Economics and Management
ISSN: 2683-1325
Volume: XVII, Issue: 2, Year: 2020, pp. 26-39

EFFECT OF PLAYER AND TEAM CHARACTERISTICS' EFFECTS ON PLAYERS' SALARIES: A STUDY OF STATISTICAL METHODS

Zhen-Jia-Liu[1]

Received: 15.05.2020 Accepted: 30.06.2020

Abstract

This study investigated factors that influence the remuneration of professional baseball players in Asian (Japan, Korea, and Taiwan) and US professional baseball leagues. The empirical results obtained by investigating 5289 baseball players as the study sample revealed that the support vector machine model was the most accurate in the Japanese and Korean leagues for predicting players' annual total remuneration, whereas the SVM model and the logit model were the most accurate in the Chinese (Taiwan) and US leagues, respectively.

Keywords: *baseball; salaries; players; logistic model; Support vector machines; Rough set theory*

JEL Codes: *C202, J01*

1. Introduction

Human capital and hedonic wage function theories are the major theories adopted to determine factors influencing remuneration. Human capital theory explains the influence of employees' individual characteristics, such as education (Renkas, 2013), age (Renkas, 2013), tenure (Barth, 1997), and training (Sun, 2012), on remuneration, and hedonic wage function theory explains the influence of work characteristics, such as industry (Sun, 2012), gender (Sun, 2012; Chzhen and Mumford, 2011; Ahmed and Mcgillivray, 2015), marital status (Chzhen and Mumford, 2011), party membership (Sun, 2012), number of family members (Sun, 2012), and number of young children in the household (Popli, 2012), on remuneration and reflects employee preferences for wages and work characteristics.

Remuneration structures in professional sports leagues vary by country because of the differences in their operational environments, such as the number of teams and fans. Athletes are paid a fixed remuneration along with bonuses and other incentives. Some players sign multiyear contracts (i.e., long-term contracts not signed on a yearly basis but

[1] Tan Kah Kee college, Xiamen University, China, E-mail:2240328588@qq.com

renewed upon expiry). Discussions of differences in baseball player remuneration in the literature have mainly focused on human capital, reporting seniority and experience (Stone and Pantuosco, 2008; Holmes,2011; Haupert and Murray, 2012; Hake and Turner, 2011; Hill and Jolly, 2017), contract length (Pedace and Hall, 2012), population of the team's metropolitan statistical area (Holmes, 2011), nationality (Pedace and Hall, 2012), team performance (Wang and Cheng and Jane), total team revenue (Holmes, 2011), player performance (Haupert and Murray, 2012), and player age (Hake and Turner, 2011) as factors influencing athlete remuneration. However, these studies only applied human capital theory or other theory and did not conduct diagnostic tests, including those for determining the accuracy of the applied model.

Research has demonstrated that logistic regression, support vector machine (SVM), and rough set theory (RST) models are applicable to multifarious practical problems pertaining to economic and financial result prediction; however, this research has not applied them to athletes' remuneration. In this study, we adopted the "total sum" (i.e., fixed remuneration and bonus) to measure remuneration; developed logistic regression, support vector machine (SVM), and rough set theory (RST) models; and evaluated the accuracy of the developed models in predicting remuneration in professional baseball leagues in Asia (Japan, Korea, and Taiwan) and the United States

2. Literature Review

2.1. Player remuneration

In sporting competitions, coaches and trainers are crucial for individual and team performances (Wester and Weiss, 1991). They influence player styles and techniques and are involved in player role distribution, practice scheduling, on-site strategy, player dispatching, and other related decisions. Therefore, coach characteristics affect player performance and thus player remuneration (Kahn, 1993b). Idson and Kahane (2000) argued that the seniority of professional sports team coaches is positively related to player remuneration. Professional sports coaches are extensively involved in decision making, such as team composition and analyzing opponent strategy. Their experience confers them with strong leadership skills, which enhance team performance and positively contribute to player performance, thus leading to higher player remuneration. For success in team sports, athletes not only rely on individual prowess but also on team strategy. Idson and Kahane (2000, 2004) have asserted that team performance influences managerial assessment (e.g., remuneration and incentives) of players.

Krautmann (2017) reported that risk-averse owners pay a premium in salary bids for free agents in MLB because free agents' contract terms are negatively related to the degree of variability in their performance; this suggests a heretofore unrecognized factor affecting the market for talent in professional sports. Liu and Zong and Wang and Zhang (2017) reported factors that determine salary among professional baseball players in the Asian (Japan, Korea and Taiwan) and US. Empirical results show that US baseball players are the highest paid and those in Taiwan receive the lowest salary, with a large difference of 75 times. Experience, age, education and other variables, as well as population of teams' home cities are influential. Whether or not a player transfers to other teams, their health conditions (measured as BMI) and other variables significantly affect salary, and there is no significant variance arising from different management environments in different countries. Although team age and player salary are correlated significantly in the three countries, the correlation is positive in the US and negative in Asian (Japan, Korea and Taiwan). However, their health conditions (measured as BMI), training, education and population of teams' home cities have insignificantly correlation with the total salary from multi-year contracts.

2.2 Predicting economic and financial results by using the logistic model, support vector machines, or rough set theory

The logistic model, SVMs, or RST can be used to predict financial results such as firm performance (Zhang and Yang and Li, 2015), corporate failures (Kim and Mun and Bae, 2018), fraudulent financial statements (Yeh and Chi and Lin, 2016). For example, Zhang and Yang and Li (2015) employed an SVM to predict the profitability of the construction companies listed on the A-share market in China. Additionally, Kim and Mun and Bae (2018) used an SVM to predict corporate bankruptcy in Korea and demonstrated that the proposed method offers greater accuracy for corporate bankruptcy prediction than do existing methods. By contrast, Yeh and Chi and Lin (2016) integrated the RST and SVM approaches and employed both financial and nonfinancial ratios for detecting fraudulent financial statements.

Analogous economic results may be predicted using the logistic model, SVMs, or RST, such as stock market or price index (Cheng and Yang, 2018), financial asset returns (Taylor and Yu, 2016), futures index (Das and Padhy, 2017), economic resource price or consumption (Kaytez and Taplamacioglu and Cam, 2015; Yu and Zhang and Wang, 2017; Wu and Shen, 2018). For example, Cheng and Yang (2018) developed a novel fuzzy time-series model based on rough set rule induction for forecasting stock indexes, with TAIEX, Nikkei, and Hang Seng Index stock prices forming the experimental dataset. By contrast, Taylor and Yu (2016) provided an empirical illustration by using daily stock index data and presented a new autoregressive logistic model for forecasting the probability of a time series of financial asset returns. Das and Padhy (2017) combined an unsupervised extreme learning machine with SVM-based clustering and support vector regression (called USELM- SVR) to forecast an energy commodity futures index on the Indian multi-commodity exchange. Furthermore, Kaytez and Taplamacioglu and Cam (2015) employed SVMs and LSSVMs as new techniques for energy consumption forecasting. Yu and Zhang and Wang (2017) demonstrated that an SVM performs effectively in forecasting, implying that it is a favorable candidate for crude oil price forecasting. Wu and Shen (2018) used an LSSVM model based on gray relational analysis to predict energy demand.

3. Methodology

3.1 Data sources and sampling

Comprehensive data are unavailable for professional athletics; therefore, we ensured data reliability by obtaining data directly from the investigated professional baseball leagues[1]. In addition, data were obtained from websites[2] and sports media[3]. Baseball players who have played at least one game in the top division in professional baseball leagues in Japan, Korea, Taiwan, and the United States between 2009 and 2017 were included in the study. Per-season contracts were investigated. Data on 5289 players were obtained.

Remuneration is generally of two types: direct (basic remuneration, stipends, incentives, and living expenses) and indirect (health insurance, vacation benefits, welfare, and services). Professional baseball players define their remuneration structure as consisting of remuneration and bonuses. However, many professional baseball leagues do

---

[1] http://mlb.mlb.com(MLB, American Professional Baseball League)
http://www.npb.or.jp(Nippon Professional Baseball League)
http://www.koreabaseball.com(Korea Professional Baseball League)
http://www.cpbl.com.tw(Chinese Professional Baseball League)

[2] http://home.a07.itscom.net/kazoo/pro/pro.htm (Salaries of Japanese Baseball Players)
http://content.usatoday.com/sports/baseball/salaries/default.aspx(Salaries of American baseball Players)
http://mlbcontracts.blogspot.com(contracts of American baseball Players)
http://twbsball.dils.tku.edu.tw/wiki/index.php?title=%E9%A6%96%E9%A0%81(Wikiped ia: Baseball)

[3] Data on the remuneration of professional baseball players are not public in Taiwan and Korea. The figures used in this study were extrapolated from media reports (e.g., the Baseball Match Handbook published by the Chinese and Korean professional baseball leagues).

not reveal the additional incentives offered. In this study, total annual remuneration[1]is defined as the remuneration (in US dollars) received by the sampled athletes during the 2009–2017 seasons. Accordingly, two levels of remuneration, 1 and 0, are defined as the output labels, representing annual remuneration higher and lower than the average remuneration of players in the same league at year t, respectively. In addition, the remuneration equations for professional baseball players are measured using player team transfer[2], players' degree of obesity[3], the length of training[4],the population of the city in which the team is located[5][6] (Liu and Zong and Wang and Zhang ,2017).

3.2 Model

The paper adopts logistic regression model, support vector machines (SVM) and rough sets theory (RST) to estimate parameters.

(1) Logistic regression model

$$\lambda=\ln\frac{p}{1-p}=\beta_0+\beta_1x_1+\cdots\ldots\beta_ix_i \text{ ………………………………(1)}$$

According to the definition of logistic function

$$p=\exp(\alpha+\beta x)/[1+\exp(\alpha+\beta x)] \text{ ……………………….. (2)}$$

$$1\text{-}P=1/[1+\exp(\alpha+\beta x)] \text{ ………………………………. (3)}$$

α, β_1 ……β_m are return parameters in the model

When the dependent variable is 0, 1 variable, the results are in two situations of occurrence (the dependent variable is 1) or non-occurrence (the dependent variable is 0). The model expressions are as follows:

$$P(Y=1)=\frac{\exp(\lambda)}{1+\exp(\lambda)} \text{ …………………………………………………… (4)}$$

---

[1] Calculated as the sum of all payments guaranteed in the contract.

[2] It is a dummy variable which is equal to one if the player i at time t belonged to a new team after the previous season otherwise, is equal to zero

[3] body-mass index is the health of player i at time t measured in weight divided by height (in meters)

[4] It is the player's lagged degree of on-the-job training measured as the number of games through division 2 or minor league

[5] It is the population of the city in which the team is located at time t measured as the natural logarithm of the population of the team's home city.

[6] Statistics released by the government in a city of each team and estimated based on information obtained from Wikipedia (English version).

$$P(Y=0) = \frac{\exp(\lambda)}{1+\exp(\lambda)} \quad \text{...} \quad (5)$$

Equation (4) and (5) show that P(Y=1) =1- P(Y=0)

(2) Support vector machines (SVM)

Support vector machines are a set of related supervised learning methods used for classification and regression. Viewing input data as two sets of vectors (two classes classification) in an high dimensional transformed space, an SVM seeks to construct a separating hyper-plane in that space, one which maximizes the margin between the two data sets. To calculate the margin, two parallel hyper-planes are constructed, one on each side of the separating hyper-plane, which are "pushed up against" the two data sets. Intuitively, a good separation is achieved by the hyper plane that has the largest distance to the neighboring data points of both classes, since in general the larger the margin the better the generalization error of the classifier. That is, based on the structured risk minimization principle, SVMs seek to minimize an upper bound of the generalization error instead of the empirical error as in neural networks.

$$y = \text{sign}\,(w^T \emptyset(X) + b),\ y \in \{-1.1\} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (6)$$

where y is output (1 for type A , -1 for type B) ; $\emptyset(X)$ is a nonlinear mapping form the input space to the high dimensional transformed space. SVMs exploit the idea of mapping input data into a high dimensional reproducing kernel Hilbert space (RKHS) where classification could be easily performed. Coefficients W and b are estimated by the following optimization problem

$$\min_{w,b} R(w, \xi) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m} \xi_i \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (7)$$

$$\text{with } y_i\,(w^T \emptyset(X) + b) \geq 1+\xi_i\ ,\ i=1 \ldots\ldots\ldots\ldots m \quad \xi_i \geq 0$$

c is a prescribed parameter to evaluates the trade-off between the empirical risk and the smoothness of the model.

(3) Rough set (Yeh et al.,2010)

Rough sets theory (RST) is a machine-learning method has proved to be a powerful tool for uncertainty and has been applied to data reduction, rule extraction, data mining and granularity computation. Here, we illustrate only the relevant basic ideas of RST that are relevant to the present work. By an information system we understand the 4-tuple S=(U,A,V, f), where U is a finite set of objects, called the universe, A is a finite set of attributes , V= $\mathrm{U}_{a \in A} V_a$ is a domain of attribute a , and $f : U \times A \rightarrow V$ is called an information function such that $f(x, a) \in v_a$, for $\forall a \in A$, $\forall x \in U$. In the classification problems, an information system is also seen as a decision table assuming that $A = C \cup D$ and $C \cap D = \emptyset$, where C a set of condition attributes and D is a set of decision attributes

Let S= (U, A, V,f) be an information system, every P ⊆A generates an in-discernibility relation IND (P) on U, which is defined as follows:

IND (P) = $\{(x, y) \in U \times U : f(y,a), \forall a \in P\}$…………………………(8)

U/IND (P) = $\{C_1, C_2 \ldots\ldots\ldots . C_K\}$ is a partition of U by P, every C_i is an equivalence class. For $\forall x \in U$ the equivalence class of x in relation to U/IND (P) is defined as follows:

$[x]_{U/IND(P)}= \{y \in U : f(y,a) = f(x,a), \forall a \in P\}$…………………………… (9)

Let P ⊆A and X ⊆U. The P-lower approximation of x (denoted by $P_*(x)$) and the P-upper approximation of x (denoted by $P^*(x)$) are defined as follows:

$P_*(x)= \{ y \in U : [y]_{U/IND(P) \subseteq X}\}$

$P^*(x)= \{ y \in U : [Y]_{U/IND(P) \cap X} \neq \emptyset\}$…………………………………… (10)

where $P^*(x)$ is the set of all objects form U which can certainly be classified as elements of x employing the set of attributes P. $P^*(x)$ is the set of objects of U which can be classified as elements of x using the set of attributes P. Let P, Q⊆ A, the positive region of classification U/IND (Q) with respect to the set of attributes P, or in

Short, P-positive region of Q is defined as POS (Q) $=U_{X \in U/IND(Q)}\ p(X)$.

$POS_p(Q)$ contains objects in U that can be classified to one class of the classification$U/IND(Q)$ by attributes P. The dependency of Q on P is defined as

$\gamma_p(Q)$= card ($POS_p(Q)$)/ card(U)……………………………………..(11)

An attribute a is said to be dispensable in P with respect to

Q, if $\gamma_p(Q)=\lambda_P - \{a\}(Q)$; otherwise a is an indispensable

attribute in P with respect to Q. ⊂

Let S= (U, A, V,f) be a decision table, the set of attributes

P(P ⊆C) is a reduce of attributes C, which satisfies the following

Conditions:

$\gamma_p(D) = \gamma_c(D), \gamma_p(D) \neq \gamma_{P'}(D)\ \ \forall P' \subset P$. ……………………..(12)

A reduce of condition attributes C is a subset that can discern decision classes with the same Accuracy as C, and none of the attributes in the reduced can be eliminated without decreasing its distrainable capability

A reduction of condition attributes C is a subset that can discern decision classes with the same accuracy as C, and none of the attributes in the reduced attributes can be eliminated without reducing its distrainable capability (Pawlak,2002).

3.3. Confusion matrix and type I, II errors of predicting models

True positive (TP) and true negative (TN) are correct classifications. A false negative (FN) occurs when the outcome is incorrectly predicted as negative when it is actually positive. a FN, also called a Type I error, occurs when a null hypothesis is rejected when it is actually true. A false positive (FP) occurs when the outcome is incorrectly predicted as positive when it is actually negative. A FP, also called a Type II error, occurs when a null hypothesis is accepted when it should have been rejected.

3.4. Robustness Test

In order to avoid possible bias from extreme values, the study also adopt those samples only include the sample data of from the 5th percentile to the 95th percentile as measures for the robustness test (Huang & Liu, 2011)

4. Results

4.1. Descriptive Statistics

The summary statistics of the player' salary among nations (for convenience of analysis, salaries' currency conversion is based on the exchange rate in 2009-2017 (Table1). The rate was obtained though www.oanda.com, those from the US receive the highest total salary, followed by Taiwan with the lowest. Based on country, overall salary level can be compared, such that US players receive 7.55 times that of Japanese professional baseball players, while the Japanese receive 1.76 times the salary of their Korean counterparts and the Korean receive 5.37 times the salary of their Taiwanese counterparts. The Americans earn 71.46 times of what Taiwanese professional baseball players receive. This large disparity in salary perhaps explains growing ambition among Asian players to find employment in the US. Comparing average salaries for 2009-2017, there is significant difference in the Asian and US Professional Baseball Leagues, reflecting perhaps that salary levels among professional baseball players have changed with the overall management environment.

It shows that the characteristics of baseball players or team from different countries vary. Of them, in terms of players' obesity (BMI)[1], baseball players in the US average as obese. The Japanese and Korean tend to be overweight. Taiwanese baseball players have standard BMI. Taiwanese baseball players have standard BMI. US players show lower training, reflecting the fact that American professional baseball players enjoy longer professional life. A possible reason maybe because US players have better physical conditions than their Asian counterparts. Another reason is that major league affords better

[1]BMI equal to or larger than 27 is considered obese, "overweight" is 24 - 27, and standard is 18.5 - 24.

training that helps maintain stamina muscle power and playing skills. Taiwan players have higher training. Reasons maybe that Taiwan baseball players suffer from sports injuries due to too-frequent dispatching, resulting in a shorter professional life.

Table 1. Descriptive statistics (average, annual salary, in US ten thousand dollars).

	U.S	JAPAN	Taiwan	Korea
TR_{it}	428.05	56.72	5.99	32.17
BMI_{it}	28.92	25.82	22.12	26.19
$TRAINING_{it-1}$	15.2	19.6	28.75	24.45

where TR_{it} is the sum of all payments guaranteed in the contract;. BMI_{it} is the health of player i at time t measured in weight divided by height (in meters); $TRAINING_{it-1}$ is the player's lagged degree of on-the-job training measured as the number of games through division 2 or minor league.

4.2. Empirical Test

Comparisons of predicted and actual classifications are shown in Tables 2–5. As indicated in Table 2 (Japanese professional baseball league), the SVM model had the highest value (the accuracy was 61.32%), and the RST model had the lowest value (the accuracy was 60.88%). As indicated in Table 3 (Korean professional baseball league), the SVM model had the highest value (the accuracy was 52.18%), and the RST model had the lowest value (the accuracy was 51.095). As indicated in Table 4 (Chinese professional baseball league, Taiwan), the SVM model had the highest value (the accuracy was 61.09%%), and the logit model had the lowest value (the accuracy was 53.68%).As indicated in Table 5 (Major League Baseball, United States), the logit model had the highest value (the accuracy was 59.78%), and the RST model had the lowest value (the accuracy was 49.78%). Furthermore, the empirical results show that these models exhibited higher accuracy (>50%). The significance of the difference provides strong evidence of the most accurate trend predictions regarding athlete remuneration in professional baseball leagues.

Table 2 The accuracy of every prediction model: Japan (N=1189)

	Logit	SVM	RST
Overall Correct Rate	62.28%	61.32%	60.88%
Overall Incorrect Rate	37.72%	38.68%	39.12%

Table 3 The accuracy of every prediction model: Korea (N=833)

	Logit	SVM	RST
Overall Correct Rate	52.22%	52.18%	51.09%
Overall Incorrect Rate	47.78%	47.82%	48.91%

Table 4 The accuracy of every prediction model: Taiwan (N=435)

	Logit	SVM	RST
Overall Correct Rate	53.68%	61.09%	56.11%
Overall Incorrect Rate	46.32%	38.91%	43.89%

Table 5 The accuracy of every prediction model: American (N=2832)

	Logit	SVM	RST
Overall Correct Rate	59.78%	50.08%	49.78%
Overall Incorrect Rate	40.22%	49.92%	50.22%

4.3. Robustness Test

We repeat the same analyses to tackle a possible sample specific issue and get general robust results (i.e., the study adopt those samples only include the sample data of from the 5th percentile to the 95th percentile as measures for the robustness test). Comparisons of predicted and actual classifications are shown in Tables 6–9.

As indicated in Table 6 (Japanese professional baseball league), the logit model had the highest value (the accuracy was 62.90%), and the RST model had the lowest value (the accuracy was 61.48%). As indicated in Table 7 (Korean professional baseball league), the logit model had the highest value (the accuracy was 53.26%), and the RST model had the lowest value (the accuracy was 52.11%). As indicated in Table 8 (Chinese professional baseball league, Taiwan), the SVM model had the highest value (the accuracy was 62.61%), and the logit model had the lowest value (the accuracy was 55.02%). As indicated in Table 9 (Major League Baseball, United States), the logit model had the highest value (the accuracy was 60.85%), and the RST model had the lowest value (the accuracy was 50.67%). Overall, analysis of the prediction model through robustness testing indicated that the significant difference provides strong evidence of more accurate trend predictions for athlete remuneration in professional baseball leagues. Additionally, the sample revealed that the empirical results still show that these models exhibited higher accuracy (>50%)

Table 6 The accuracy of every prediction model: Japan (N=1070)

	Logit	SVM	RST

Overall Correct Rate	62.90%	61.93%	61.48%
Overall Incorrect Rate	37.10%	38.07%	38.52%

Table 7 The accuracy of every prediction model: Korea (N=749) ((Robustness test)

	Logit	SVM	RST
Overall Correct Rate	53.26%	53.22%	52.11%
Overall Incorrect Rate	46.74%	46.78%	47.89%

Table 8 The accuracy of every prediction model: Taiwan (N=392)

	Logit	SVM	RST
Overall Correct Rate	55.02%	62.61%	57.51%
Overall Incorrect Rate	44.98%	37.39%	42.49%

Table 9 The accuracy of every prediction model: American (N=2549)

	Logit	SVM	RST
Overall Correct Rate	60.85%	50.98%	50.67%
Overall Incorrect Rate	39.15%	49.02%	49.33%

5. Discussion

This study examined salary structures among professional baseball players in Asia (Korea, Japan and Taiwan) and the United States. The results indicate that the SVM model predicted annual player salary in the Japanese and Korean professional baseball leagues the most accurately; the SVM model predicted annual player salary in the Chinese professional baseball league (Taiwan) the most accurately; and the Logit model predicted annual player salary in Major League Baseball (American) the most accurately.

The remuneration contracts of baseball players differ from those in other industries. In the United States and Japan, baseball player remuneration contracts are primarily multiyear contracts. In Taiwan, multiyear contracts were not used until 2006. Multiyear and annual contracts have distinct advantages and drawbacks. Although multiyear contracts guarantee exclusive access to the player, they financially burden the team sponsors because of the guaranteed payment regardless of player performance. Moreover, the team must bear the negative impact (cost efficiency) of a player being injured or unavailable because of unexpected events during the contract validity period. Therefore, seeking the optimal remuneration contract that benefits both the team management and the player is critical. However, few studies have focused on these two types of contracts, and the underlying theory is weak.

In summary, this study has the following implications. Although the empirical results of this study can serve as a reference for future academic research and team administration, the characteristics of our data are relatively restrictive. For example, data on player remuneration in Asian baseball leagues are incomplete, thus necessitating corroboration through further research. Nevertheless, remuneration data in professional sports are more easily accessible than those in other industries, but some variables, such as bonuses and incentives, are not easily accessible. Including such data in our analysis may have yielded clearer results. In addition, we developed cross-national models for forecasting remuneration structures for professional baseball players, which exhibited accurate outcomes with high goodness-of-fit. These models and results can serve as a reference to regulators and policy makers; however, our models are subjective, and an optimal model should be determined in the future.

REFERENCES

Ahmed, S., & Mcgillivray, M. (2015). Human capital, discrimination, and the gender wage gap in Bangladesh. World Development, 67,506-524.

Barth, E. (1997). Firm-specific seniority and wages. *Journal of Labor Economics,* 15(3), 495-506.

Chzhen, Y., & Mumford, K. (2011). Gender gaps across the earnings distribution for full-time employees in Britain: Allowing for sample selection. *Labour Economics*, 18(6), 837-844.

Das, S.P.,& Padhy, S.(2017). Unsupervised extreme learning machine and support vector regression hybrid model for predicting energy commodity futures index. Memetic Computing,9(4):333-346.

Haupert, M., & Murray, J. (2012). Regime switching and wages in major league baseball under the reserve clause. *Cliometrica,* 6(2),143-162.

Hill, J.R.,& Jolly, N.A .(2017). Revenue Sharing and Player Salaries in Major League Baseball. Journal of Sports Economics,18(8): 831-849.

Holmes, P. (2011). New evidence of salary discrimination in major league baseball. *Labour Economics*, 18(3),320-331.

Huang, D.T. & Liu, Z.C. (2011). The relationships among governance and earnings management: An empirical study on non-profit hospitals in Taiwan", *African Journal of Business Management*, vol.5,no.14,pp.5468-5476.

Idosn, T.L., & Kahane, L.H. (2000). Team effects on compensation: An application to salary determinants in the National Hockey League. *Economic Inquiry,* 38(2), 345-357.

Idosn, T., & Kahane, L. (2004). Teammate effects on pay. *Applied Economics Letters*, 11(12),731-733.

Kahn, L.M. (1993b). Managerial quality, team success, and individual player performance in major league baseball. *Industrial and Labor Relations Review*, 46(3): 531-547.

Kaytez, F.,& Taplamacioglu,M. C.,&; Cam, E.(2015). Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squaressupport vector machines. Interational Journal of electrical power & energy systems, 67:431-438.

Kim, S..,& Mun, B.M.,& Bae, S.J. (2018). Data depth based support vector machines for predicting corporate bankruptcy. Applied Intelligence,48(3):791-804.

Krautmann, A. C. (2017). Risk-Averse Team Owners and Players' Salaries in Major League Baseball. Journal of Sports Economics,18(1):19-33.

Liu,Z.J., & Zong ,B.B.,& Wang, Y.S.,& Zhang, X. Wei.(2017). The relationship between player characteristics, team characteristics and athletes' remuneration: An advanced empirical study of Asian and US professional baseball leagues. Modern Economy. 8(7):865-877

Pedace, R., & Hall, C.M. (2012). Home safe: no trade clauses and player salaries in major leagues baseball. *Industrial Relations: A Journal of Economy and Society*, 51(3), 627-644.

Popli, G. K. (2012). *Gender wage differentials in Mexico: A distributional approach. Journal of the Royal Statistical Society*: Series A (Statistics in Society). Blackwell Publishing Ltd.

Renkas, J. (2013). Wage expectations in light of human capital measurement theory. *Argumenta Oeconomica Cracoviensia,* 9, 29-42.

Stone, G., & Pantuosco, L. J. (2008). Estimating baseball salary equations from 1961-2005: A look at changes in major league compensation. International Journal of Sport Finance, 3*(*4), 228-238.

Sun, H. (2012). A study of determination in Chinese peasant workers' wage: The Role of human capital and market segmentation. *International Journal of Business Administration*, 3(4), 31-36.

Taylor, J.W., & Yu, K.(2016). Using auto-regressive logit models to forecast the exceedance probability for financial risk management. Journal of The Royal statistical society series A-statistics in society,179(4):1069-1092.

Wang, J.S., Cheng, C.F., & Jane, W.J. (2014). Buying success or redistributing payment: bidirectional causality in Korean Professional Baseball League. *Eurasian Business Review*, 4,247-260.

Wester, K.R., & Weiss, M.R. (1991). The relationship between perceived coaching behavior and group cohesion in high school football teams. *Sport Psychologist*, 5, 41-54.

Wu, Y.H.,& Shen, H.(2018). Grey-related least squares support vector machine optimization model and its application in predicting natural gas consumption demand.Journal of Computational and applied mathematics, 338:212-220.

Yeh, C.C.,& Chi, D.J.,& Lin, T.Y.(2016). A Hybrid Detecting Fraudulent Financial Statements Model Using Rough Set Theory and Support Vector Machines. Cybernetics and systems,47(4):261-276.

Yeh,C.C., & Chi, D.J., & Hsu, M.F..(2010). A hybrid approach of DEA, rough set and support vector machines for business failure prediction. Expert Systems with Applications. 37(2), 1535-1541.

Yu, L.,& Zhang, X.,& Wang, S.(2017). Assessing Potentiality of Support Vector Machine Method in Crude Oil Price Forecasting. Eurasia Journal of mathematics Science and technology education, 13(12):7893-7904.

Zhang, H.,& Yang, F.,& Li, Y.(2015). Predicting profitability of listed construction companies based on principal component analysis and support vector machine-Evidence from China. Automation in consturction, 53:22-28.