



Blaseball Is a Mess, Fair Play Is the Future: An analysis of odds, outcomes, and accuracy in the Peace and Prosperity Era

Games Pergame (@TwinCanine#1333) (he/they)*,1

*Society for Internet Baseball Research

ABSTRACT Analysis of game odds and matchups between Season 2 and Season 12 indicate that the new management is making good on promises to increase fair play in Internet League Blaseball, though that process did not begin when they took over.

KEYWORDS

Odds
Fairness
Schedule

INTRODUCTION

At the end of Season 10 after the defeat of the Peanut, the Hall Monitor introduced the fans of Internet League Blaseball to the Boss². She informed the gathered mass of confused, surprised, and delighted fans that “Blaseball is a Mess.” Among other things, the Boss assured the fans that, with the new ownership of ILB in the Peace and Prosperity Era, “Fair Play is the Future.” Now that the Expansion Era has begun, this paper will examine historical trends in both the Discipline Era and the Peace and Prosperity Era and assess the Boss’s promise.

METHODOLOGY

The sample examined in this paper is comprised of 10860 regular season games of blaseball in Seasons 2 - 12.³ To examine Fairness in blaseball, we examine the historical schedule of ILB play and assessing the odds determined for the matchup to see how evenly matched the teams were before gameplay.⁴ Based on the odds presented, we assign a Fairness Value (FV) to the matchup based on how far the odds are from a 50/50 split. Under this system, a 45/55 split would result a FV of 5, a 90/10 split would result in an astronomical FV of 40, and a 50/50 split would return a desired FV of 0.⁵

Each game has an FV calculated, and then games are summed to find a score for each day. Once grouped by season, trends in

Fairness begin to emerge. For example, the recently completed Season 12 shows a fairly robust range of FVs, but a fairly slight downward trend as the season progresses in Figure 1 (For full figures of trends by season, see Appendix B). This paper will also be examining the presence and frequency of Fair Games, or games with an FV of 0, across the seasons.

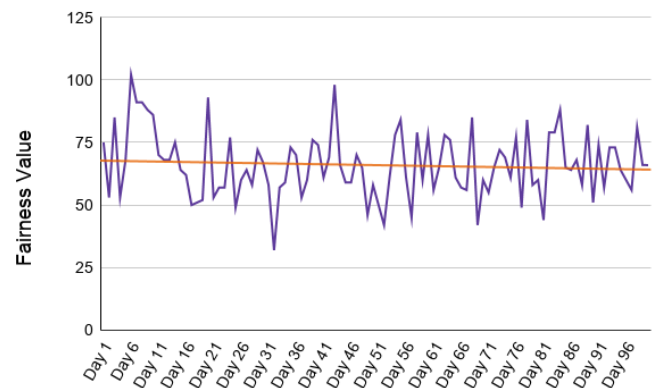


Figure 1 Total FV per day (purple line) during Season 12. Trendline is plotted in orange.

For the purposes of this paper, the seasons will be separated by

Copyright © 2021 by the Society for Internet Baseball Research
Manuscript compiled: Wednesday 31st March, 2021

¹@gamespergame on Twitter.

²AKA the Coin, AKA new ownership (Girllboss, Gatekeep, Gaslight)

³3 days worth of games are being excluded because of Waveback-related reasons, and other data gaps

⁴Current Society understanding is that the odds presented by the ILB are useful and statistically reflective of outcomes, and analysis of odds versus outcome in this sample supports that understanding

⁵Each game’s Fairness Value was rounded to match the odds presented to fans.

Era. Season 2 to 10 (8880 games) will comprise the Discipline Era, and Seasons 11 and 12⁶ (1980 games) will comprise the Peace and Prosperity Era (PaP Era). By comparing these Eras, we can begin to understanding the history, change, and present of Fairness.

RESULTS

The Discipline Era

An average day in the Discipline era would have an FV sum of 83.12, which roughly equates to games having an 58 / 42 odds (See Table 1). These values spike for seasons 4 and 5, but after the High Filter Decree was passed during the Season 5 elections and league and division placement were changed to match historical performance, FV begins to decrease as the Era progresses. Fair Games count is also interesting to observe for this period. Seasons 4 and 5, the seasons with the highest average FV, also possess the lowest count of Fair Games.

■ **Table 1** Fairness Values and Fair Games in the Discipline Era

| Season | Average FV | Fair Games |
|----------------|------------|------------|
| Season 2 | 81.62 | 27 |
| Season 3 | 89.87 | 28 |
| Season 4 | 95.03 | 17 |
| Season 5 | 94.81 | 19 |
| Season 6 | 85.23 | 26 |
| Season 7 | 74.21 | 32 |
| Season 8 | 78.24 | 29 |
| Season 9 | 77.63 | 23 |
| Season 10 | 71.47 | 26 |
| Discipline Era | 83.12 | 25 |

A better understanding of how FV are distributed amongst days is to separate the values into quartiles, as in Table 2 and Figure 2. A stark decrease is present in the higher bounds of Daily FV, and though the lower bound move up quite a bit in Season 4, all categories trend down after that point.

⁶ The Expansion Era began during Latesiesta Season 12, but in the interest of inclusion and some semblance of statistical significance, Lateseaon 12 was included.

■ **Table 2** Distribution of Daily Fairness Values in the Discipline Era

| Season | Max FV | Upper Quart. | Average | Lower Quart. | Min FV |
|----------------|--------|--------------|---------|--------------|--------|
| Season 2 | 144 | 94.36 | 81.62 | 69.62 | 50 |
| Season 3 | 121 | 98.66 | 89.87 | 77.4 | 53 |
| Season 4 | 135 | 108.74 | 95.03 | 84.77 | 70 |
| Season 5 | 126 | 105.58 | 94.81 | 81.96 | 61 |
| Season 6 | 132 | 97.25 | 85.23 | 74.57 | 54 |
| Season 7 | 107 | 84.93 | 74.21 | 64.1 | 51 |
| Season 8 | 114 | 89.63 | 78.24 | 67.48 | 47 |
| Season 9 | 110 | 87.13 | 77.63 | 67.86 | 40 |
| Season 10 | 102 | 80.6 | 71.47 | 61.33 | 34 |
| Discipline Era | 121.22 | 94.1 | 83.12 | 72.12 | 51.11 |

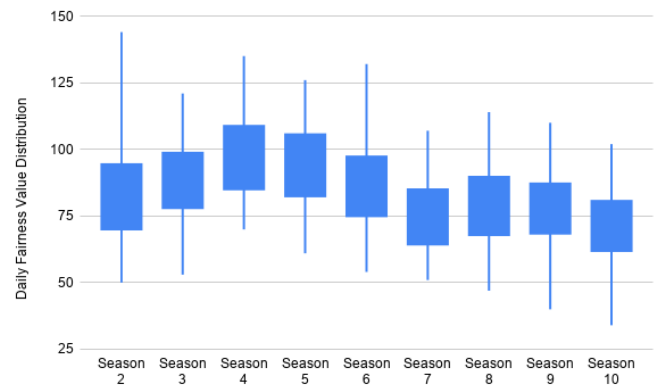


Figure 2 Daily Fairness Values of each Season from Table 4 as Box Plot

On an individual game basis, FV distribution also provides some interesting data. If you refer to Figure 3, you will have seen the FV for all 10860 games in the sample.⁷

These FVs necessarily bottom out at 0 for the Fair Games discussed above in the daily analysis, and largely cluster around the bottom of the chart, rarely rising above 20. A vast majority of ILB games occur in a 70/30 split or closer odds, regardless of era.

Fairness Values can also be grouped into 5 point categories to see how each season fares on intensity of Unfairness in Table 5 in Appendix A. We again see how deeply unfair Season 4 and 5 were. Those two seasons have almost twice as many games with a FV over 20, in the Deeply Unfair category, than any other single season, and have more games in that category than some seasons combined. These seasons didn't have significantly fewer games in the Mostly Fair category (FVs between 1 and 5), but those deeply unfair games really effect the overall divide of the season (see Figure 4). These breakdowns can also be by day in Appendix C.

⁷ Congratulations, and I'm sorry.

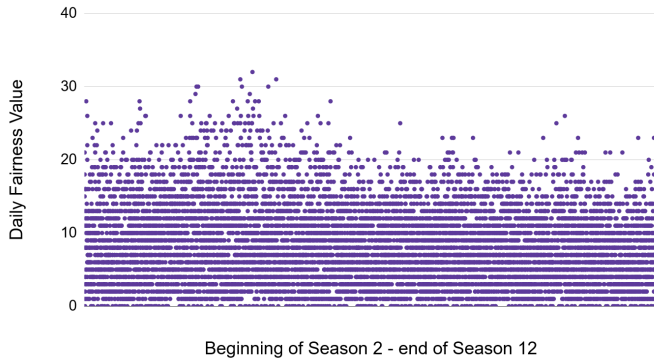


Figure 3 Daily Fairness Values across the sample size, sorted by season and day horizontally.

The Peace and Prosperity Era

Though smaller in size, the Peace and Prosperity Era has a fairly tight grouping of Average FV by day, and follow the general trend of distribution of Fair Games, seen in Table 3. Average FV across both seasons roughly equates to a 57 / 43 split on an average day.

Table 3 Fairness Values and Fair Games in the Discipline Era

| Season | Average FV | Fair Games |
|-----------|------------|------------|
| Season 11 | 74.04 | 27 |
| Season 12 | 65.94 | 32 |
| PaP Era | 69.99 | 29.5 |

Season 11 shows a slightly higher Daily FV than the previous season in the Discipline Era, though this is potentially attributable to the under-developed Tokyo Lift’s entry into ILB. With an election under their belts, Season 12 had a historically low Average FV of 66, and it matched the previous record for Fair Games (32 games across 99 days, held by Season 7). Looking at Daily FV distribution in Table 3, we see the same pattern play out - a slight increase to individual distribution in Season 11 versus Season 10, but well below Discipline Era averages and fitting the downward trend. Season 12 shows historic lows in all categories.

Table 4 Distribution of Daily Fairness Values in the Discipline Era

| Season | Max FV | Upper Quart. | Average | Lower Quart. | Min FV |
|-----------|--------|--------------|---------|--------------|--------|
| Season 9 | 114 | 83.7 | 74.04 | 65.22 | 40 |
| Season 10 | 102 | 77.11 | 65.94 | 56.62 | 32 |
| PaP Era | 108 | 80.41 | 69.99 | 60.92 | 36 |

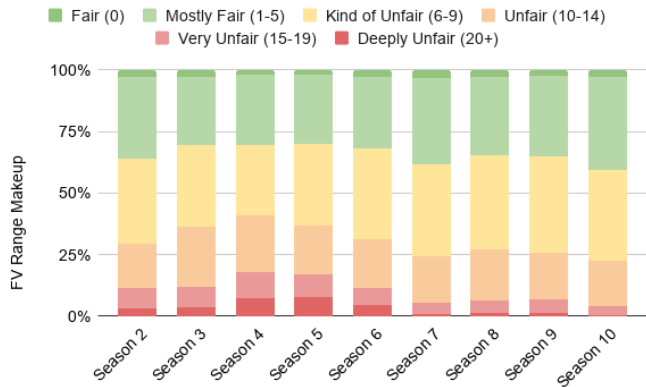


Figure 4 Percent of Games in each FV Range by Season in the Discipline Era

On a per-game basis, Seasons 11 and 12 also showed a noticeable decrease. While combined with the previous era in Figure 3, the end of that plot shows a clear reduction of games around and just below an FV of 20.⁸

The Peace and Prosperity Era continues the trends set in seasons 9 and 10 of decreasing the amount of games that fall in higher categories (10+), and increasing the proportion of games falling in the lower categories (9-), as illustrated in Table 6 in Appendix A. Season 12 returns the Very Unfair (15-19) range down to the point where it doesn’t show up in Figure 5. These breakdowns can also be observed more fully in Appendix C.

ANALYSIS

Baseball really *was* a mess. The distribution of FV categories shifting began during the Discipline Era, but the games played in the Peace and Prosperity Era were far more fair on average than the seasons before. At time of recording, it seems that new management intends to make good on their promise of Fairness, at least in terms of scheduling fair matchups across a season.

⁸ Again, I’m sorry.

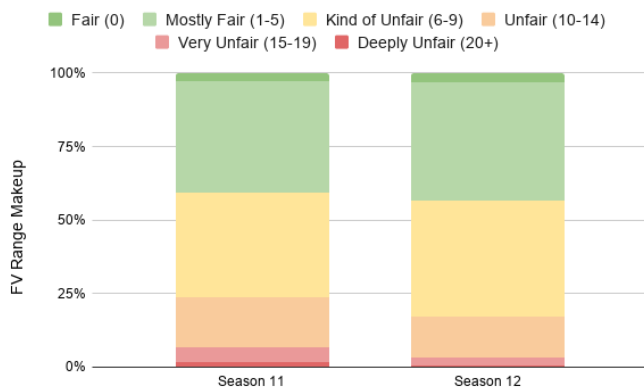


Figure 5 Percent of Games in each FV Range by Season in the Peace and Prosperity Era

FUTURE WORK

There is more Fairness to be gleaned in a few areas that are immediately identifiable. Just outside the scope of this paper is an analysis of Home Field Advantage, and, if there is an advantage, whether teams are scheduled for Home Games at roughly even rates. Additionally, based on initial work done by CowboyMoth#8639 and a few other researchers on Strength of Schedule and related metrics⁹, there is opportunity to look at how “easy” a given team’s schedule is in each season, if teams are regularly given challenging schedule, and if so, if specific teams are being picked on by the ILB. Finally, there is some work to be done on weather. Sun II and Black Hole in particular do not appear to effect the usefulness of the presented odds, but there is a chance they are impacting the strength of the odds’ evaluations. Other weathers too have opportunities for unfairness - a given team may see specific weather types more frequently than other teams do, or more frequently than they see other weather types, and while it is unlikely that weather would affect either odds or outcome, there is always potential to investigate.

ACKNOWLEDGEMENTS

Special thanks to BaronBliss (@baronbliss#7135) (any) for creating this template. Thanks to Astrid (@Ske#6201) (she/her) and the developers of Chronicer for the robust API tool that made gathering game data extremely easy. Thanks to the SIBR community for encouraging inquisitiveness and providing excellent feedback.

REFERENCES

⁹ SIBRmetrics?

APPENDIX A - GAME FV CATEGORY COUNTS

■ **Table 5 FV Category Counts in the Discipline Era**

| Season | Deeply Unfair (20+) | Very Unfair (15-19) | Unfair (10-14) | Kind of Unfair (6-9) | Mostly Fair (1-5) | Fair (0) |
|----------------|---------------------|---------------------|----------------|----------------------|-------------------|----------|
| Season 2 | 30 | 82 | 178 | 343 | 330 | 27 |
| Season 3 | 33 | 83 | 242 | 329 | 275 | 28 |
| Season 4 | 70 | 100 | 221 | 275 | 277 | 17 |
| Season 5 | 75 | 91 | 198 | 327 | 280 | 19 |
| Season 6 | 45 | 68 | 198 | 364 | 289 | 26 |
| Season 7 | 9 | 45 | 187 | 368 | 349 | 32 |
| Season 8 | 11 | 51 | 208 | 379 | 312 | 29 |
| Season 9 | 11 | 54 | 191 | 387 | 324 | 23 |
| Season 10 | 5 | 35 | 181 | 367 | 376 | 26 |
| Discipline Era | 32.11 | 67.67 | 200.44 | 348.78 | 312.44 | 25.22 |

■ **Table 6 FV Category Counts in the Discipline Era**

| Season | Deeply Unfair (20+) | Very Unfair (15-19) | Unfair (10-14) | Kind of Unfair (6-9) | Mostly Fair (1-5) | Fair (0) |
|----------------|---------------------|---------------------|----------------|----------------------|-------------------|----------|
| Season 11 | 18 | 47 | 172 | 350 | 376 | 27 |
| Season 12 | 7 | 24 | 138 | 393 | 396 | 32 |
| Discipline Era | 12.5 | 25.5 | 155 | 371.5 | 386 | 29.5 |

APPENDIX B - DAILY FV PER DAY CHARTS FOR EACH SEASON

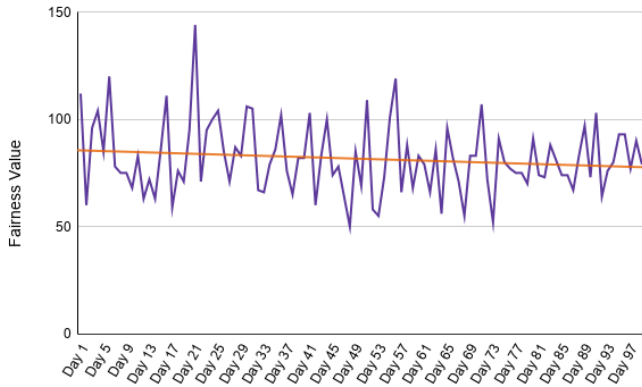


Figure 6 Total FV per day (purple line) during Season 2. Trendline is plotted in orange.

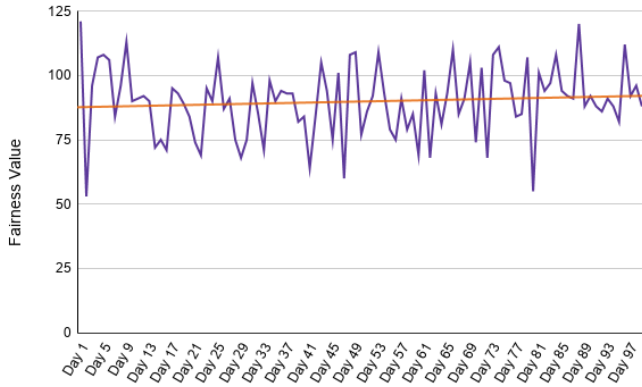


Figure 7 Total FV per day (purple line) during Season 3. Trendline is plotted in orange.

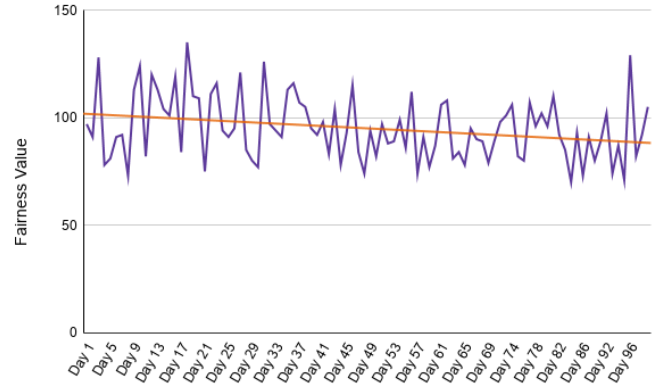


Figure 8 Total FV per day (purple line) during Season 4. Trendline is plotted in orange.

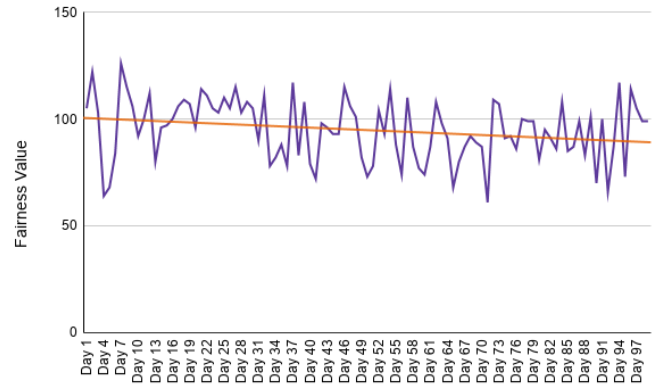


Figure 9 Total FV per day (purple line) during Season 5. Trendline is plotted in orange.

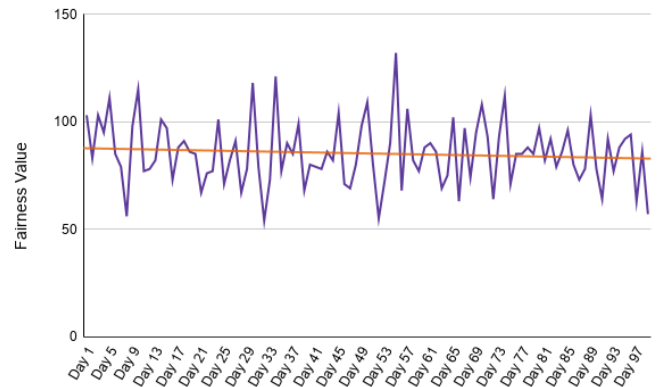


Figure 10 Total FV per day (purple line) during Season 6. Trendline is plotted in orange.

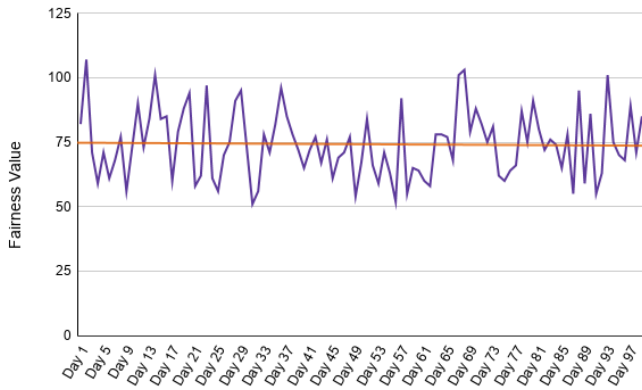


Figure 11 Total FV per day (purple line) during Season 7. Trendline is plotted in orange.

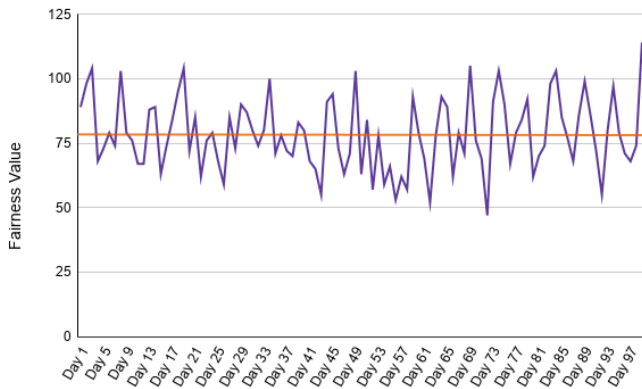


Figure 12 Total FV per day (purple line) during Season 8. Trendline is plotted in orange.

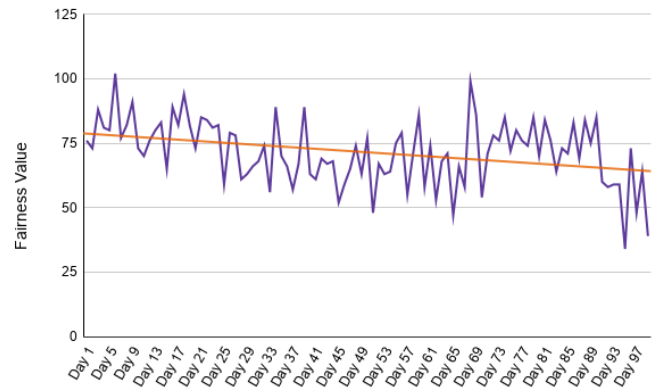


Figure 14 Total FV per day (purple line) during Season 10. Trendline is plotted in orange.

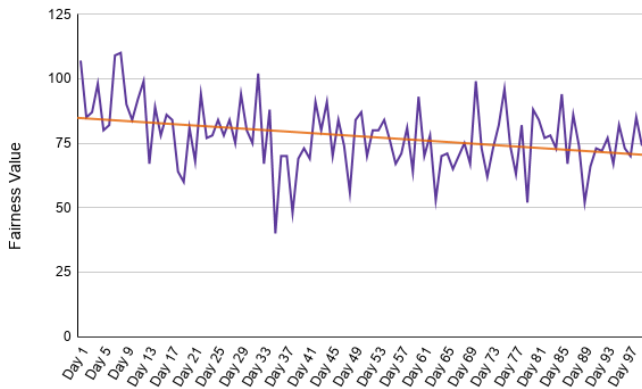


Figure 13 Total FV per day (purple line) during Season 9. Trendline is plotted in orange.