Final Data Analysis Project:
Examining a starting pitcher’s change in FIP over the course of the 2015 MLB season

Paul Brendel

Introduction

In 2015, the Major League Baseball (MLB) season lasted from April 5 to October 4. A starting pitcher, who generally pitches every fifth game and around 80-100 pitches per game, may have varying amounts of success over the course of the season. One commonly used statistic to measure a pitcher’s success is Fielding Independent Pitching (FIP), which measures what a pitcher’s Earned Run Average would look like over a given period of time if the pitcher were to have experienced league average results on balls in play [1]. The higher the FIP, the worse the pitcher was performing. A couple factors may influence how a pitcher’s FIP changes over the course of the season: age and salary. Younger pitchers might have a declining FIP towards the end of the season, whereas older pitchers may be more likely to have a consistent FIP over the season due to their experience in handling the long grind of the season. A higher-paid pitcher may have a greater tendency to perform at a more consistent level than a pitcher with a lower salary.

There were 89 starting pitchers who recorded 150+ innings in 2015. The pitcher’s FIP was recorded at each of six time points: April, May, June, July, August, and September (including the few days in October). The pitcher’s age at the start of the season (ranging from 22 to 42) will be analyzed as a continuous, time-fixed variable and a categorical, time-fixed variable (“Young” = <27, “Prime” = 27-30, and “Old” = >30). The pitcher’s 2015 salary (ranging from $507,500 to $32,571,000) will also be analyzed as a continuous and a categorical, time-fixed variable (below $5,000,000 or above $5,000,000, where $5,000,000 was the sample’s median salary). A pitcher’s monthly FIP was considered missing if fewer than 10 innings were pitched that month. Consequently, there were 13 missing monthly values for FIP. Data was obtained from fangraphs.com and baseball-reference.com [2,3].

Analysis

A histogram of the distribution of FIP values revealed that the distribution was reasonably Normal. However, there was one very extreme (>5 SD) outlier (Drew Hutchison's October FIP of 11.28). Removal of this outlier caused the overall mean FIP to change from 3.80(1.17 SD) to 3.78(1.12 SD). Table 1 of the Appendix includes the mean values of overall FIP and the FIP of each of the age and salary categories. The mean FIP amongst old pitchers (4.06) appeared to be noticeably larger than that of young (3.62) or prime (3.65) pitchers, and a two-sample t test (old vs. not old) confirmed that this difference is significant ($p<.001$). Table 1 also includes the overall slope (FIP/month) and the slope for each of the age and salary categories based on a linear, random intercept model. All slopes (overall and category-specific) were small but positive, corresponding to decreasing performance over the season. The slope for the old pitchers (.119) was noticeably larger than that of the young (.029) or prime (.033) pitchers.

A common circumstance for a pitcher to have a missing month is due to injury (i.e. being placed on the disabled list). A unique trend may appear in these circumstances, in which the pitcher will have an increasing FIP leading up the injury month, and then the pitcher's FIP improves (i.e. decreases) after the recovery period. A profile plot was created for the eight pitchers in this sample who had a missing month in the middle of the season (i.e. not April or Sept./Oct.; Figure 1). Their months were rescaled so that the missing month was set to zero. No consistent trend was observed for this sample, but further analysis should examine a larger sample size (thus encompassing injured players from multiple seasons).

A profile plot of 25 randomly selected pitchers from this sample seemed to indicate observations had random intercepts with many fluctuations, but no obvious increase or decrease over time (Figure 2). An empirical summary plot similarly showed a lack of change in FIP over time, with the standard deviation increasing slightly over time (Figure 3). Lastly, a correlation matrix of monthly FIP showed a range of correlation coefficients from .36-.13 with no obvious pattern by lag time (Table 2).

To model the trend in FIP over time, two different fixed effects models were compared: unstructured means and a linear model (with random intercepts). These models were compared based on -2 log likelihood, AIC, and BIC (Table 3). The linear model was most appropriate based on these criteria.

Therefore, although there appears to be large fluctuations in FIP over the season for each pitcher, the time trend is not so difficult that it cannot be described with a linear model.

Model specification using this linear model involved comparing different combinations of the covariates: month, age, and salary (Table 4). Month and age were significant ($p<.05$) predictors of FIP, whereas salary was not. When a month*age product term was added to the model with month and age, this product term was not significant, indicating that the effect of month on FIP does not vary over levels of age. Table 5 includes the model parameters when age is changed to a binary variable (old vs. not old). This model indicates that an old pitcher will, on average, have a FIP that is 0.43 points higher than non-old pitchers over the course of the season.

To determine the optimal covariance structure for the linear model, eight different covariance types were compared based on -2 log likelihood, AIC, and BIC (Table 6). Based on these criteria, the random intercept model fit the data best. This model indicates that pitchers have a random "baseline" FIP value, and that throughout the course of the season the pitcher's FIP will fluctuate around a line with very small positive slope with equicovariance over time. In other words, "good" pitchers will generally end the season as "good" pitchers and "bad" pitchers will generally end the season as "bad pitchers"; rarely will a pitcher substantially improve or worsen throughout the season.

Conclusion

A linear, random intercept model best describes the change in FIP amongst MLB starting pitchers over the six months of the 2015 season. On average, FIP is expected to increase about .06 points every month and pitchers over the age of 30 will have a FIP that is 0.43 points higher than pitchers under the age of 30. A player's change in FIP over month is not expected to differ by pitcher age (i.e. no effect measure modification by age), and salary was not a useful predictor of FIP. Based on our model (with age as a continuous variable), a 25-year-old pitcher is expected to have a FIP of 3.46 for the first month of the season and a FIP of 3.77 for the last month of the season. Future analyses should consider additional covariates to improve the precision of FIP estimates.

Appendix

Group	N	FIP Mean (Stdev)	Slope (FIP/month)
Overall	89	3.78 (1.12)	.062
Young	28	3.62 (1.05)	.029
Prime	30	3.65 (1.12)	.033
Old	31	4.06 (1.14)	.119
<$5,000,000	46	3.71 (1.06)	.048
>=$5,000,000	43	3.86 (1.18)	.076

Table 1. Mean and slope values for the overall sample and the age/salary subgroups.

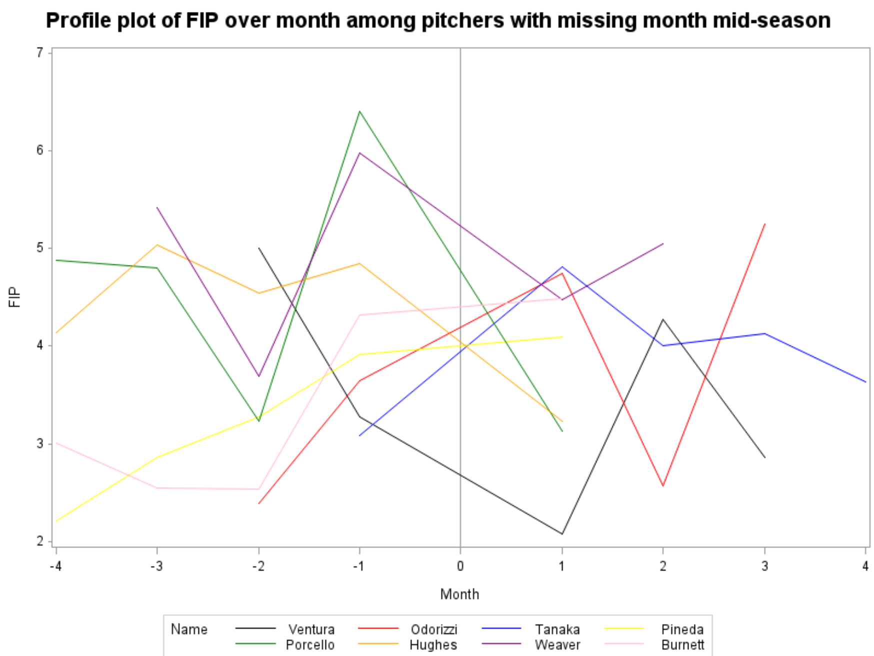

Figure 1. Profile plot of pitchers with mid-season missing months rescaled with missing month = 0.

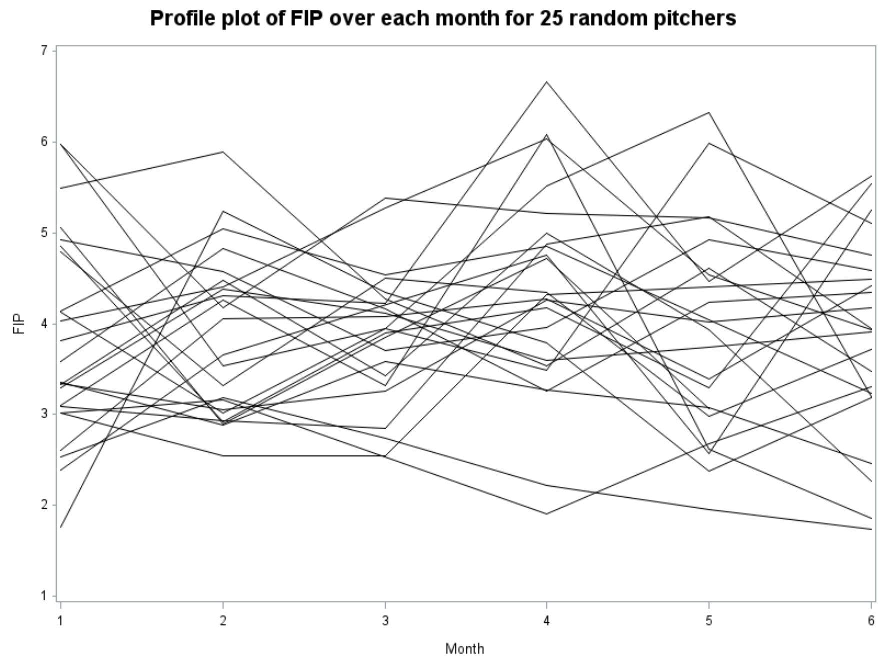

Figure 2. Profile plot of FIP vs. month for 25 randomly selected pitchers.

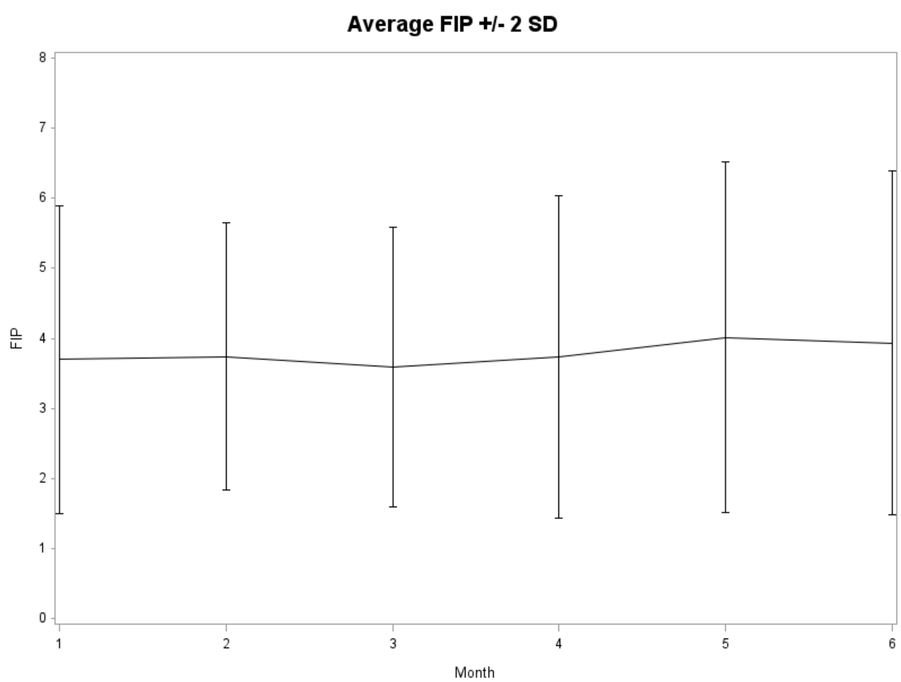

Figure 3. Empirical summary plot of FIP vs month +/- 2 standard deviations.

Month	1	2	3	4	5	6
1	1.00	.22	.25	.16	.30	.13
2	.	1.00	.23	.20	.15	.23
3	.	.	1.00	.34	.36	.15
4	.	.	.	1.00	.20	.13
5	.	.	.	.	1.00	.33
6	.	.	.	.	.	1.00

Table 2. Correlation matrix of FIP by month.

Model	-2 log likelihood	AIC	BIC
Unstructured means	1535.3	1539.3	1544.2
Linear	1524.1	1528.1	1533.0

Table 3. Comparison of fixed effects models. The linear model has the best fit.

Parameter	Estimate	Standard error	T statistic	p-value
Model 1				
Intercept	3.78	.071	53.68	<0.001
Model 2				
Intercept	3.57	0.11	31.46	<0.001
Month	0.062	0.026	2.42	0.016
Model 3				
Intercept	2.12	0.51	4.17	<0.001
Month	0.062	0.026	2.42	0.016
Age	0.051	0.017	2.94	0.0035
Model 4				
Intercept	3.62	0.14	26.26	<0.001
Month	0.055	0.026	2.11	0.036
Salary/$100,000	-0.00029	0.0010	-0.29	0.77
Model 5				
Intercept	1.76	0.56	3.15	0.0023
Month	0.055	0.026	2.11	0.035
Age	0.069	0.020	3.44	<0.001
Salary/$100,000	-0.0019	0.0010	-1.80	0.073
Model 6				
Intercept	2.84	0.83	3.44	<0.001
Month	-0.15	0.19	-0.77	0.44
Age	0.025	0.028	0.89	0.37
Month*Age	0.0072	0.0065	1.11	0.27

Table 4. Model specification. Model 3 was selected as the best model based on the significance ($p<.05$) of its covariates.

Parameter	Estimate	Standard error	T statistic	p-value
Intercept	3.42	0.12	28.00	<0.001
Month	0.062	0.026	2.42	0.016
Age>30	0.43	0.14	3.04	0.0025

Table 5. Model 3 with age changed from a continuous to a categorical variable.

Covariance name	Number of parameters	-2 log likelihood	AIC	BIC	LRT vs independence (X^2)	LRT vs unstructured (X^2)
Unstructured	21	1536.4	1578.4	1630.6	54.5	-
Independence	1	1590.9	1592.9	1595.4	-	54.5
Random Intercept	2	1556.2	1560.2	1565.1	34.7	19.8
Autoregressive	2	1565.8	1569.8	1574.8	25.1	29.4
ARMA(1, 1)	3	1554.5	1560.5	1567.9	36.4	18.1
Ante-dependence	11	1553.6	1575.6	1603.0	37.3	17.2
Factor Analytic 1	12	1544.6	1568.6	1598.4	46.3	8.2
Toeplitz	6	1551.2	1563.2	1578.2	39.7	14.8

Table 6. Covariance structure selection. The random intercept structure fit the data best.

References

1. http://www.fangraphs.com/library/pitching/fip

2. http://www.fangraphs.com

3. http://www.baseball-reference.com