Analysis of Pitching Skills of Major League Baseball Players

Michiko Miyamoto(✉) and Akihiro Ito

Department of Management Science and Engineering,
Akita Prefectural University, 84-4 Aza Ebinokuchi Tsuchiya,
Yurihonjo 015-0055, Japan
{miyamoto,B16D003}@akita-pu.ac.jp

Abstract. This study examines the pitcher's deciding ball after pushing a batter with two strikes of an aged pitcher group (31 to 43 year-old) and a younger pitcher group (20 to 30 year-old) by using an actual tracking data of the Major League Baseball in 2015. The regression analyses are conducted for all pitchers and for each age group on different pitch types; *i.e.*, two-seam, cutter, splitter, forkball, straight and so on. We also analyze relationships between pitchers' knocking out batters and their pitching characteristics measured by pitch movements by using a framework and empirical analyses. The results of the research model using Structural Equation Modeling show what makes the pitcher advantageous over the batter.

Keywords: Major league baseball · Tracking data · Regression analyses · Structural equation modeling

1 Introduction

Baseball games always focus on players' actual strengths and accomplishments. Winning a game is certainly related to a player's individual performance. The performance of the pitcher, particularly, becomes crucial to the outcome of a game [1–4], and pitcher performance projection is a fundamental area in baseball analysis [5]. Karakolis [6] states that pitchers are evaluated by their abilities, performances and contributions.

Best players in American (AL) and National League (NL) of Major Baseball League are young. For example, Bryce Harper of NL is twenty-three-year-old, becoming the third-youngest player to win the Baseball Writers' Association of America's National League Most Valuable Player Award in 2015, and is expected to be like Frank Robinson, who became the only player to win league MVP honors in both NL and AL, as well as winning the Triple Crown, leading the league in batting average, home runs and runs batted in. Keating [7] studied the past three decades of elite seasons by players, and found that the proportion of elite seasons by position players ages 25 and under declined sharply, beginning in the early 1990s and bottoming out at 5.9% in 2002, then it started to rise, and it has jumped sharply in 2014 and 2015, hitting a huge 34.4% in 2015.

T. Ahram (ed.), *Advances in Human Factors in Sports, Injury Prevention and Outdoor Recreation*, Advances in Intelligent Systems and Computing 603,
DOI 10.1007/978-3-319-60822-8_2

While younger players have a significant impact on the Major League scene, many 35-year-old-plus players are still contributing at a high level.

In 2007, Julio Franco played after his 49th birthday, also Roger Clemens (45), Jamie Moyer (44) and David Wells (44) are included. Barry Bonds set the record at the age of thirty-six, thirty-seven, and thirty-nine. Table 1 depicts Major League Baseball clubs by average player age in 2016. The Boston Red Sox had the roster with the highest average player age of 31 years in 2014 [8].

Table 1. Major league baseball rosters by average player age in 2016

Clubs	Ages	Clubs	Ages
Seattle Mariners	30.1	Chicago White Sox	28.7
Washington Nationals	29.8	Los Angeles Angels	28.7
Pittsburgh Pirates	29.7	New York Mets	28.6
Toronto Blue Jays	29.6	Los Angeles Dodgers	28.6
San Francisco Giants	29.5	Texas Rangers	28.5
Kansas City Royals	29.4	Houston Astros	28.4
Detroit Tigers	29.3	Cincinnati Reds	28.3
Oakland Athletics	29.3	Colorado Rockies	28.2
Atlanta Braves	29.2	Baltimore Orioles	28.1
Boston Red Sox	29	Minnesota Twins	28.1
New York Yankees	29	St. Louis Cardinals	28
Miami Marlins	29	Milwaukee Brewers	27.9
San Diego Padres	28.9	Philadelphia Phillies	27.8
Chicago Cubs	28.8	Tampa Bay Rays	27.8
Cleveland Indians	28.8	Arizona Diamondbacks	26.9

(The authors created the table based on data from Statistia [8])

Gibbs et al. [9] find that relatively older players outperform relatively younger players for the average Canadian NHL player for a period of 2000 and 2009; however, the relative age effect reversal happened among ALL-Star (2007–2009) and Olympic (1998–2010) team rosters, *i.e.*, younger players outperform older players.

Although MLB players decline after their peak in the late 20’s due to declining health or skills [10], some play into their 40s.

In order to contribute to their teams and stay competitive in the MLB, the older players of MLB should perform differently from the younger players because of deteriorating their physical condition. Their performances need to be considered their physical strengths and experiences.

We have two purposes for this study. The first one is to examine the pitcher’s deciding ball after pushing a batter with two strikes of an aged pitcher group (31 to 43 year-old) and a younger pitcher group (20 to 30 year-old) by using an actual tracking data of the Major League Baseball in 2015. Another purpose of this study is to investigate relationships between pitchers’ striking out batters and their pitching characteristics measured by pitch movements. We use the data from the PITCHf/x®, whose service tracks and digitally records the full trajectory of live baseball pitches.

PITCHf/x® is a pitch tracking system, created by Sportvision, and has been installed in every MLB stadium since around 2006. The data includes pitch type, speed, and movement information. Pitch types are defined by mathematical models that are built around velocity, spin, and movement. It is a constantly evolving, sophisticated system.

2 Literature Review

Using longitudinal data from 86 seasons of Major League Baseball, Bradbury [10] indicates that hitters and pitchers peak around the age of 29 – later than previous estimates of 27. Fair [11] created a model which looked at peak age and how performance deviates from this high point by age. His most intriguing result was that, of players who performed a standard deviation above their expected level of performance for four seasons after the age of 28 (peak age of the study), 14 of the 17 examples played all of these seasons. Sommer [12] attempts to find the number of seasons of major league experience it takes for a player to reach his peak, by examining 5 different seasons over the past fifty years (1966, 1976, 1986, 1996, and 2006) to see how this has changed over time. A ballplayer's batting average in year t for each of his n years in the majors with a minimum of 100 at bats per season was regressed against career year [12, 13]. Sommer [12] found that the profiles have changed dramatically since the 1960s, with conceivable stronger ballplayers reaching a higher peak several years after the batting average reached a peak for regulars in 1966.

Some studies use a statistic called WAR (Wins above Replacement) as a proxy for a player's performance. WAR compares the number of wins that a player adds to his team over a replacement level player at the same position [14–18], which is an attempt by the sabermetric baseball community to summarize a player's total contributions to their team in one statistic. Furnald [19] first used WAR to examine the impact of aging in baseball, concluding that it is important for the Major League Baseball team's management to properly identify how aging is currently affecting players as well as how aging impacts players at different positions. Whiteside, et al. [20] grouped pitch types into three distinct categories: hard pitches (*i.e.*, fastball, sinkers, and cutters), breaking pitches (*i.e.*, sliders, curveballs, and screwballs), and off-speed pitches (*i.e.*, changeups, splitters, and slow curveballs), and found that the proportion of hard pitches thrown decreased significantly until the 7th inning compared with the 1st inning, while the proportions of breaking and off-speed pitches increased. Significant decreases in pitch speed, increases in vertical movement, and decreases in release height emerged no later than the 5th inning, and the largest differences in all variables were generally recorded between the 1st inning and the late innings (7–9). Pitchers were most effective during the 2nd inning and significantly worse in innings 4 and 6.

3 Research Model and Hypotheses

PITCHf/x data include the three-dimensional spatial coordinates of the ball's trajectory, along with several other pitch characteristics. Pitch speed was the exit speed of the ball from the hand. Release location and movement values were reported relative to the right-handed reference frame originating at home plate (y-axis pointing to pitching mound, z-axis pointing up, x-axis orthogonal). Horizontal release and movement values were inverted for left-handed pitchers to permit statistical analyses and interpretation (all values pertain to a right-handed pitcher). Vertical and horizontal release locations were the z and x displacements of the ball, respectively, when it left the pitcher's hand. Vertical and horizontal ball movements were the z and x displacements of the ball between the time it left the pitcher's hand and the time it crossed home plate. Zone percentage represented the percentage of pitches that were thrown in the strike zone. Each of these parameters was recorded using the PITCHf/x ball-tracking system, provided by Sportvision, Chicago, IL, which is installed in all 30 MLB stadiums [21]. PITCHf/x system has home plate as its point of origin, $\hat{y}$ points towards the pitcher, $\hat{z}$ points vertically upward, and $\hat{x} = \hat{y} \times \hat{z}$ (*i.e.*, the x axis points to the catcher's right) [22].

This paper empirically investigates factors affecting pitchers' striking out batters (hereafter, we define it as "close"), *i.e.*, "strike out," as well as "set out." "The set out" includes Called Strike, Swinging Strike, Swinging Strike - Blocked, Swinging on Pitchout, Foul Tip, Foul Tip on Bunt, Automatic Strike, Hit Into Play, Missed Bunt Attempt and Pitchout (For detailed descriptions of variables, see Table 4). Those factors are considered to affect "close" are number of pitches (pitch per at bat and balls), and amount of change in balls (pitch deflection break and pitch arc break) as shown in Fig. 3.

3.1 The Regression Models

First, we perform regression analyses to see which pitch types are closely associated to strike out and set out for the young group and the aged group as shown in model (1), and then, we conduct the structure equation modeling based on four hypotheses. Pitch types are listed in Table 2.

$$\begin{aligned} y_i = \alpha + \beta_1 FF + \beta_2 SL + \beta_3 CU + \beta_4 CH + \beta_5 FA + \beta_6 FC + \beta_7 FO + \beta_8 FS \\ + \beta_9 FT + \beta_{10} KC + \beta_{11} KN + \beta_{12} EP + \beta_{13} SI + \varepsilon \end{aligned} \quad (1)$$

where y_i : *strike out or set out*.

β = weight of each attribute and

ε = residuals.

3.2 The Structural Models

We perform the structure equation modeling to see what is affecting to "close", *i.e.*, pitchers' striking out batters. Since for a left-handed pitcher, everything goes in the opposite direction from a right-handed pitcher, we use absolute values for the analysis. The research model for the structure equation modeling will be as Fig. 1.

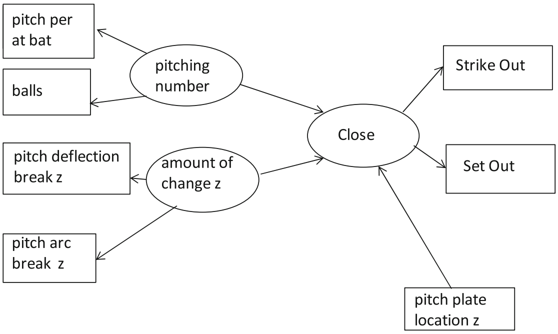

Fig. 1. The research model

More specifically, we will investigate the following three hypotheses regarding factors affecting close;

- H1: Number of pitching will affect close
- H2: Amount of change in z will affect close
- H3: Pitch plate location z will affect close

For estimating a fit between factors, advanced quantitative techniques of structural equation modeling (SEM) [23] have been employed. SEM has been established as an analytical tool, leading to hundreds of published applications per year. Overviews of the state of the method can be found in Cudeck et al. [24], Jöreskog [25], Mueller [26], and Yuan and Bentler [27]. Based on these results of analyses, we will measure how such factors, *i.e.*, shoot chance, cross front goals, players' skills, and in origination area, affect shooting.

In structural equation modeling, we consider the causalities among all variables, especially between the result and the latent variables. A latent variable enables us to find many compiled observed variables at the same time based on the notion of structure. This works for generating and verifying hypotheses to find factors and causalities.

4 Data

Pitching parameters (*i.e.*, pitch type, pitch speed, horizontal release location, vertical release location, horizontal movement, vertical movement, and percentage of pitches in the strike zone) were obtained directly from the PITCHf/x database that is made available by Data Studio Japan Inc., Japan's leading sports information provider. Each pitch is classified into 13 types: four-seam fastball, two-seam, sinker, cut ball, slider, curveball, screwball, knuckle, knuckle curve, change up, splitter, or Eephus pitch (*i.e.*, slow ball). A list of pitch types is shown in Table 2.

Descriptive statistics of variables for Pitchf/x 2015 data are shown in Table 3. An average age for this sample is 29.58 year-old. The youngest pitcher is 21, and the oldest is 43. A list of variables is shown in Table 4. Table 5 contains the Pearson correlation

Table 2. A list of pitch types

Variables	Pitch types
FF	Four-seam fastball
SL	Slider
CU	Curve
CH	Change up
FA	Straight
FC	Cut ball
FO	Fork
FS	Splitter
FT	Two-seam fastball
KC	Knuckle curve
KN	Knuckle
EP	Eephus Pitch
SI	Sinker

Table 3. Descriptive statistics

	N	Min.	Max.	Mean	Deviation
Set_out	87,048	0	1	0.74	0.438
Strikeout	87,048	0	1	0.41	0.491
FF	87,048	0	1	0.32	0.467
SL	87,048	0	1	0.19	0.396
CU	87,048	0	1	0.10	0.298
CH	87,048	0	1	0.11	0.315
FA	87,048	0	1	0.00	0.062
FC	87,048	0	1	0.05	0.219
FO	87,048	0	1	0.00	0.027
FS	87,048	0	1	0.02	0.146
FT	87,048	0	1	0.11	0.311
EP	87,048	0	1	0.00	0.020
SI	87,048	0	1	0.06	0.234
pitch_arc_break_xx	87,048	0.00	8.66	2.1442	1.25081
pitch_per_atbat	87,048	3	15	5.09	1.495
balls	87,048	0	3	1.59	1.049
pitch_plate_location_x	87,048	−4.70	4.23	−0.0184	0.75907
pitch_plate_location_ z	87,048	−2.04	6.88	2.2040	0.85922
age	87,048	21	43	29.58	3.821
		ALL	20–30	31–43(yr old))	
Number of pitches		716	482	234	

Table 4. A list of variables

Variables		Descriptions
pitch_plate_location_z		The height of the pitching position from the ground when a ball reaches the homebase
pitch_deflection_break_x		An estimated change amount in the horizontal direction; measuring the change caused by a ball rotation
pitch_arc_break_z		An estimated change amount in the vertical direction; measuring the change caused by a ball rotation
pitch_per_atbat		Number of throws in the bat
balls		Number of balls during pitching
set out	C	Called Strike
	S	Swinging Strike
	W	Swinging Strike - Blocked
	Q	Swinging on Pitchout
	T	Foul Tip
	O	Foul Tip on Bunt
	A	Automatic Strike
	X	Hit Into Play - Out(s)
	M	Missed Bunt Attempt
	Y	Pitchout - Out(s)
strike out	event_code	Batting result: if event_code = 1, then strike out = 1; else, strike out = 0

Table 5. Correlation matrix

	Set_out	Strike_out	FF	SL	CU	CH	FA	FC	FO	FS
Set_out	1	.487**	−.036**	.040**	.038**	.009**	0.006	−0.005	−0.002	.016**
Strikeout	.487**	1	−.035**	.068**	.052**	−0.004	.008*	−.010**	0.002	.008*
FF	−.036**	−.035**	1	−.338**	−.228**	−.244**	−.043**	−.159**	−.019**	−.103**
SL	.040**	.068**	−.338**	1	−.162**	−.174**	−.030**	−.113**	−.013**	−.074**
CU	.038**	.052**	−.228**	−.162**	1	−.117**	−.021**	−.076**	−.009**	−.049**
CH	.009**	-0.004	-.244**	-.174**	-.117**	1	-.022**	-.082**	-.010**	-.053**
FA	0.006	.008*	−.043**	−.030**	−.021**	−.022**	1	−.014**	−0.002	−.009**
FC	−0.005	−.010**	−.159**	−.113**	−.076**	−.082**	−.014**	1	−0.006	−.034**
FO	−0.002	0.002	−.019**	−.013**	−.009**	−.010**	−0.002	−0.006	1	−0.004
FS	.016**	.008*	−.103**	−.074**	−.049**	−.053**	−.009**	−.034**	−0.004	1
FT	−.036**	−.061**	−.240**	−.171**	−.115”	−.123**	−.022**	−.080**	−.009**	−.052**
KC	.019**	.032**	−.115**	−.082**	−.055**	−.059**	−.010**	−.039**	−0.005	−.025**
EP	−.008*	−.011**	−.014**	−.010**	−0.007	−.007*	−0.001	−0.005	−0.001	−0.003
SI	−.026**	−.043**	−.171**	−.122**	−.082**	−.088**	−.015**	−.057**	−.007*	−.037**
pitch_arc_break_xx	−.007*	−.019**	−.219**	−.269**	.075**	.223**	.040**	−.215**	.007*	.029**
pitch_per atbat	−.185**	−.161**	.059**	−.058**	−.091**	.023**	.009**	.013**	−0.004	−.009**
balls	−.211**	−.167**	.073**	−.069**	−.109**	.017**	0.001	.019**	−0.002	−.013**
pitch_plate_location_z	−.044**	−.123**	.350**	−.200**	−.174**	−.145**	−.029**	.028**	−.015**	−.082**
pitch_plate_location_x	.024**	.043**	−.038**	.101**	.022**	−.042**	−.024**	.032**	−.022**	−.043**

	FT	KC	EP	SI	pitch_arc_break_xx	pitch_per_ atbat	balls	pitch_plat_e_location	pitch_plat_e_location
Set_out	−.036**	.019**	−.008*	−.026**	−.007*	−.185**	−.211**	−.044**	.024**
Strikeout	−.061**	.032**	−.011**	−.043**	−.019**	−.161**	−.167**	−.123**	.043**
FF	−.240**	−.115”	−.014**	−.171**	−.219**	.059**	.073**	.350**	−.038**
SL	−.171**	−.082**	−.010**	−.122**	−.269**	−.058**	−.069**	−.200**	.101**
CU	−.115”	−.055**	−0.007	−.082**	.075**	−.091**	−.109**	−.174**	.022**
CH	−.123**	−.059**	−.007*	−.088**	.223**	.023**	.017**	−.145**	−.042**

(*continued*)

Table 5. (*continued*)

	FT	KC	EP	SI	pitch_arc_break_xx	pitch_per_ atbat	balls	pitch_plat_e_location	pitch_plat_e_location
FA	−.022**	−.010**	−0.001	−.015**	.040**	.009**	0.001	−.029**	−.024**
FC	−.080**	−.039**	−0.005	−.057**	−.215**	.013**	.019**	.028**	.032**
FO	−.009**	−0.005	−0.001	−.007*	.007*	−0.004	−0.002	−.015**	−.022**
FS	−.052**	−.025**	−0.003	−.037**	.029**	−.009**	−.013**	−.082**	−.043**
FT	1	−.058**	−.007*	−.087**	.325**	.044**	.054**	.082**	−.040**
KC	−.058**	1	−0.003	−.042**	.024**	−.048**	−.052**	−.101**	.022**
EP	−.007*	−0.003	1	−0.005	.009**	−0.003	−.007*	−.007*	0.005
SI	−.087**	−.042**	−0.005	1	.227**	.035**	.045**	.047**	−.023**
pitch_arc_break_xx	.325**	.024**	.009**	.227**	1	.012**	.019**	−.059**	−.015**
pitch_per atbat	.044**	−.048**	−0.003	.035**	.012**	1	.822**	.029**	−.033**
balls	.054**	−.052**	−.007*	.045**	.019**	.822**	1	.029**	−.034**
pitch_plate_location_z	.082**	−.101**	−.007*	.047**	−.059**	.029**	.029**	1	−.104**
pitch_plate_location_x	−.040**	.022**	0.005	−.023**	−.015**	−.033**	−.034**	−.104**	1

** The correlation coefficient is significant (two sides) at the 1% level. * at the 5% level.

coefficient between all pairs of twenty-one variables with the two-tailed significance of these coefficients. All variables correlate fairly well and are statistically significant, and none of the correlation coefficients are particularly large; therefore, multicollinearity is not a problem for this data.

5 Results of Analyses

5.1 The Regression Models

We set "strike out," "set out" or "being struck (including four balls)" as an event. The missing values were excluded. Data are limited only when the events occurred and when the pitchers push a batter with two strikes. A regression analysis is performed for each pitch type.

As for the independent variable, a dummy variable is created for each pitch type. The target variable is set to "1" for strike out or set out, and "0" for otherwise. In other words, either the strike out or set out indicate whether the pitcher struck the batter in any way. The results of the regression on the 13 different pitch types are summarized in Table 6 (dependent variable: set out) and Table 7 (dependent variable: strike out).

The results for "set out" shows that all ball types are positive and statistically significant. For all and both age groups, the four-seam fast ball has the highest coefficient. Those which the younger group has the higher coefficient than the older group are Four-seam, Slider, Curve, Straight, while the older group has the higher coefficient in Change up, Cut ball, Two-seam fastball, Knuckle curve, Knuckle, and Sinker.

Table 6. The result of regression analysis (dependent variable: set out)

		Overall			20–30 years old			31–43 years old		
		Coef	t.stat	p.value	Coef	t.stat	p.value	Coef	t.stat	p.value
FF	Four-seam fastball	0.473	275.741	.000	0.483	225.594	0.000	0.456	158.593	0.000
SL	Slider	0.398	231.735	.000	0.407	190.348	0.000	0.380	132.169	0.000
CU	Curve	0.289	168.131	.000	0.299	139.712	0.000	0.269	93.525	0.000
CH	Change up	0.292	170.046	.000	0.287	134.098	0.000	0.301	104.597	0.000
FA	Straight	0.057	32.985	.000	0.067	31.242	0.000	0.031	10.603	0.000
FC	Cut ball	0.191	111.118	.000	0.169	78.930	0.000	0.225	78.278	0.000
FO	Fork	0.022	12.893	.000	0.028	12.884	0.000	–	–	–
FS	Splitter	0.135	78.888	.000	0.110	51.321	0.000	0.173	59.980	0.000
FT	Two-seam fastball	0.266	155.035	.000	0.264	123.390	0.000	0.270	93.888	0.000
KC	Knuckle curve	0.152	88.335	.000	0.149	69.738	0.000	0.156	54.235	0.000
KN	Knuckle	0.045	26.352	.000	–	–	–	0.076	26.391	0.000
EP	Eephus Pitch	0.013	7.749	.000	0.013	5.860	0.000	0.015	5.093	0.000
SI	Sinker	0.195	113.489	.000	0.186	86.919	0.000	0.210	73.052	0.000
R		0.862			0.862[a]			0.863[a]		
R Square[b]		0.743			0.743			0.745		
Adjusted R Square		0.743			0.743			0.745		
Std. Error of the Estimate		0.436			0.437			0.436		

[a] Dependent Variable: set out. [b] Linear Regression through the Origin

Table 7. The result of regression analysis (dependent variable: strike out)

		Overall			20–30 years old			31–43 years old		
		Coef	t.stat	p.value	Coef	t.stat	p.value	Coef	t.stat	p.value
FF	Four-seam fastball	0.339	130.524	0.000	0.342	106.251	0.000	0.333	75.891	0.000
SL	Slider	0.328	126.223	0.000	0.336	104.301	0.000	0.312	71.125	0.000
CU	Curve	0.238	91.669	0.000	0.251	77.998	0.000	0.212	48.291	0.000
CH	Change up	0.210	80.885	0.000	0.209	64.830	0.000	0.213	48.424	0.000
FA	Straight	0.046	17.616	0.000	0.053	16.366	0.000	0.030	6.728	0.000
FC	Cut ball	0.135	52.137	0.000	0.117	36.350	0.000	0.165	37.465	0.000
FO	Fork	0.019	7.177	0.000	0.023	7.169	0.000	–	–	–
FS	Splitter	0.100	38.693	0.000	0.091	28.182	0.000	0.118	26.927	0.000
FT	Two-seam fastball	0.165	63.713	0.000	0.165	51.305	0.000	0.166	37.817	0.000
KC	Knuckle curve	0.129	49.820	0.000	0.125	38.852	0.000	0.137	31.224	0.000
KN	Knuckle	0.026	10.121	0.000	–	–	–	0.045	10.150	0.000
EP	Eephus Pitch	0.004	1.733	0.083	0.004	1.310	0.190	0.005	1.141	0.254
SI	Sinker	0.121	46.786	0.000	0.118	36.775	0.000	0.128	29.061	0.000
R		0.643			0.647[a]			0.637[a]		
R Square[b]		0.413			0.418			0.405		
Adjusted R Square		0.413			0.418			0.405		
Std. Error of the Estimate		0.488			0.488			0.486		

[a] Dependent Variable: strike out. [b] Linear Regression through the Origin

A two seam fastball, much like a sinker or cutter (cut fastball), is gripped slightly tighter and deeper in the throwing-hand than the four-seam fastball. This pitch generally is thought of as a “movement pitch,” as opposed to the four-seam fastball, which is primarily thought of as a “straight pitch” [28]. The results imply that younger pitchers are throwing more straight pitches, while the older pitchers are throwing more movement pitches.

The results for “strike out” show that all ball types, except Eephus Pitch, are positive and statistically significant. Eephus Pitch is positive and statistically significant at a 10% level for an overall result, but positive and not significant for both age groups.

A coefficient for the Eephus Pitch is very small, as well. Those of the younger group that has the higher coefficient than the older group are Four-seam, Slider, Curve, Straight, while the older group has the higher coefficient in Change up, Cut ball, Knuckle curve, Knuckle, and Sinker. There is almost the same level of coefficient in Two-seam fastball for both groups. The results for strike out also imply that younger pitchers are throwing more straight pitches, while the older pitchers are throwing more movement pitches.

5.2 The Structural Equation Models – Results of Hypotheses

Based on the research model depicted in Fig. 1, we test the efficacy of the structural equation model that was conducted by AMOS 24. Among different pitch types, we select the four-seam and the change up, which had higher coefficients on the regression analyses for the young and the aged group respectively. The major results of analysis for the four seam ball for the young group is shown Fig. 2, and those for the aged group is shown in Fig. 3, respectively. The results for the four seam ball for the young group and those for the aged group, and those for the change up for the young and the aged group are shown in Table 8.

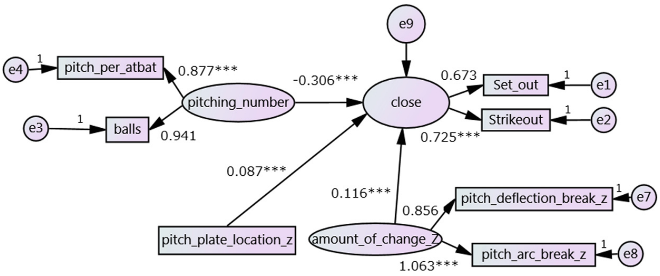

Fig. 2. Four seams (young)

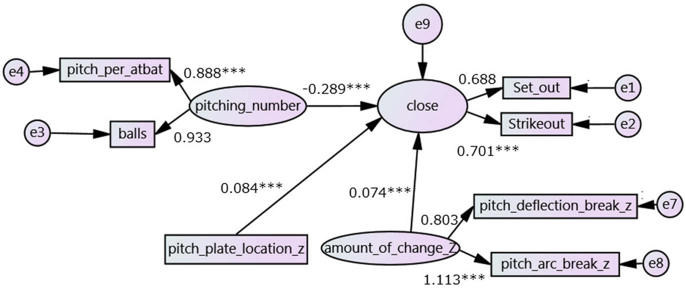

Fig. 3. Four seams (aged)

Table 8. The path coefficients of research models (standard weights)

Construct		Young	Aged	Young	Aged
		Four-seam first ball (FF)		Change up (CH)	
close	← pitching_number	−0.306***	−0.289***	−0.169***	−0.068***
close	← amount_of_change_z	0.116***	0.074***	−0.026	−0.012
close	← pitch_plate_location_z	0.087***	0.084***	−0.233***	−0.229***
Set out	← close	0.673	0.688	0.557	0.419
Strikeout	← close	0.725***	0.701***	0.852***	1.082***
balls	← pitching_number	0.941	0.933	0.887	0.744
pitch_per_atbat	← pitching_number	0.877***	0.888***	0.887***	1.051***
pitch_deflection_break z	← amount_of_change_z	0.856	0.803	0.951	1.334
pitch_arc_break_z	← amount_of_change_z	1.063***	1.113***	0.847***	0.631

*** Denotes significance at 1%

The path diagram highlights the structural relationships. In these diagrams, the measured variables are enclosed in boxes, latent variables are circled, and arrows connecting two variables represent relations, and open arrows represent errors.

When SEM is used to verify a theoretical model, a better goodness of fit is required for SEM analysis; the better the fit, the closer the model matrix and the sample matrix. By means of various goodness-of-fit indexes, including the Goodness-of-Fit statistic (GFI) and the adjusted goodness-of fit (AGFI) [29], the comparative fit index (CFI) [30], and the root mean squared error of approximation (RMSEA) [31], estimated matrix can be evaluated against the observed sample covariance matrix to determine whether the hypothesized model is an acceptable representation of the data. In general, incremental fit indexes (*i.e.*, GFI, AGFI, CFI) above 0.90 signify good model fit. RMSEA values lower than 0.08 signify acceptable model fit, with values lower than 0.05 indicative of good model fit [31]. The research model is shown in Table 8 as GFI = 0.975, AGFI = 0.942, CFI = 0.973, RMSEA = 0.049 for the young group and GFI = 0.979, AGFI = 0.951, CFI = 0.976, RMSEA = 0.045 for the aged group (see Table 9).

The path coefficient for structural models of the four-seam first ball suggested that the regression coefficient between close and pitching number; close and amount of change z; close and pitch plate location z, close and strike out; pitch per at bat and pithing number; pitch arc break z and amount of change z show significance for both the young and the aged group.

Those of the change up suggested the regression coefficient between close and pitching number; close and pitch plate location z, close and strike out; pitch per at bat and pithing number; show significance for both the young and the aged group, while pitch arc break z and amount of change z show significance for the young group but not for the aged group. Since all of the indexes satisfy the cut-off values, these results are regarded as acceptable.

Table 9. Reliability tests

FIT indices	Recommended level	Young	Aged
CMIN/DF	5.0 (Wheaton et al. [32]) ~ 2.0 (Tabachnick and Fidell [29])	89.925	37.995
GFI	>0.90 (Tabachnick and Fidell [29])	0.975	0.979
AGFI	>0.90 (Tabachnick and Fidell [29])	0.942	0.951
CFI	>0.90 (Bentler [30])	0.973	0.976
RMSEA	<0.08 (Browne and Cudeck [31])	0.049	0.045
AIC	Smaller values suggest a good fitting (Akaike [33])	3333.309	1463.827
p-value	>0.05	0.000	0.000

The results of the research models for the young group and the old group for the four-seam first ball show the following three findings;

- H1: Number of pitching is significantly, negatively affecting close
- H2: Amount of change in z is significantly, positively affecting close
- H3: Pitch plate location z is significantly, positively affecting affect close

And those for the change-up show the following three findings;

- H1: Number of pitching is significantly, negatively affecting close
- H2: Amount of change in z is significantly, negatively affecting close
- H3: Pitch plate location z is significantly, positively affecting affect close for the young group, but not statistically significant for the aged group.

The results of the structure models imply that there is not so much of a difference between two age groups in terms of factors relating to the close.

6 Conclusion and Future Study

We conducted two different analyses, *i.e.*, the regression analyses and the structural equation models, in this study. The results from the regression analyses imply that younger pitchers are throwing more straight pitches, while the aged pitchers are throwing more movement pitches. The results of the structure models imply that there is not so much difference between the two age groups in terms of how pitchers' striking out batters.

Movement refers to the spin-induced deflection, and break refers to the maximum bend in the pitch. The arc of a curveball bends much more than a fastball [34]. Older pitchers differ from their younger counterparts in a variety of physical and mental dimensions. Older pitchers may lose their physical strength somewhat, while they have gained their skills in pitching through their careers. We did not study the data in terms of differences in left-arms and right-arms, and that will be our future study.

References

1. Kindall, J.: Sports Illustrated Baseball: Play the Winning Way. Time, Inc., Lanham, MD (1993)
2. Lopez, A., Kirkgard, J.: Coaching Baseball Successfully. Human Kinetics, Champaign, IL (1996)
3. Yeh, C.S., Lee, M.S., Zhu, F.L.: Baseball Bible. Business Weekly Publication, Inc, Taipei (2003)
4. Chiu, S., Hsiao, C., Wu, H.: Measuring pitchers' performance using data envelopment analysis with advanced statistics. Contemp. Manage. Res. **11**(4), 351–384 (2015). doi:10.7903/cmr.14157
5. Shu, P.E.S.: Arsenal/zone rating: a PitchF/X based pitcher projection system. In: 2016 MIT Sloan Sports Analytics Conference Research Paper (2016)
6. Karakolis, T.: Measuring a Pitcher's Ability, Performance, and Contribution, Fan Graphs Community Research. http://www.fangraphs.com/community/measuring-a-pitchers-ability-performance-and-contribution/ (2013). Accessed 16 Feb 2017
7. Keating, P.: Better, Faster, Younger: Why Baseball's Young Stars are Its Best in 20 years, ESPN The Magazine. http://www.espn.com/mlb/story/_/id/14946288/better-faster-younger-mlb-best-stars-20-years (2016). Accessed 17 Feb 2017

8. Statistia: Major League Baseball Rosters by Average player age in 2016. https://www.statista.com/statistics/236223/major-league-baseball-clubs-by-average-age-of-players/ (2016). Accessed 17 Feb 2017
9. Gibbs, B.G., Jarvis, J.A., Dufur, M.J.: The rise of the underdog? The relative age effect reversal among Canadian-born NHL hockey players: a reply to Nolan and Howell. Int. Rev. Sociol. Sport. **47**(5), 644–649 (2012)
10. Bradbury, J.: Peak athletic performance and ageing: evidence from baseball. J. Sports Sci. **27**(6), 599–610 (2009)
11. Fair, R.C.: Estimated age effects in baseball. J. Quant. Anal. Sports **4**(1), 1 (2008). Article 1
12. Sommers, P.: The changing hitting performance in major league baseball, 1966-2006. J. Sports Econ. **9**(4), 435–440 (2008)
13. Scully, G.W.: The Business of Baseball. University of Chicago Press, Chicago (1989)
14. Gennaro, V.: Diamond Dollars: The Economics of Winning in Baseball. Maple Street Press, Hingham, MA (2007)
15. Forman, S.: Player Wins Above Replacement. http://www.baseball-reference.com/blog/archives/ 6063 (2010) Accessed 19 Feb 2017
16. Slowinski, S.: What is WAR? http://www.fangraphs.com/library/index.php/misc/war/ (2010). Accessed 17 Feb 2017
17. Bradbury, J.: Hot Stove Economics: Understanding Baseball's Second Season. Copernicus Books, New York (2011)
18. Baumer, B.S., Jensen, T.: openWAR: an open source system for evaluating overall player performance in major league baseball. J. Quant. Anal. Sports **11**(2), 69–84 (2015)
19. Furnald, N.A.: The Impact of Age on Baseball Players' Performance, Unpublished manuscript. Colgate University (2012)
20. Whiteside, D., Martini, D.N., Zernicke, R.F., Goulet, G.C.: Changes in a starting pitcher's performance characteristics. Across the duration of a major league baseball game. Int. J. Sports Physiol. Perform. **11**(2), 247–254 (2016)
21. Fast, M.: Four plus pitches. The hardball times. http://www.hardballtimes.com/main/article/four-plus-pitches/ 23 June 2010. Accessed 21 Feb 2017
22. Kogan, D.: The anatomy of a pitch: doing Physics with PITCHf/x data. Phys. Teach. **47**(7), 412–416 (2009)
23. Byrne, B.M.: Structural Equation Modeling Using AMOS. Basic Concepts, Applications, and Programming, 2nd edn. Taylor & Francis, New York (2010)
24. Cudeck, R., du Toit, S., Sörbom, D. (eds.): Structural Equation Modeling: Present and Future. Scientific Software International, Chicago (1992)
25. Jöreskog, K.G.: New developments in LISREL—analysis of ordinal variables using polychoric correlations and weighted least squares. Qual. Quant. **24**, 387–404 (1990)
26. Mueller, R.: Structural equation modeling: back to basics. Struct. Equ. Model. **4**, 353–369 (1997)
27. Yuan, K.H., Bentler, P.M.: Mean and covariance structure analysis: theoretical and practical improvements. J. Am. Stat. Assoc. **92**, 767–777 (1997)
28. Ellis, S.: Pitching Grips- How to Grip and Throw Different Baseball Pitches. http://www.thecompletepitcher.com/pitching_grips.htm (2017). Accessed 23 Feb 2017
29. Tabachnick, B.G., Fidell, L.S.: Using Multivariate Statistics, 5th edn. Allyn and Bacon, New York (2007)
30. Bentler, P.M.: Comparative fit indexes in structural models. Psychol. Bull. **107**, 238–246 (1990)
31. Browne, M.W., Cudeck, R.: Alternative ways of assessing model fit. Sociol. Methods Res. **21**, 230–239 (1993)

32. Wheaton, B., Muthen, B., Alwin, D.F., Summers, G.: Assessing reliability and stability in panel models. Sociol. Methodol. **8**(1), 84–136 (1977)
33. Akaike, H.: A New Look at the Statistical Model Identification. IEEE Trans. Autom. Control **19**(6), 716–723 (1974)
34. Fast, M.: What the heck is PITCHf/x. In: Distelheim, J., Jacobs, B. (eds.) The Hardball Times Baseball Annual 2010, pp. 153–158. ACTA Sports, Chicago (2009)

http://www.springer.com/978-3-319-60821-1

Advances in Human Factors in Sports, Injury Prevention
and Outdoor Recreation
Proceedings of the AHFE 2017 International
Conference on Human Factors in Sports, Injury
Prevention and Outdoor Recreation, July 17-21, 2017,
The Westin Bonaventure Hotel, Los Angeles, California,
USA
Ahram, T. (Ed.)
2018, X, 168 p. 52 illus., Softcover
ISBN: 978-3-319-60821-1