Evaluation model revisited: Korean baseball Players Salary Calculation Model (PSCM)

Sung-Bae Roger Park[1*], Tae-Geun Kwon[2], & Jong-Hwan Jeon[2]

[1]*Faculty of Department of Global Sport Industry, Hanyang University, Republic of Korea*
[2]*Ph.D. Candidate of Department of Global Sport Industry, Hanyang University, Republic of Korea*

Abstract

The purpose of this study is three-fold. Firstly, it revisits and updates the prevailing model (Park et al., 2018), by adding players' field performance data from the past three seasons (2017 to 2019). Secondly, it verifies the effectiveness of the modified salary calculation model by comparing the top 20 ranked players' actual salaries with model Expected Monetary Value (EMV). Lastly, it elaborates and refines the model to enhance its applicability for the Korean market. The results indicate that all 10 Korea Baseball Organization (KBO) teams paid players much more than what was required. Also, according to the proposed salary calculation model, SK Wyverns was the most efficient team (i.e., smallest difference between total team salary and EMV) while Lotte Giants was the least efficient. Finally, expert panels suggest that the salary calculation model be used as a supplementary tool because the model does not count such factors as defensive position, leadership, contribution to winning, physical conditions, past injuries, and defensive skills.

Key words: salary calculation, expected monetary value, leadership, field performance data

Introduction

Determining professional baseball players' salaries has become a topic of great interest, not only academically but also practically. From labour economics perspectives, players' salaries are supposed to be determined solely based on their past athletic performances. However, in reality players' salaries are decided not only by their field performances, but also by payroll system (e.g., free agent or rookie), their

personalities and characteristics, reputation, popularity, years of experiences, leadership, defensive position, age, past injuries, and even agents' negotiation skills. This complexity and ambiguity of salary calculation causes continuous labour disputes between baseball teams and players. Sport mediators and arbitrators are often invited to solve and mitigate the labour disagreements between two parties. Various performance records and statistics (e.g., batting, pitching, defensive position, base running) make it more difficult to accurately and adequately measure the relevant market value of professional baseball players. More importantly, it is difficult to objectively calculate the salary of a professional baseball player because it is a prediction game rather

than a simple evaluation based on his past athletic performances.

In 1970s, statisticians and economists stepped into the sport industry to help translate players' field performance data into monetary value in the professional baseball market. Various studies have been conducted on the estimation and valuation of professional baseball players, both in the United States and in Korea (Krautmann, 1999; Oh & Lee, 2013; Rockerbie, 2010; Scully, 1974; Zimbalist, 1992). The most prominent formula is Bill James' Sabermetrics. The Sabermetrics model helped process player primary data (hits, RBIs, steals, etc.) and has proved to be very powerful for adequately assessing a player's actual value by calculating an individual player's contribution and impact on scoring and winning. Furthermore, several researchers agreed that evaluating the value of athletes from various angles such as converted Sabermetrics indices, rather than a single performance indicator, is appropriate to accurately assess a player's market value (Choi, 2009; Lee, 2014; Lee & Kim, 2016).

Park, Kwon, and Jeon (2018) developed a Korean baseball Player' Salary Calculation Model (PSCM) with use of 12 Sabermetric indices (i.e., OPS, GPA, SECA, TA, RC, RC/27, XR, ISO, PSN, wOBA, %OW, BABIP) based on 20 years of field performance records of Korean baseball players who played from 1997 to 2016. The PSCM was developed by meeting the statistical assumptions as follows. As a reminder from Park et al. (2018)'s study, Kaiser-Meyer-Olkin (KMO) and Bartlett tests were used to verify that the data used for principal component regression met the basic assumptions of factor analysis. The sample fit KMO value was acceptable at 0.73 and Bartlett's spherical test showed a significant probability less than .001, which was deemed suitable for factor analysis. In addition, the results of the Power Hitter Index (PHI) and the Contact Hitter Index (CHI), and the preliminary salary decision model showed that the coefficients were slightly changed.

The main purpose of this current study is three-fold. Firstly, it revisits the suggested PSCM (Park et al., 2018) and updates by adding the past three years of records from 2017 to 2019. Secondly, it verifies the effectiveness of the PSCM and its applicability to the baseball market. Lastly, it refines and elaborates the model through the panel study after identifying the potential issues and problems within the PSCM.

Methods

The original PSCM was developed by Park et al. (2018) using players' field performance and salary data from 1997 to 2016. In order to update and modify it, players' field performance records and salary data were excerpted from KBO Yearbooks from 2017 to 2019 (after the 2019 Korean Series games). The dataset has been stored in the MS Excel program and then imported into IBM SPSS v.23 for descriptive analyses and assumption checking such as multicollinearity. Then three equation models [Power Hitter Index (PHI), Contact Hitter Index (CHI), and Expected Monetary Value (EMV)] were regenerated using IBM SPSS Modeler v.17.

In order to test if the modified PSCM was effective and could be applicable to the current Korean baseball labour market, researchers compared the actual salary of top 20 ranked players among those who met the minimum number of at-bats in 2019 with EMV by the PSCM. In addition, the percentages of the annual salary of Free Agent (FA) players by the total salary per team were calculated in order to identify and rank the payroll efficiency.

To further refine and elaborate the model, the expert panel was established with eight members (three sports marketing professors, two former professional baseball team front-level staff members, two enthusiastic baseball fans for over 20 years, and one baseball-related business entrepreneur). They were provided with the actual salary of top 20 players with the EMV by the PSCM and asked to provide their professional ideas,

thoughts, and in-depth opinions. Also they were to evaluate the modified PSCM and the potential issues and problems that might exist within the model. The expert panels advised not only on the field performance data and salary amounts of players, but also on a variety of external variables that could influence the salaries of professional baseball players in Korea.

Results

Revisiting Player Salary Calculation Model (PSCM)

Park et al. (2018)'s PSCM was developed based on the players' field performance records and salary data from 1997 to 2016. In order to answer the first research purpose of this current study, the PSCM was updated and adequately modified by adding players' field performance records and salary data from 2017 to 2019 from Korea Baseball Organization (KBO) Yearbooks (see Table 1). The assumption of independence has been confirmed by Durbin-Watson test (see Park et al., 2018). As you can see in the Table 1, the players' age varied from 19 to 41 years old. The median age was 29, with

eight years of experiences in the KBO league.

As in Park et al. (2018)' study, the number of years of experiences with the KBO, Free Agent (Yes/No), Power Hitter Index (PHI), and Contact Hitter Index (CHI) were used as independent variables, while the Expected Monetary Value (EMV) was the dependente variable. The good explanatory power from the variables of the athlete's salary and the player's annual salary, FA (yes/no), PHI, and CHI has been confirmed, with $r^2 = .527$ and the adjusted $r^2 = 0.525$, $F(4, 933) = 259,83$, $p < .000$. (see Table 2)

Based on the validation model and statistical results, the standardized equations were generated as follows.

Power Hitter Index (PHI) $= 0.072 \times OPS + 0.032 \times GPA + 0.314 \times SECA + 0.012 \times \%\,OW + 0.455 * BABIP$

Contact Hitter Index (CHI) $= 0.079 \times OPS + 0.122 \times GPA - 0.201 \times SECA + 0.134 \times \%\,OW + 0.574 \times BABIP$

Expected Monetary Value (EMV) $=$ Annual $\times 971.3 +$ FA $\times 74460.7 +$ PHI $\times 7121.9 +$ CHI $\times 7921.9 + 6881.0$ (Unit: KRW 10,000)

Table 1. Descriptive statistics of the modified PSCM variables

	Min.	Max.	Average	Median	SD	Skewness	Kurtosis	Effective No.
Age	19	41	28.84	29	4.17	0.13	-.40	938
Years	1	22	8.19	8	3.945	0.44	-.20	938
Salary (₩ 10,000)	1162	375000	30036	12090	48057	3.42	14.46	938
FA (yes/no)	0	1	.21	0	.41	1.45	.10	938

Table 2. Validation model coefficients and statistical results

Variables	B	SE	β	t	p
Constant	6881.05	2770.29		2.48	.01
Year	971.32	360.25	.08**	2.7	.01
FA (Yes/No)	74460.7	3512.73	.62***	21.2	0
PHI	7121.89	1091.56	.15***	6.52	0
CHI	7921.93	1091.6	.16***	7.26	0

*** $p < .001$, ** $p < .01$, * $p < .05$

Top 20 Players Overpaid?

The second purpose of this current study was to apply the modified PSCM into a total of 273 players who met a minimum number of at-bats in 2019 season to verify the effectiveness of the PSCM model. The discrepancy table has been generated to compare and see the actual salary of top 20 players in 2019 and their EMV. As seen below (Table 3 & Figure 1), the PSCM model implied all 10 KBO clubs were more likely to overpay to the players than deserved. More specifically, SK Wyverns seemed to be the most efficient team by overpaying the smallest amounts to players than the PSCM suggested, compared to other teams (see Table 3 and Figure 1). More importantly, SK Wyverns could be the exemplary case to fit the model the best by accomplishing both successes of on-field athletic performance and off-field financial management,

Table 3. *Salary Effeciency Ranking*

(Unit : ₩ 10,000)

Ranking	Discrepancy Rank	Team Name	Model Suggesting Salaries (A)	Actual Salaries (B)	Discrepancy (A-B)
1	7	Doosan Bears	150306	548500	-398194
2	2	SK Wyverns	537669	718467	-180798
3	3	Kiwoom Heroes	130307	340800	-210493
4	9	LG Twins	344586	767367	-422781
5	8	NC Dinos	537318	942967	-405649
6	4	KT Wiz	379105	638767	-259662
7	6	KIA Tigers	371566	745200	-373634
8	5	Samsung Lions	302649	619200	-316551
9	1	Hanwha Eagles	421307	553967	-132659
10	10	Lotte Giants	343663	1066833	-723170
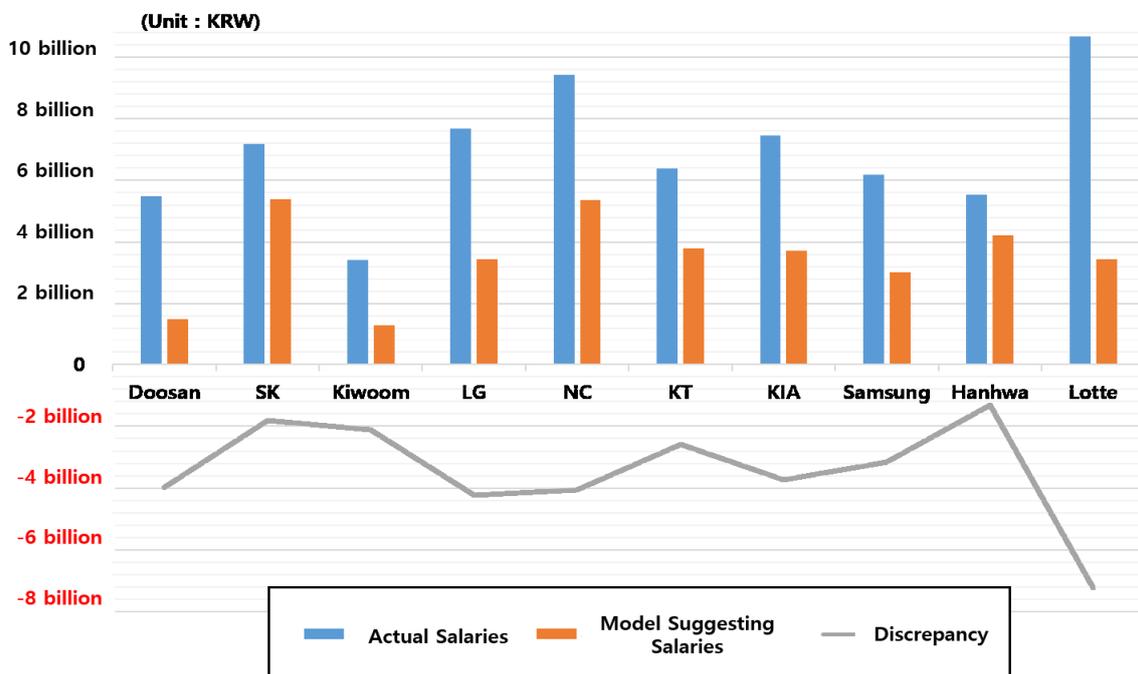
Figure 1. *Actual KBO Team Salaries and the Expected Monetary Value in 2019*

followed by Kiwoom Heroes, where it advanced to compete against the Korean Series with Doosan Bears in 2019. However, Lotte Giants ranked the bottom both on-field and off-field operations in 2019, followed by NC Dinos (see Table 3.). According to the salary efficiency ranking by discrepancy between actual salary and the EMV by the PSCM, Hanwha Eagles seemed to do the best job by overpaying the smallest amounts, approximately ₩1.3 billion. However, its on-field performance fell below their expectations in 2019.

As you can see in Figure 2, teams' overall salary is significantly affected by Free Agent (FA) players. This study verified the percentage of FA players by team's total salary in 2019. It is known that some teams rely heavily on FA players by paying enormous amounts of salaries. Lotte Giants and NC Dinos became top two teams with FA players' salary explaining over 70% and 55% of their overall payrolls, respectively. To make matters worse, FA players' performances in Lotte and NC failed to meet the teams' expectations. Meanwhile, it is worth noting that Doosan Bears won the Korean Series Championship title, but they did not much rely on FA players as figure 2 showed.

The researchers ranked top 20 players among those who met a minimum number of at-bats in 2019 based on the EMV by the PSCM (see Table 4 & Figure 3). The highest valued player was Jeong Choi, slugger of SK Wyverns, whose EMV was ₩1.07 billion while he received ₩1.76 billion. Eui-Ji Yang ranked the second and he received ₩3.13 billion that was about three times higher than the estimated EMV. However, the experts shared consensus that Eui-Ji Yang deserved to make much more salary than the EMV proposed, because he was the best catcher in the league. Byung-Ho Park of NC Dinos is not a free agent, but he has received ₩450 million more than the EMV. In addition, Lotte Giants' Dae-Ho Lee received ₩3.7 billion, but he was worth less than ₩1 billion in 2019. It implies that his underperformance contributed significantly to Lotte's bottom ranking in salary efficiency.

On the other hand, the peculiar fact is that SK Kang-Min Kim, who received an annual salary of ₩400 million, performed over ₩960 million. Calculating the market value of a player is a difficult task because salary has to be determined based on past performances. Even though a high-paid FA player performs less than expected, he gets what the contract says, even if deemed somewhat unreasonable. Therefore, the conditional
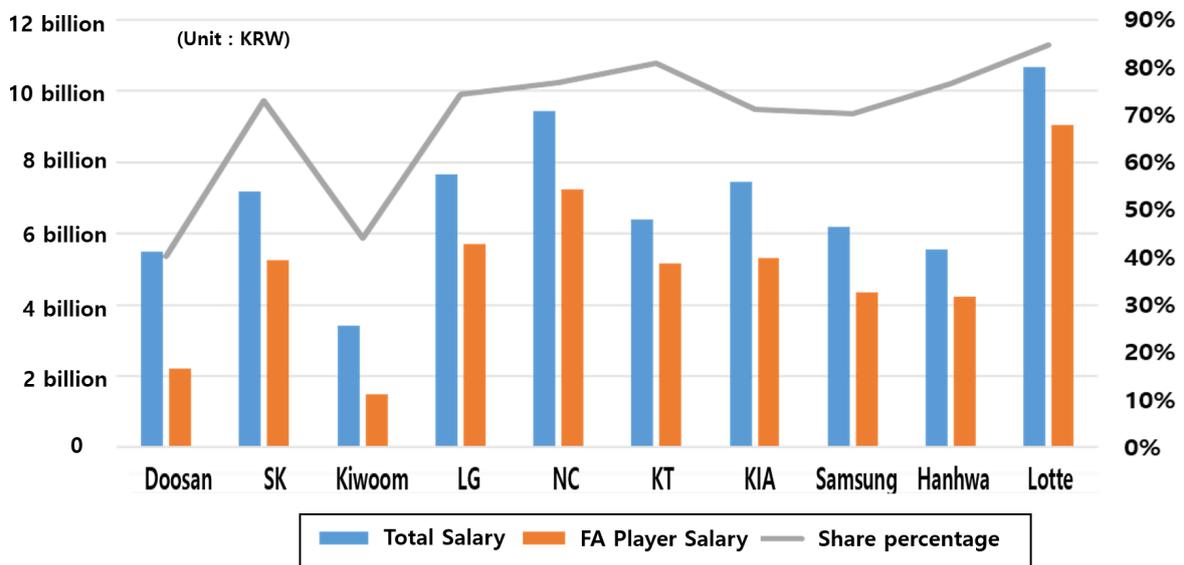

Figure 2. *Percentage of FA Player Salaries by Total Annual Team Payrolls in 2019*

contract system should be fairly usable by adjusting and modifying their field performances after the season is over. This type of payment structure will help increase salary efficiency by minimizing overpaid players and maximizing underpaid players.

Table 4. Top 20 Players by the EMV in 2019

(Unit : 0.1 billion KRW)

No.	Player	H	HR	RBI	AVG	OBP	SLG	Model Suggesting Salary	Actual Salary
1	Jeong Choi (SK)	147	29	99	.292	.399	.519	10.7	16.7
2	Eui-Ji Yang (NC)	138	20	68	.354	.438	.574	10.6	31.3
3	Byeong-Ho Park (Kiwoom)	121	33	98	.280	.399	.560	10.5	15
4	Hyeong-Woo Choi (KIA)	137	17	86	.300	.413	.485	10.4	25
5	Seok-Min Park (NC)	96	19	74	.267	.393	.478	10	24
6	Seong-Yeol Lee (Hanhwa)	106	21	85	.256	.343	.464	10	32
7	Tea-Gyun Kim (Hanwha)	132	6	62	.305	.382	.395	10	21
8	Han-Joon Yu (KT)	159	14	86	.317	.383	.443	9.8	15
9	Byung-Hun Min (Lotte)	112	9	43	.304	.380	.453	9.8	20
10	Jae-Gyun Hwang (KT)	127	20	67	.284	.358	.468	9.7	22
11	Dae-Ho Lee (Lotte)	138	16	88	.285	.355	.435	9.6	37.5
12	Kang-Min Kim (SK)	114	8	50	.270	.324	.370	9.6	4
13	Hyun-Soo Kim (LG)	160	11	82	.304	.370	.437	9.5	28.8
14	A-Seob Son (Lotte)	151	10	63	.295	.360	.400	9.5	24.5
15	Chang-Min Mo (NC)	104	10	55	.305	.358	.466	9.4	67
16	Joo-Chan Kim (KIA)	101	3	32	.300	.340	.374	9.4	9
17	Han-Yi Park (Samsung)	19	2	13	.257	.333	.378	9.4	25
18	Won-Seok Lee (Samsung)	97	19	76	.246	.327	.441	9.4	16.3
19	Eui-Yun Jeong (NC)	107	13	59	.282	.342	.447	9.4	73
20	Jae-Ho Kim (Doosan)	101	4	48	.268	.379	.358	9.4	12.5

H : Hit / HR : Homerun / RBI : Runs Batted In / AVG : Hitting Average /
OBP : On Base Percentage / SLG : Slugging Percentage
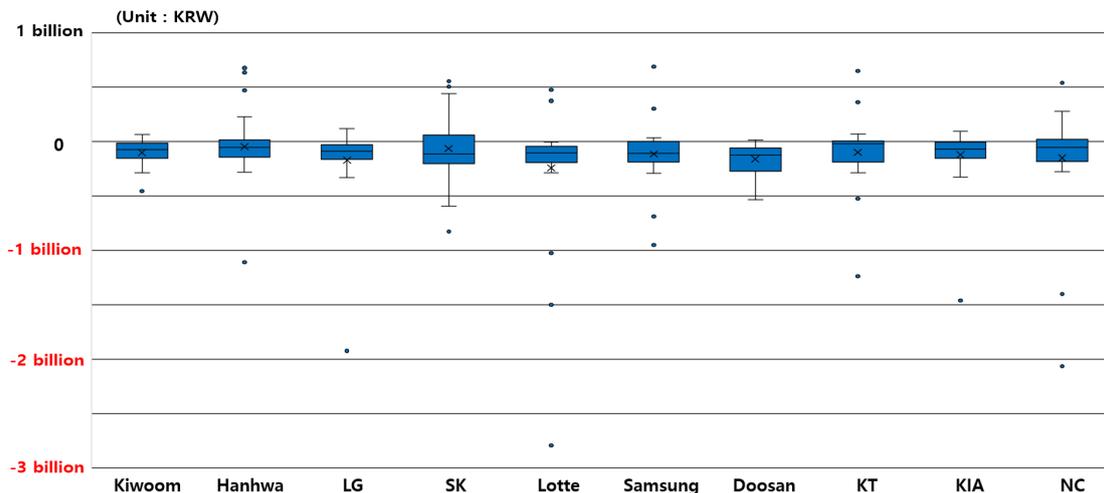

Figure 3. Salary Position of Free Agent Players

Consideration of David Ross Effect

The last research question was to verify the effectiveness of the salary calculation model and its potential applicability into the real baseball labour market, examining its accuracy and adequacy. It also revises and thus improves this current model by accepting expert panel opinions and advice that would better reflect the field of baseball labour market. Furthermore, expert panels were willing to provide the external variables affecting players' salary amounts, other than field performance data.

First of all, many panels point out that unbalanced supply and demand of professional baseball players in Korea caused and prompted the vicious circle of overpaying FA players, in recent years. A limited number of top-class FA players (e.g., Dae-Ho Lee, Hyung-Woo Choi, Suk-Min Park) entering the labour market received enormous spotlight from the media and

teams, which powerfully boosted their market value in the past few years. For instance, Hyung-Woo Choi (Kia Tigers) became the first person to break the invisible ₩10 billion (4 years) barrier in 2016, and Dae-Ho Lee was signed by Lotte Giants to the seemingly absurd figure of ₩15 billion (4 years). Also, the players' reputation, popularity, recognition rate, and major league experience were indicated as key factors to induce overestimation of players' value.

Expert panels also pointed out factors other than field performance records and statistics, including defensive position, leadership, contribution to winning, physical conditions, past injuries, and defensive skill (see Table 5). In fact, baseball specialists often assert that some position players (i.e., pitcher, catcher, short-stop, center-fielder) are very important to win a game and thus deserve additional consideration for higher salary, compared to counterpart players. More importantly, several panel members mentioned that players'

Table 5. Summary of Panel Study

Panel No.	Why they are overvalued?	External Variables to Consider	Suggestions
Panel #1	• Large amounts of down payment • Contribution to wins • Inaccurate calculation of field performance	• Game winning RBIs	Differential payment by years of experiences
Panel #2	• Unbalance of player supply and demand • Field performance-oriented team management	• Defense position • Base running skill • Clutch skill • Defense skill	Development of salary calculation indices to compare players in the same positions
Panel #3	• Reputation • Major League experience	• Leadership • Celebrity power • Physical condition	Inclusion of other factors than field performance
Panel #4	• Player awareness rate • Soaring player salary after cash trade	• Overall batting skill	Inclusion of defense skill
Panel #5	• Market supply and demand of players	• Defense position • Physical condition	Batting performance-weighted salary calculation
Panel #6	• Club's annual budget • FA market condition • Expected performance in the future	• Age • Celebrity power • Defense skill	Development of salary calculation indices to compare players in the same positions
Panel #7	• FA market condition • Existence of celebrity athlete • Judgement by team standing	• Defense skill • Years of experience • Positions	-
Panel #8	• FA Market condition • Major League experience	• Contributions to the winning • Celebrity power	-

personalities and characteristics such as leadership, teamwork, spirit of sacrifice truly matter to building a prestigious baseball club and thus should be considered as key factors to calculate players' salary. Brave, Butters, and Roberts (2017) argued that team chemistry proved to have a high influence on team performance by 44% while each individual player's baseball skill affected the team's performance by only 20%. This is called *David Ross effect* after David Ross, a former catcher and Chicago Cubs manager, whose personal leadership and teamwork experience contributed to winning the World Series of Baseball in 2016. Panels suggested that players' physical conditions and past injuries should also be considered in negotiating players' salaries. In summary, expert panels were not shy to conclude that the PSCM could be usable and applicable to the current baseball labour market in Korea, but only as a supplementary tool due to the fact that there might exist other internal and external factors to consider as mentioned previously. Furthermore, they recommended the PSCM model to be continuously elaborated and refined to reflect intangible characteristics of players, such as leadership, teamwork, and celebrity power.

Conclusions

This current study completed the first research objective by successfully updating the original PSCM developed by Park et al. (2018). Field performance records and salary data of players who met the minimum number of at-bats from 2017 to 2019 were inserted into the existing model. The three equations of PHI, CHI, and PSCM were generated with new coefficients of variables. The researchers also tried to verify the effectiveness and applicability of the PSCM into the market. The results showed that most KBO teams paid much more salary than the EMV estimated by the PSCM. It is reasonable to conclude that some players were overpaid because of absence of well-designed player salary calculation model. Also, this study

identified SK Wyverns to be the most effective and efficient team in terms of both on-field and off-field operation, while Lotte Giants was the worst. There was a strong correlation between the number of FA players and the overall team payrolls. If FA players do not perform well on the field, then the salary efficiency goes down, and vice versa. Lotte Giants and NC Dinos heavily relied on FA players, but their athletic performances seemed to be far below the expectations. As mentioned previously, the conditional contract system can be an option to increase salary efficiency by minimizing overpaid players and maximizing underpaid players after the season is over.

Determining the player's salary has never been an easy task because of various factors and intangible variables affecting the market value of players, such as leadership, teamwork, and spirit of sacrifice. The panel study stood on the same page with Brave, Butters, and Roberts (2017)'s research in a sense that the proposed PSCM has a genuine limitation by excluding team chemistry factor. Expert panels recommended that the model better be utilized as a supplementary tool mainly due to the unique nature of complex salary determination in the professional sport labour market. Furthermore, it was pointed out that the PSCM needs to be elaborated and refined by including the variables of defensive position, leadership, contribution to winning, physical conditions, past injuries, and defence skill into the current formula.

This study has several limitations. First of all, senior managers in KBO teams, who are deeply involved in salary determination and negotiation, declined to share their own valuation systems and know-hows, due to conflicts of interest with their scope of employment. Secondly, it was not easy to recruit the baseball expert panels, who must stand on the neutral line. Thirdly, this study only covers hitters, not pitchers that could be categorized as starting pitcher, relief pitcher, or closing pitcher.

In conclusion, it is believed that this study can contribute to the future development in the field of

player performance evaluation in Korea. Researchers who are interested in player valuation are recommended to focus on how non-field performance data and information (i.e., leadership, teamwork, reputation, physical conditions, MLB/NPB experiences, and celebrity power) can be inserted into the mathematics equation modelling.

Acknowledgments

References

Brave, S. A., Butters, R. A., & Roberts, K. (2017). In search of the holy grail: Team chemistry and where to find it. Retrieved from https://pdfs.semanticscholar.org/1f6d/ec8b0e1108ed647c6eaa485f600b63aacc9b.pdf

Choi, K. (2009). The introduction of Sabermetrics and Record Analysis in Korea. *Journal of Social Science*, **25(1)**, 129-139.

Krautmann, A. C. (1999). What's wrong with Scully-estimates of a player's marginal revenue product. *Economic inquiry*, **37**, 369-381.

Lee, J. (2014). Measurements for hitting ability in the Korean pro-baseball. *Journal of the Korean Data and Information Science Society*, **25(2)**, 349-356.

Lee, J., & Kim, H. (2016). Suggestion of batter ability index in Korea baseball-focusing on the sabermetrics statistics WAR. *Korean Journal of Applied Statistics*, **29(7)**, 1271-1281.

Oh, T., & Lee, Y. (2016). Value Evaluation Model for Korean Professional Baseball Players. *Journal of Labour Economics*, **39(2)**, 113-139.

Oh, T., & Lee, Y. (2013). Measurement of monopsony exploitation in salary determination; case of Major League Baseball. *Korean Journal of Sport Management*, **18(3)**, 1-15.

Park, S. R., Kwon, T., & Jeon, J. (2018). The development of a valuation model to determine the true market value of professional baseball players. *Korean Journal of Sport Science*, **29(3)**, 520-533.

Rockerbie, D. W. (2010). Marginal revenue product and salaries: Moneyball redux.

Scully, G. W. (1974). Pay and performance in major league baseball. *The American Economic Review*, **64(6)**, 915-930.

Zimbalist, A. S., & Storey, J. (1992). *Baseball and billions: A probing look inside the big business of our national pastime*. New York: BasicBooks.