9454	-1.89	0.0744
X**3	1	4.0671	1.4191	2.87	0.0099
Salary	1	-6.761E-9	4.2613E-9	-1.59	0.1291

Estimate of Lag Distribution

Variable	Estimate	Standard Error	t Value	Approx Pr > t	-3.031	0	5.8021
X(0)	-3.031318	2.2262	-1.36	0.1892	*****		
X(1)	4.171272	2.6509	1.57	0.1321		*****	
X(2)	5.624462	2.9892	1.88	0.0753		*****	
X(3)	4.359700	3.1382	1.39	0.1808		*****	
X(4)	3.408434	2.9752	1.15	0.2662		*****	
X(5)	5.802112	2.4950	2.33	0.0313		*****	

Summary: Trade Up...or Say Goodbye?

Lag models use regressors that include previous values of a predictor variables

Distributed lag models incorporate effects over successive time periods

In baseball, the most predictive months occur in different parts of the season

The Trade Deadline is best modeled by current place and the win rate in May and July

Questions

David J. Corliss

davidjcorliss@peace-work.org

www.peace-work.org