**Pitcher Fatigue Analysis. Regression Results**

NOTE

When these colors are used to shade cells, they refer to that coefficient’s level of statistical significance. A lover p-value means higher significance (More confidence that the coefficient is not 0). Generally any p-value below 0.05 is considered enough to be statistically significant

Table

Description automatically generated

**Part 1: Continuous regression of pitch count against Spin Rate, Zone Speed, and Horizontal and vertical break for five pitch types.**

Table

Description automatically generated

Each additional pitch thrown by a pitcher is associated with a .19 rpm decrease in spin rate and a 0.013 MPH decrease in Release Speed for fastballs. The rest of the table is interpreted similarly with respect to the pitch type and measurement.

**Part 2: Factor Level Regression against Spin Rate, Zone Speed, and Horizontal and vertical break for five pitch types.**

To cut down on variance, pitch counts were grouped into groups of 10. 0 to 9, 10 to 19 and so on. Each table now has coefficients for a single measurement.

These coefficients represent the change in that measurement associated with the pitch count being in that bin as opposed to the expected measurement for pitches in the first bin (0 to 9). The first bin, 0 to 9 is treated as the intercept. Thus the value for that bin is the expected differences in these measurements for a pitcher throwing his first ten pitches of the outing compared to his expected measurements without knowing the pitch count.

Table

Description automatically generated

For instance, if a pitcher has already thrown 70 pitches in a game, we would expect the spin rate on his fastball to be, on average, 15.319 rpm slower than if he were just coming into the game We would expect the Release speed on his cutter would be, on average, 0.609 MPH slower than his first ten pitches.

The first column (intercept) is interpreted differently than the other 9. The coefficient for intercept in this context represents the expected difference between a pitchers first ten pitches and their average for that year.

NOTE: The red in the bottom table for horizontal break means that many of these coefficients are not statistically significant. Thus we don’t have convincing evidence that pitch count has an effect on the horizontal break of a pitch. The Zone Speed table, however, has lots of blue meaning that we often do have strong evidence that the pitch count has an effect on zone speed.

Below we have the same information represented graphically. Again we have one chart for each measurement, and one line for the effect that each pitch count bin (coefficient) has on that measurement for each pitch type.

The shape in the middle of each line is the point estimate for the effect of being in that pitch bin. The colored line on either side represents the uncertainty. Small intervals are estimates that were calculated with a high degree of confidence. Usually because of large sample size or low variation. If the interval covers 0, the dotted line, we wouldn’t conclude that that bin has a significant effect. The corresponding cells in the grid above will often be red or yellow if this is the case (High P-Value)

**Spin Rate Chart:**

A picture containing graphical user interface

Description automatically generated

**Release Speed Chart:**

A picture containing diagram

Description automatically generated

**Vertical Break Chart:**

Diagram

Description automatically generated with medium confidence

I probably wont even include the horizontal break chart. None of the effects were really significant anyways. If you want to see it I can make it real quick.

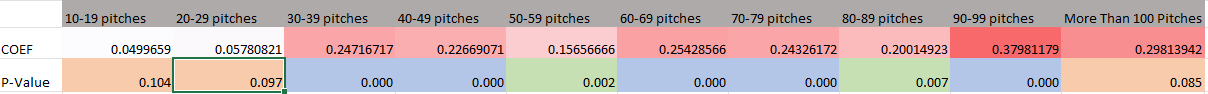
Even with evidence that these measurements change with pitch count, we can’t immediately infer that it has a significant effect on the outcomes of these pitches. Statistical significance doesn’t always mean the effect is significant for practical purposes.

The following regressions aim to show whether these differences in measurements actually make a difference on pitch effectiveness by looking at the relationships between pitch count and stats such ad WHIP and FIP.

**PART 3a: Factor Level Regression of Pitch Count Against WHIP**

|  |  |  |
| --- | --- | --- |
| first\_pitch\_group | count(\*) | whip |
| <chr> | <int> | <dbl> |
| 0-9 pitches | 88714 | 1.775129 |
| 10-19 pitches | 58247 | 1.730861 |
| 20-29 pitches | 41325 | 1.663266 |
| 30-39 pitches | 30163 | 1.752321 |
| 40-49 pitches | 22821 | 1.726143 |
| 50-59 pitches | 17811 | 1.666874 |
| 60-69 pitches | 14384 | 1.693023 |
| 70-79 pitches | 10794 | 1.669259 |
| 80-89 pitches | 7328 | 1.573573 |
| 90-99 pitches | 3858 | 1.610526 |
| > 100 Pitches | 1579 | 1.53816 |

If we group all the plate appearances in the dataset by what the pitch count was, we might be surprised to see that WHIP decreases as the pitch count increases. However, it would be a mistake to conclude that pitch count decreases WHIP. This is likely biased because the better pitchers are usually the ones who go deep into games so the higher bins are comprised almost entirely of datapoints from only the leagues best pitchers. This is why a regression gives us better insight.



Interpretation here is similar to part 2. For instance, Pitchers with pitch counts in the 70’s would be expected to be throwing at a level that would produce a WHIP that’s 0.2433 higher than when that pitcher first comes into the game and has a pitch count between 0 and 9.

The same chart as before is below. Estimate (X-axis) is how much we’d expect a pitcher’s WHIP to change if they have a pitch count in that bin as opposed to a pitch count In the first, lowest bin. Error bars show uncertainty.

**WHIP Coefficient Chart:**

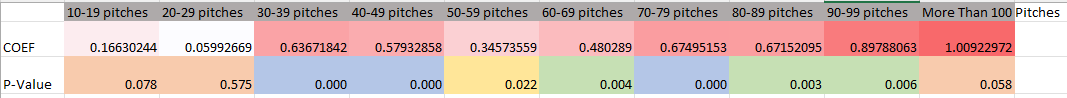
Chart, box and whisker chart

Description automatically generated

The only difference between the blue and orange bars is that for the blue bars, I compared a pitcher’s WHIP in a particular bin to that pitchers WHIP for that season and for the orange bars I compared a pitcher’s WHIP in a particular bin to that pitchers WHIP for all seasons (every Datapoint we have). It probably makes more sense to do it the blue way since a player’s performance can vary from season to season, but obviously it looks like it didn’t make much of a difference.

**PART 3b: Factor Level Regression of Pitch Count Against FIP**

Very similar process here as for the WHIP regression above. Pitcher FIPs in each bin were compared to their overall FIP. These differences were the dependent variable in a regression against pitch count. Results are below:



As with whip, FIP tends to increase with the pitch count. A number of the pitch count bins shows statistically significant differences, particularly 30-49, likely because those bins have relatively large sample sizes.

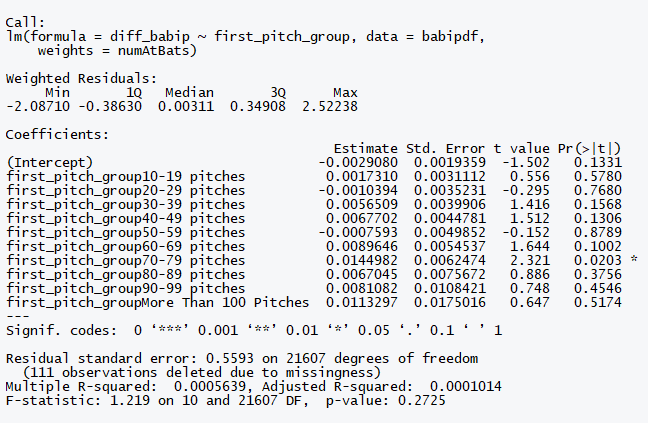
**FIP Coefficient Chart:**

Chart, box and whisker chart

Description automatically generated

**BABIP Regression**

The regression of pitch count against Batting Average for Balls in Play (BABIP) showed no association. The regression output is below. It’s interesting that pitchers do perform better with lower pitch counts but balls put into play aren’t more likely to be hits. The regression output is below.



**Conclusion:**

As expected, the data shows evidence that with a higher pitch count, certain measurements such as Spin Rate and Release Speed decrease as the pitcher becomes more fatigued. It’s difficult to say how these small differences in speed or RPM with effect the effectiveness of a pitch. This is why we looked into the effect pitch count has on a results-based metric such as WHIP and FIP. Here we saw evidence that a higher pitch count is associated with a higher whip and fip (Worse performance).

When we interpret these results, we must always consider if there’s any bias. In this case, the most likely cause for bias is selection bias. A pitcher who is having a good day is more likely to reach a higher pitch count. Therefore, it’s possible that the coefficients for the high pitch count bins are biased downward since those bins are comprised mostly of pitchers who are doing well that day while struggling pitchers were likely pulled out of the game before they could reach a pitch count that high. For this reason, I think that these estimates could be used as a low estimate for the effect of pitch count, with the true effect being at least that great or possibly greater.