Data analysis of financial burden index through KBO league FA pitcher's performance and contract amount size

Semin Kim¹, Kangsoo You²

¹Department of Computer Education, Jeonju National University of Education, Republic of Korea
²School of Liberal Arts, Jeonju University, Republic of Korea

ABSTRACT

Due to the recent pandemic of COVID-19, sports events around the world are being hit hard in 2020. However, with the KBO league having been opened even though it is late, it has attracted the attention of baseball fans around the world as it becomes an overseas broadcast. However, since the Republic of Korea is also affected by the global economic crisis, the financial burden of professional baseball teams that rely on their parent company has become a hot topic. Therefore, in this study, we collected records such as IP, ERA, and WAR of KBO league FA pitchers and news related to FA contracts. Then, we preprocessed the data and analyzed it. To this end, we analyzed the performance and financial burden indicators according to the pitcher's age and contract years. In addition, it was found that the performance of athletes was maintained until the mid 30's, and through rational FA contracts, the team and athletes were able to suggest win-win strategies. And with this basic data analysis in the sports industry, the possibility of utilizing these strategies was confirmed.

Keywords: Big data of baseball, Free agent, KBO league, Performance, Sports data analysis

This is an open access article under the CC BY-SA license.

Corresponding Author:
Kangsoo You
School of Liberal Arts
Superstar College, Jeonju University
303 Cheonjamro, Wansangu, Jeonju 55069, Republic of Korea
Email: gsyou@jj.ac.kr

1. INTRODUCTION

Since the beginning of 2020, COVID-19 has emerged as a pandemic epidemic, and many sports events worldwide have been hampered. Additionally, even in the spring of 2020, COVID-19 did not end, and several winter sports events ended their seasons early, and even major league baseball (MLB) and Nippon professional baseball (NPB) have not been opened as of May 2020. However, in the Republic of Korea, the KBO league has been opened and has held games since May 4, although there are no spectators. As a result, the ESPN broadcasting station in the United States is broadcasting KBO's baseball games, and the KBO is gaining attention as a part of 'the Korean Wave'. Therefore, in this study, we intend to conduct the study in a way that introduces the records of FA pitchers who were active in professional baseball throughout the world.

Sports is an area that contains vast amounts of data. Through the current information society, sports have been computerized and accumulated a lot of data [1, 2]. Most of the game contents that are generated in the game are encoded and recorded in a systematic way by the scorer [3, 4]. Baseball is called a 'sport of record'. Baseball can mark records in a variety of ways [5-8]. Accumulated records such as home runs, hits, saves, and steals can also be saved, and can be used as records using simple formulas such as batting average, on-base percentage, slugging percentage and average ERA. Baseball records even use complex formulas and
statistics, such as walks plus hits divided by innings pitched (WHIP), wins above replacement (WAR), and fielding independent pitching (FIP), as records [9, 10]. In professional baseball, pitching is a very important power. The importance of a pitcher in a baseball game cannot be overstated. The phrase ‘baseball is a pitcher's game’ is a quote spoken among sports media and baseball fans around the world. In other words, baseball is a sport that presupposes that pitchers are basically advantageous [11, 12]. Teams without an ace pitcher have a very low chance of achieving good results. Therefore, whenever the FA market opens, pitchers accrue high popularity and value [13]. Therefore, this study intends to conduct research mainly on pitchers from among FA players.

The professional baseball teams in Korea operate in a structure that relies on financial support provided by the parent company of the team. The KBO originally had no purpose in generating profits because it operated in the early days as a concept of the effects of advertising and social contribution. However, as the size of the professional sports market grows year by year, it spends a small amount every year, is rapidly expanding the size of the FA contract market, and the deficits of the clubs are intensifying day by day [14]. Accordingly, in recent years, various methods have been devised to improve the profit structure and reduce costs to create profits in the clubs. Therefore, the demand for objective analysis results that both teams and players can accept at the same time is increasing [3, 15-18].

Research that analyzes the annual salary and performance of professional baseball pitchers has been ongoing. As an example of previous studies, one study grasped the relationship between annual salary and game performance by using sabermetrics using WHIP, FIP, WAR and classical records such as earned runs average (ERA) or winning games [19, 20]. Additionally, in relation to the calculation of annual salary, there were also studies that analyzed variables that significantly correlated with the annual salary increase among the demographic indicators, the joining and transfer information of athletes, and traditional performance indicators and sabermetric indicators [21]. There is also a study that analyzed the records that a pitcher can control himself, which are the main factors affecting the salary of Korean professional baseball pitchers [15].

In Korea’s professional baseball, however, there are many cases in which players recruited through the FA market have demonstrated good performance, but there are also many that do not. Therefore, the clubs have made great efforts to prepare a plan to set a salary at a reasonable level. In this study, we tried to compare the financial burden indicators against age and the financial burden indicators against annual FA base contract amount, average ERA, and win contribution to substitute players to propose and report reference materials for reasonable FA contracts.

2. RESEARCH METHOD
2.1. Pitcher’s records

The various indicators used in this study among the records for evaluating pitchers are as follows, and the initial notation is ERA and WAR. In addition, in order to express the performance efficiency compared to the annual salary, the financial burden index (FBI) using the total contract amount of FA and the contract period WAR was expressed. The various records used in this study are as [22, 23].

2.1.1. Earned runs average (ERA)

ERA means the ratio of earned runs (ER), which is the score that the pitcher is responsible for, against the scores allowed by the pitcher in 9 innings. ERA was the most commonly used pitcher evaluation index with Winning Games, but there was a contradiction in treating the overall team's performance as a pitcher's only statistics. The formula for calculating the average ERA is ERA=9(ER)/IP [22, 23].

2.1.2. Wins above replacement (WAR)

WAR is a representative record of the value of players in Sabermetrics. If the value of WAR is 1, it means that the player added 1 win to the team rather than a player who could not help the team win. For example, if player A has a WAR value of 3.5, this means that he added 3.5 wins to the candidate who played in the season or a player at level 1.5. The formula for obtaining the pitcher's WAR is as: first, the formula for obtaining the starting pitcher WAR, and then the method for obtaining the relief pitcher WAR [24, 25].

The Starting Pitcher’s WAR=((Expected Winnings per 9 Innings-Expected Winnings per Replacement Player per 9 Innings)×Pitched Innings)/9

1) Expected Winning per 9 Innings=((A League’s Lose Points-Corrected Loss Points)/Runs per Winning)+0.5
2) Corrected Loss Points=(FIP/0.92)/Park Factor
(3) \[ \text{Runs per Win} = \left( \frac{\text{Opposing Team’s Innings} \times \text{Opposing Team’s Lose Point Level} + (\text{Own Team’s Innings} \times \text{Own Team’s Lose Point Level})}{(\text{Opposing Team’s Innings} + \text{Own Team’s Innings})} \right) + 2 \times 1.5 \]

(4) \[ \text{Replacement Player’s Expected Winnings} = 0.38 \]

The Starting Pitcher’s WAR = \[ \frac{(\text{Expected Winnings per 9 Innings} - \text{Expected Winnings per Replacement Player per 9 Innings}) \times \text{Pitched Innings}}{9} \]

(1), (2), (3) are the same example of the starting pitcher’s WAR

(4) Leverage index (LI) = 0.47

2.2. Operational definition of financial burden indicator (FBI)

In this study, the amount of money that was invested to produce WAR 1 in order to find out the efficiency of a player was compared to the annual salary. To this end, the total contract amount (down payment + annual salary) was calculated by dividing it by the total contract amount. Then, by dividing the value by WAR again, an indicator was created. For example, if a player named B signed a four-year contract on the condition of a deposit of 1 billion won and an annual salary of 500 million won, the total amount of the FA contract is 3 billion won. Recalculating this as the average annual amount is 750 million won (7.5). However, if a player named B has a WAR of 2.5 in a particular year, the player’s FBI is \((10 + 4 \times 5)/4/2.5\), so 3.0. In other words, it means that player B’s team has invested 300 million won to record player B’s WAR by one. Based on these results, the formula for calculating the efficiency index for the year was defined as:

\[ \text{FBI} = \frac{(\text{Deposit} + \text{Total Salary})}{\text{Total Contract Period} / \text{WAR of the year}} \]

2.3. Research limitations

The limitations of this study are as:

- First, age standards and annual standards do not necessarily match. Baseball players in Korea have a variety of stories, including high school graduation to professional entry, college graduation to professional entry, military service, and military service exemption from international competitions. Therefore, it may be the first year of the FA contract at the age of 30, or the first year of the FA contract at the age of 35. Therefore, age standards and annual standards were analyzed at the same time to complement each other.

- Second, the records of starting pitchers and relief pitchers were calculated. Unlike the MLB, the KBO has only 10 teams. Therefore, there are not many players. If the starting pitcher and the relief pitcher are separated, it is considered that the statistical error may be greater due to the specific player’s record and contract amount.

- Third, in this study, the study subjects were recorded from 2010 to 2019. This was considered to be appropriate for the study from 2010 by comparing the period when the FA market is rapidly increasing in amount and price in Korea. In the early days of the FA system, it was judged that the contract size was too small compared to the present time, which could lead to statistical errors.

2.4. Big data analysis process

As shown in Figure 1, the procedure for analyzing sports data in this study was conducted in four steps:

- data collection
- data pre-processing
- results analysis
- data visualization

![Figure 1. Model of sports big data analysis](image-url)
In the data collection stage, pitchers' ERA and WAR records were collected from the official Statiz website (http://www.statiz.co.kr), and press releases were collected during the FA contract from various news sources to sum up the total of players' deposits and salaries. These sums were collected and used for research. The collected data was pre-processed in a spreadsheet as shown in Figure 2. The contract was made for pitcher FA contracts between 2010 and 2019. In other words, they signed an FA contract after 2010, and are players who completed the FA contract period by 2019.

![Figure 2](image.png)

### 3. RESULTS AND DISCUSSION

In this study, the trends of ERA, WAR, and FBI were examined for pitchers who started and completed an FA contract between 2010 and 2019.

#### 3.1. Result analysis

In this study, the performance of FA players was analyzed based on age-standards and annual-standards. Age-standards were analyzed by dividing them into categories of 32 or younger, 33 to 34 years old, 35 to 36 years old, and 37 years or older. Annual-standards were analyzed by dividing players into the 1st, 2nd, 3rd and 4th years of their contracts.

#### 3.1.1. Result of average of innings pitched (IP)

Playing innings is the basic virtue of a pitcher in baseball. In the sports media of Korea, the expression 'Playing innings' is expressed as 'The pitcher eats innings'. The more innings the pitcher eats, the easier the other pitchers in the team can prepare for other matches. And if you finish the game with a small number of pitchers, it will be easier for the coaching staff to run a season because the relief pitchers will not be consumed. Therefore, in this study, because of the importance of these matters, we analyzed how many innings the FA pitchers could afford for each age group and how many innings they played each year. Thus, the higher the graph height, the more innings they played. As shown on the left side of Figure 3, FA pitchers averaged 91 2/3 innings in the year if their age was under 32. Additionally, 56 2/3 innings were played at the age of 33-34 years old, 51 2/3 innings at the age of 35-36 years old, and 43 innings on average at the age of 37 years or older. In addition, as shown on the right side of Figure 3, FA pitchers averaged 69 1/3 innings in
the first year of their contract, 65 1/3 innings were played in the second year, 56 innings were played in the third year, and in the fourth year, an average of 36 2/3 innings were played.

![IP of Age-Standard](image1)

![IP of Annual-Standard](image2)

Figure 3. (a) Average of IP according to age-standard, and (b) Annual-standard

3.1.2. Result of average of ERA

Equally important for pitchers to play innings is to cause the opposing team to score less. This is because scores are the determining factor for victory and defeat. The higher the height of the graph, the more points are scored. The lower the graph, the less points are scored. As shown on the left side of Figure 4, FA pitchers averaged an ERA of 4.68 under the age of 32. In addition, 5.44 was recorded at the age of 33-34, 5.11 was recorded at the age of 35-36, and an average of 5.88 was recorded in ERA over the age of 37 years old. Additionally, as shown on the right side of Figure 4, FA pitchers averaged an ERA of 4.97 in the first year of signing, 4.89 in the second year, 5.43 in the third year, and in the year of its fourth year, players averaged a recorded ERA of 6.26.

![ERA of Age-Standard](image3)

![ERA of Annual-Standard](image4)

Figure 4. (a) Average of ERA according to age-standard and (b) Annual-standard

3.1.3. Result of average of WAR

WAR is a record that efficiently expresses the performance of baseball players by applying statistics. This record occupies a large proportion in this study, and the higher the graph height, the higher the value that contributes to victory compared to alternative players. In addition, it was used as a basic data point for expressing the financial burden index (FBI) defined operationally in this study. As shown on the left side of Figure 5, FA pitchers averaged 1.86 WARs under the age of 32. A WAR record of 0.80 was recorded in the ages between 33 to 34 years old, 0.87 in ages between 35 to 36 years old, and an average of 0.46 WAR was recorded in the years after 37 year of age. In addition, as shown on the right side of Figure 5, FA pitchers averaged an ERA of 1.12 in the first year of signing. They recorded 1.28 in the second year, 0.88 in the third year, and in the fourth year, 0.31 WAR was recorded.
3.1.4. Result of average of FBI

In this study, an indicator called FBI was proposed using WAR. This is an indicator of how much burden the FA players have on the club's finances, which means that the higher the graph's height, the greater the burden on the club's finances. As shown on the left side of Figure 6, FA pitchers averaged an FBI of 7.82 when under the age of 32. In other words, on average, it is necessary to spend about 782 million won to record 1 WAR. In addition, 12.96 was recorded in the years of 33-34, 8.42 in the years of 35-36, and an average of 12.81 FBI was recorded in the 37 years old and later group. In addition, as shown on the right side of Figure 6, FA pitchers averaged an FBI of 7.44 in the first year of signing, as compared to 6.53 in the second year, 11.32 in the third year, and an FBI of 37.53 was recorded in the fourth year. This means that pitchers who spend four years in the FA contract will require an average of 3.75 billion won in order to record 1 WAR.

3.2. Discussion for analysis

In this study, pitchers with experience in FA contracts in the KBO league were analyzed, and based on this, suggestions are as:

- First, good results can be expected when a pitcher who is 32 years of age or younger is contracted with the FA contract. However, pitchers older than 37 were difficult to expect good results. However, the remarkable fact was that the performances of players aged between 35-36 was generally better than those of 33-34. For this reason, the KBO league showed that there were several pitchers whose scores had dropped significantly in the year they turned 33. And considering that players aged 35-36 are inferior in the number of innings, it can be said that players aged 33-34 and 35-36 were evaluated to be almost the same.
Second, considering the fact that first year players managed more innings, and second year players recorded a higher WAR, with a similar ERA, a comparison between first year contract players and second year players in the FA contract could be considered to be very different. There was not, however, a big difference. However, the decline started to appear by the third year, and the FBI increased significantly in the fourth year.

Third, through this study, it can be seen that until the age of 36, players show stable performance. Often, athletes in their mid-thirties have been told that they are going to suffer aging or to prepare for retirement. However, it can be said that through this study, athletes in their mid-30s also proved that the life of the athletes was longer than in the past. However, it can be seen that in the 4th-year of a FA contract, athletes also suffered a decrease in performance and increased the financial burden of the club. Therefore, it can be observed that it is much more reasonable for the club to set the contract period to about 3 years rather than 4 years. However, when signing an ace player, the contract period for 3 years may not be an attractive condition, so it is necessary to prepare a way to actively utilize the plus option.

4. CONCLUSION

Recently, with the global economic outlook worsening due to COVID-19, the clubs in the KBO league are characterized by relying on the support of their parent company. Therefore, if the contract size of the FA market expands too much, it can create a crisis that could lead to the KBO league's collapse. In addition, the number of clubs that want to make reasonable investments is increasing. Therefore, in this study, based on the data of pitchers who have experienced FA contracts in the KBO league, and through big data analysis, the players' performance and the financial burden of the team were identified. To this end, data on the amount of the contracts of players, contract duration, ERA and WAR was collected, and after preprocessing, data analysis was conducted. Based on these results, this study tried to find out how to improve the operational efficiency of the club by operating the financial burden index.

The expectation of this study is that the importance of data analysis in the field of professional sports can be further emphasized. It can also raise awareness that players in their mid-30s, who are starting to turn away, are maintaining good performance. In addition, through efforts to adjust the terms of the FA contract, the financial burden indicators can be adjusted to reduce the financial burden on the clubs and can be used as basic data for preventing the league from being destroyed.

Finally, as a future research plan, the analysis of players in other positions as well as the pitcher can be performed to further improve the performance of clubs and to continue to make basic data for professional baseball teams available to improve the quality of the KBO league through reasonable investment.

REFERENCES