

Periodicity of the Franchise Performances of NBA, NHL and MLB

2019 년 6 월

서울국제학교

김찬열

국문초록

주제어: (5단어)

주제어: 주기성, 푸리에변환, 표준편차, MLB, 드래프트제

본 연구는 MLB (Major League Baseball) 프랜차이즈의 연차별 퍼포먼스 의 주기성을 조사하고, 이를 NBA (National Basketball Association)와 NHL (National Hockey League)에서 나타난 경향과 비교하는 것을 목적으로 한다. 먼저, 주기성을 측정하기 위해 FFT (Fast Fourier Transform)를 이용하여 NBA와 NHL에서 가장 긴 역사를 가진 팀의 성과를 분석하였다. 그리고 FFT와 표준편차를 이용하여, 드래프트 전 퍼포먼스의 주기성과 드래프트 후 퍼포먼스의 주기성을 비교하기 위해 아마추어 드래프트보다 더 오랜 역사가 있는 MLB 팀을 분리하였다. 드래프트제의 보상 특성으로 인해 퍼포먼스의 주기성을 감소시킬 것이라는 가설은 MLB 프랜차이즈에 관련해 사실임이 입증되었다. 마지막으로 세 종목의 데이터의 결과를 모두 비교하고, MLB 데이터를 통해 도출된 결론이 다른 종목에 어떻게 적용 될 수 있는 지 분석 하였다.

Abstract

Key Words: Periodicity, Fourier transform, Standard deviation, Draft system, MLB

The objectives of this research are to investigate the possible periodicities of the annual performances of the MLB (Major League Baseball) franchises, and compare them with the trends found in the NBA (National Basketball Association) and the NHL (National Hockey League). I first analyzed the performances of the teams with the longest history in the NBA and the NHL by using FFT (Fast Fourier Transform) to measure the periodicity. Then, I separated the MLB teams with histories longer than that of the amateur draft, in order to compare the periodicity of the performances before the draft to those after the draft using FFT as well as standard deviation. The hypothesis that the draft would decrease the periodicity of the performances due to the compensatory nature of the draft was proven to be true for the MLB franchises. Finally, I compared the results of the three sets of data and analyzed how the conclusion derived from the MLB data could apply to other sets of data as well.

Contents

1. Introduction

2. Theoretical Background

2.1 Fourier Transform, Discrete Fourier Transform, and Fast Fourier Transform

3. Results

3.1 NBA

3.2 NHL

3.3 MLB

4. Conclusion

5. Reference List

Periodicity of the Franchise Performances of NHL and MLB

서울국제학교

김찬열

1. Introduction

Oakland Athletics enjoyed its golden age around 2002, when the baseball franchise adopted the economic management strategy widely known as ‘Moneyball’. It also showed decent performances during the 2012 and 2013 regular seasons, as it ranked 4th and 3rd, respectively, among thirty teams [1]. This 2018 season it is having a great year, recording over 90 wins due to its innovative bullpen tactics [1]. However, between these periods of glory, it had multiple seasons of relatively poor performances. For instance, from 2007 to 2011, Oakland Athletics did not win more than 81 games each year [1]. From 2015 to 2017, its best annual performance was 75 wins [1]. Therefore, distinct fluctuation can be seen in the performance of the Oakland Athletics each season. This led to the question about whether other baseball franchises and teams of other sports have performed with such pattern. I established a hypothesis that annual performances of sports franchises exhibit periodicity over time.

Draft system is a possible factor that might influence the periodicity of annual team performances. Before each season starts, franchises pick the rookies from colleges, minor leagues, and foreign leagues. In this process, generally, teams with poorer performances in the previous year receive better draft picks. Therefore, a franchise with the least wins often suddenly rises the next year, as the first pick rookie fulfills the expectations on him and brings the team to success [10]. Meanwhile, a previously victorious team might fail to recruit any valuable rookie, while losing veteran players from trades and retirements, and gradually become less successful. In this way, the draft system can be

thought as increasing the fluctuations of the franchises' performances. However, the opposite might be also possible: draft system might simply "even out" the teams' performances, leading the periodicity to decrease. This research also aims to find the effects of the draft system on the periodicity of the annual franchise performances.

To evaluate the hypothesis that annual performances of sports franchises exhibit periodicity, I surveyed the data of three of the most popular professional sports leagues in the United States; National Basketball Association (NBA), National Hockey League (NHL), and Major League Baseball (MLB). They have long histories: NBA started in 1946, NHL in 1917, and MLB in 1869. All three leagues contain thirty or thirty-one franchises each, and each regular season consists of more than eighty games. This richness of data enables thorough statistical analysis.

Draft system was adopted by NBA from its very beginning. In NHL, draft was adopted in 1963, which is only twenty years after the currently existing franchises were created [4]. However, particularly in MLB, there are abundant data both before and after the draft system was adopted, as the amateur draft system started in 1965, 64 years after the league started with the current system [2]. Therefore, I could compare the periodicities of the MLB teams' performances before and after the draft, and analyze the impact of the draft on periodicity. The results can be compared with NBA and NHL, and tentative conclusions can be drawn about the impact of the draft on the periodicities of various sports' franchises.

Multiple indices can measure the season performance of a sports team: from basic measures, such as team rank, to complex measures, such as Pythagorean expectation for baseball. For this research, regular season win rate was chosen as the indicator. Win rate is one of the most universal measures; it can be applied to all kinds of sports, and it accounts for the expansion of the leagues over time, as it is calculated by dividing the pure number of wins by the total number of games played per season. It is also one of the most stable indicators. Playoff tournament results can be heavily affected by small variances, such as injury of a single player. However, regular season win rates reflect the team performances with high confidence, as every team plays more than eighty games evenly with each other. Lastly, it is one of the most accessible measures, as the data can be easily collected from sports statistics websites.

The flow of the research is as follows. First, I will discuss about Fourier transform, which is a mathematical tool to find the periodicity of data. Next, I will display the periodicity of the regular season win rates of NBA, NHL, and MLB franchises. For MLB, I will compare the periodicities before and after 1965, which is when the draft system was first adopted. Finally, I will analyze the results and evaluate the hypothesis that sports team performances exhibit periodicity, as well as discuss the impacts of the draft system.

2. Theoretical Background

2.1 Fourier Transform, Discrete Fourier Transform, and Fast Fourier Transform

Fourier transform is a method of changing a function in the time-domain to one in the frequency-domain. From the mathematical fact that every continuous and periodic function can be represented as a combination of sinusoidal functions, Fourier transform gets a time-domain function as an input and outputs its frequency spectrum [7]. From the frequency spectrum, one can obtain the amplitude of each frequency component. The peak of the graph, when clear enough, can be interpreted as the most significant frequency of the function.

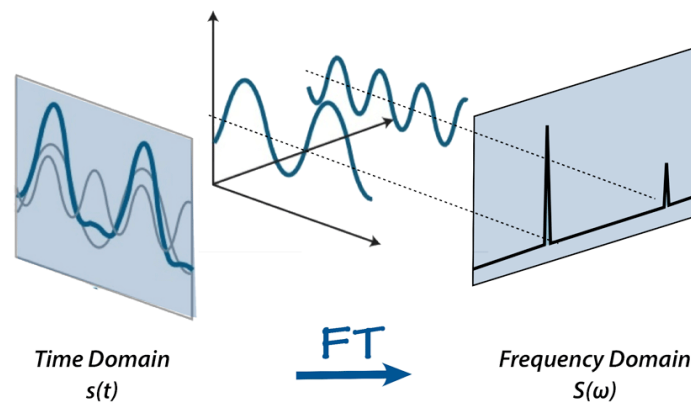


Figure 1. Graphical representation of Fourier Transform [8].

As shown in this figure, Fourier transform takes a continuous function, decomposes it into several separate sinusoidal functions, and prints out their respective amplitudes in the frequency spectrum. The equation for Fourier transform is as follows [7]:

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-2\pi i\omega t} dt$$

t is time and f(t) is the time-domain function. Similarly, w is frequency and F(w) is the frequency spectrum. Therefore, this equation converts a time-domain function into a frequency spectrum [7].

Meanwhile, the data used by this research is not continuous, but discrete, as almost all kinds of data are. Therefore, instead of the formula above, I used the Discrete Fourier Transform (DFT). This version of Fourier transform is applied to various areas, such as digital signal processing and image processing [9]. The formula for DFT is as follows [9]:

$$X_k = \sum_{n=0}^{N-1} x_n * e^{\frac{-2\pi i}{N} kn}$$

$$= \sum_{n=0}^{N-1} x_n * [\cos \cos \left(\frac{2\pi kn}{N} \right) - i * \sin \sin \left(\frac{2\pi kn}{N} \right)]$$

x_n is the discrete time-domain sequence with N samples [9]. Then the sequence can be converted via the above equation to the frequency-domain sequence X_k , where the absolute value of complex X_k represents the amplitude of k-samples frequency component [9].

In the MATLAB software, Fast Fourier Transform (FFT) algorithm was implemented in order to maximize the efficiency of the algorithm. The formula of FFT is as follows: ($W_n = e^{(-2\pi i)/n}$ and n is the number of signals) [6]

$$Y(k) = \sum_{j=1}^n X(j)W_n^{(j-1)(k-1)}$$

Fourier transform has a wide range of applications, including quantum mechanics and signal processing [7]. In this research, Fourier transform was used to obtain the periodicity of the data, which is similar to the time series analysis of statistical signal processing.

3. Results

3.1. NBA

NBA was chosen as the first subject of the periodicity analysis. NBA league was founded in 1946 with 11 franchises, and currently in 2018 has 30 franchises [3]. For data analysis, 15 teams with the longest histories, and thus with sufficient data for periodicity analysis, are chosen.

After tabularizing the regular season win rates and sorting them per team and per season, FFT was performed via MATLAB on each team's win rate over time.

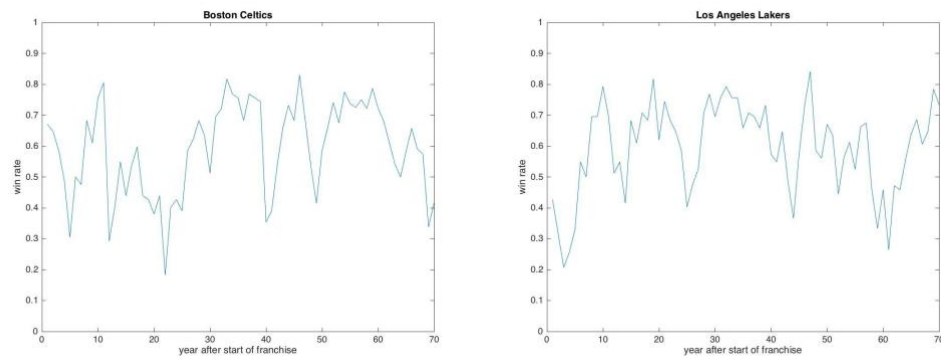


Figure 2. Time-domain plots of Boston Celtics and LA Lakers.

Above are the time-domain plots of the win rates of Boston Celtics and Los Angeles Lakers over 70 years. Certain fluctuation patterns are observed on both plots. This periodicity can be quantified by applying FFT on this data.

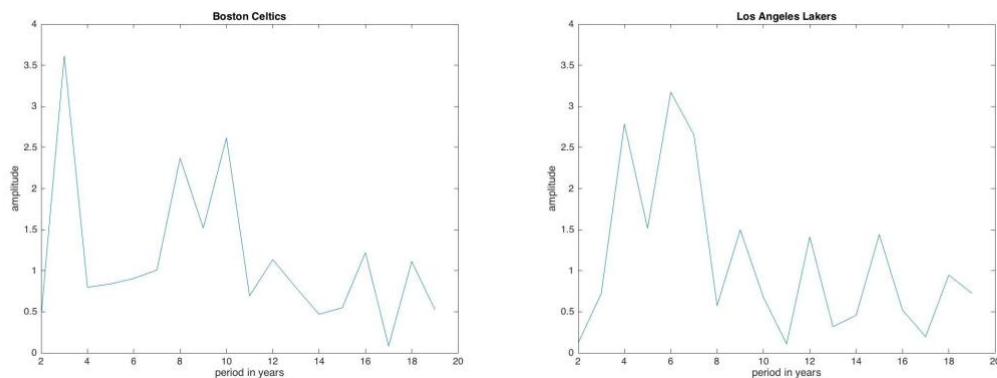


Figure 3. Frequency spectrum plots of Boston Celtics(left) and LA Lakers(right).

Above are the frequency spectrum plots of the win rates of Boston Celtics and Los Angeles Lakers. The fluctuation patterns in Figure 2 are represented as amplitude values in Figure 3. Clear peaks are shown in both frequency spectrum plots: the top two peaks are 3 and 10 years for Boston Celtics, and 6 and 4 years for LA Lakers. Those values of years can be considered as the ‘periodicities’ of each team’s win rates.

Almost all plots showed multiple peaks with varying amplitudes, so only certain peaks had to be selected for better analysis. For this research, the two highest points per each frequency spectrum were collected and displayed in a histogram. This can be used to analyze periodicity of how many years is the most common among the performances of NBA franchises.

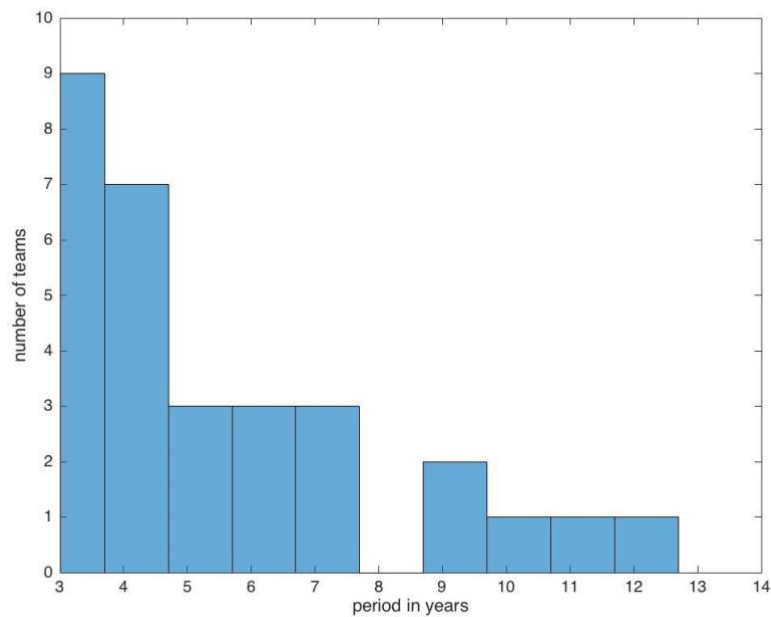


Figure 4. Histogram of the top two periodicity peaks of the frequency spectrum of each NBA franchise's win rates.

The histogram is extremely skewed to the left. 16 of the 30 peaks are peaks of either 3 years or 4 years. This is strong evidence that a substantial number of NBA teams' performances have a periodicity of 3 or 4 years.

3.2. NHL

Although NHL was first founded in 1917, the founding franchises do not exist today. Instead, six currently existing franchises that were established before 1942, known as the Original Six, were selected as subjects of data analysis, because they are the only teams with sufficiently long history for periodicity analysis [4].

As with NBA data, FFT was applied on the regular season win rates of the six NHL franchises. Frequency spectrum plots of the six teams are displayed below.

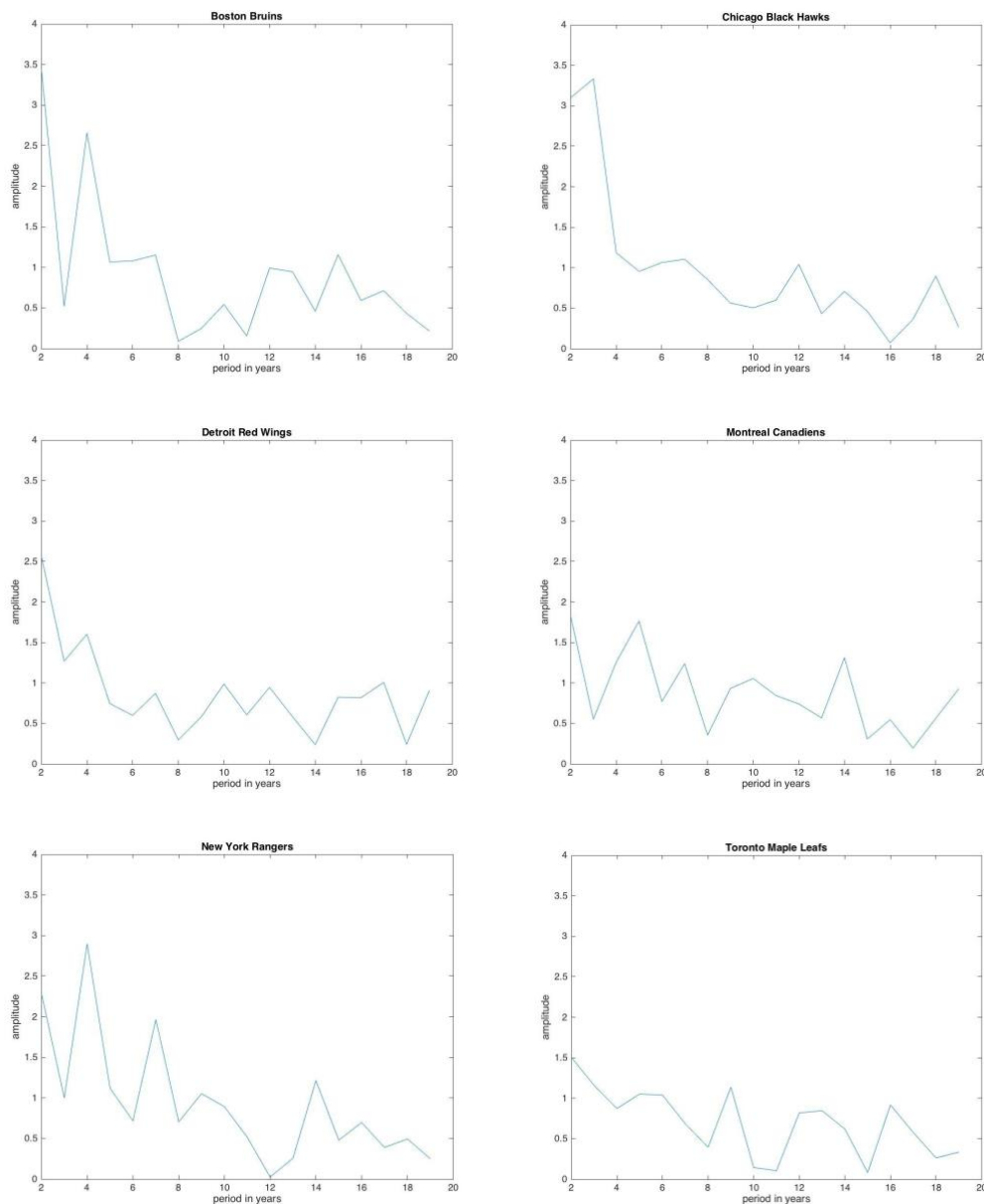


Figure 5. Frequency spectrum plots of six NHL franchises: Boston Bruins, Chicago Black Hawks, Detroit Red Wings, Montreal Canadiens, New York Rangers, and Toronto Maple Leafs (left to right, up to down).

Three franchises' win rates showed clear periodicity of 3 or 4 years. Chicago Black Hawks had a periodicity peak of 3 years, with amplitude 3.3. Boston Bruins and New York Rangers had a periodicity peak of 4 years, with amplitude 2.6 and 2.9 each. These results agree with the analysis of NBA teams, which mostly had periodicities of 3 or 4 years.

For the other three franchises, the frequency spectrums do not contain peaks of amplitude larger than 2 for period longer than 2 years. The small peaks had values similar to the neighboring frequencies. Compared to the aforementioned three teams, it is clear that these teams do not have a clear periodicity.

Of the six franchises, three had strong periodicities of 3 or 4 years, while the other three did not show a distinct fluctuation pattern.

3.3. MLB

Finally, analysis on MLB franchises' win rates is done. MLB was founded in 1869, but it started to have the current league system effectively in 1901. The thirty currently existing teams were selected as subjects of data analysis.

FFT was applied on each team's regular season win rates. The frequency spectrums could be generally classified into two types.

The first type is the plots with distinct peaks.

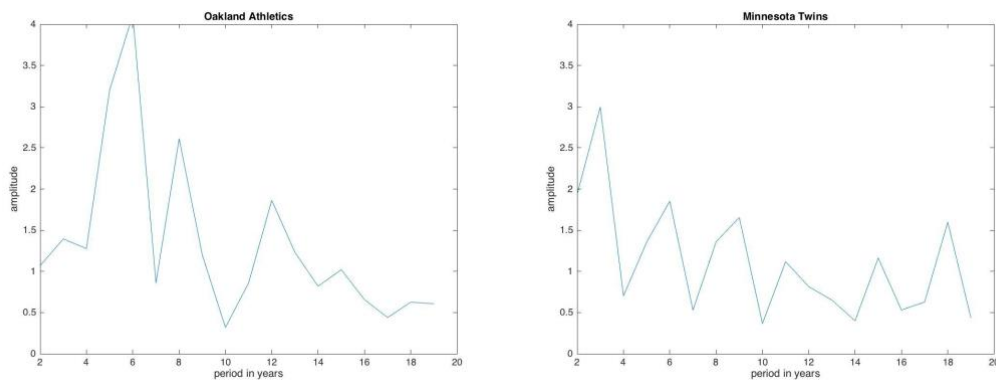


Figure 6. Frequency spectrums of the win rates of Oakland Athletics and Minnesota Twins, as examples of plots with distinct peaks.

Frequency spectrum of the Oakland Athletics is an example of a plot with clear peaks, such as when the period is 6, 8, and 12 years, and etc. The Minnesota Twins also has a plot with distinct peaks. There are clear peaks at periods of 3, 6, and 9 years, and etc. The amplitudes of the peaks are significantly larger than neighboring points. Such peaks represent the periodicities of the franchises' performances.

The second type is the 'flat' plots with no distinct peaks.

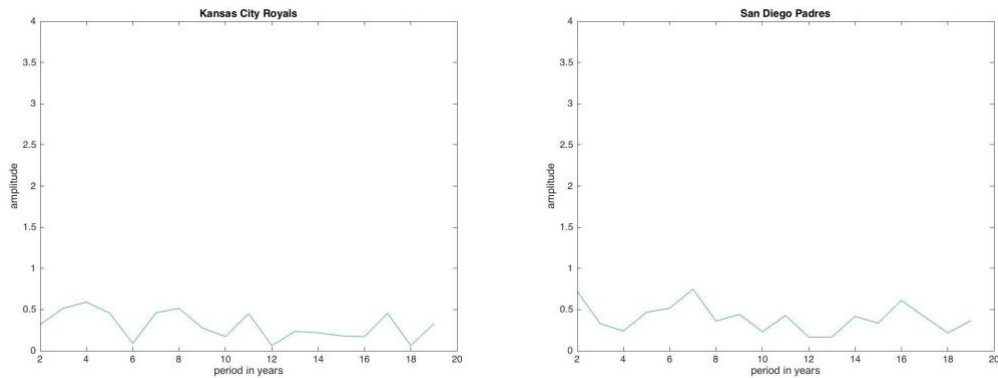


Figure 7. Frequency spectrums of the win rates of Kansas City Royals and San Diego Padres, as examples of ‘flat’ plots with no distinct peaks.

The Kansas City Royals and the San Diego Padres are franchises with win rates of ‘flat’ frequency spectrum. Although both franchises are 50 years old and have enough data to exhibit periodicity, none of the periodicity amplitudes exceeds 0.8. The local maxima points are not significantly larger than neighboring points to be considered as distinct peaks. These plots do not show strong periodicity.

To establish the standards of distinguishing the plots with clear peaks and plots with none of them, I surveyed the maximum amplitude values of the frequency spectrums of all teams.

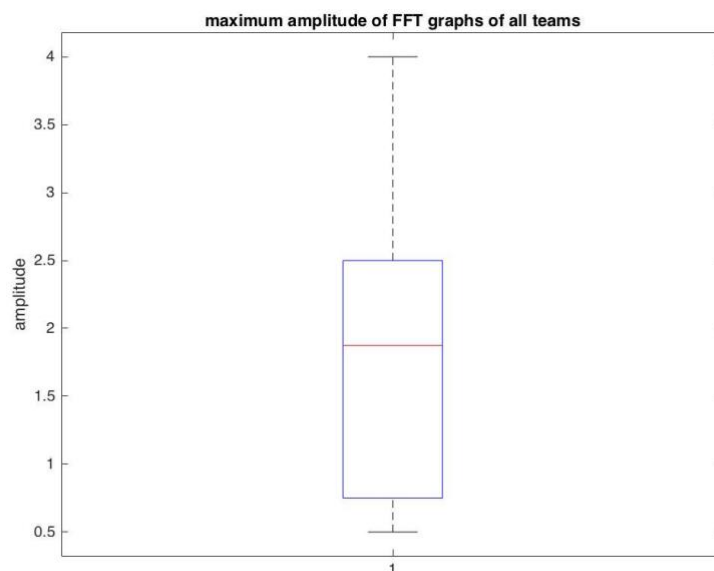


Figure 8. Distribution of the maximum values of the frequency spectrums of all teams.

The 12th Korea Scholars' Conference for Youth

The lower 25% of the maximum values lie between 0.5 and approximate 0.75. From this fact, I defined 'clear peak' as a peak with amplitude higher than 1. If a frequency spectrum contains such peak, it is classified as a plot with a distinct peak. If all of the points are of amplitude lower than 1, the spectrum is classified as a 'flat' plot. The first type plots contain significant periodicity, but the second type plots do not

From this classification, remarkable pattern is found. *Every* franchise that was founded in the beginning years of MLB contains at least one clear peak. However, *all but one* franchise with relatively shorter history, also known as 'expansion teams,' has flat plots.

From the definition above, Arizona Diamondbacks, Colorado Rockies, Houston Astros, Kansas City Royals, Los Angeles Angels, Miami Marlins, Milwaukee Brewers, San Diego Padres, Seattle Mariners, Tampa Bay Rays, Texas Rangers, Toronto Blue Jays, and the Washington Nationals are the thirteen teams with flat plots. All of these franchises are expansion teams that did not start from the beginning years of MLB. The earliest out of these teams was formed in 1961. Meanwhile, 16 of the 17 franchises with a distinct peak were founded in 1901 or 1903. The only exception was New York Mets, which was founded in 1962 but had a clear peak.

Team Acronym	Established Year in MLB	Maximum amplitude
ARI	1998	0.7581
ATL	1876	3.0396
BAL	1901	2.3360
BOS	1901	2.9659
CHC	1876	2.2373
CHW	1901	2.0197
CIN	1890	2.4125
CLE	1901	1.7419

The 12th Korea Scholars' Conference for Youth

COL	1993	0.4646
DET	1901	2.1865
HOU	1962	1.1698
KCR	1969	0.5304
LAA	1961	0.8553
LAD	1890	2.6564
MIA	1993	0.5275
MIL	1969	0.9304
MIN	1901	2.9957
NYM	1962	1.8649
NYN	1901	2.8038
OAK	1901	4.1135
PHI	1883	2.2964
PIT	1887	2.9659
SDP	1969	0.9693
SFG	1883	1.5175
SEA	1977	0.7042
STL	1892	3.1185
TBR	1998	0.9147
TEX	1961	1.0184
TOR	1977	1.0490

WSN	1969	0.8057
-----	------	--------

Table 1. Team acronym - Established year in the MLB – Maximum peak value.

The win rate periodicity is remarkably higher for older franchises than for expansion teams that were founded after 1961. Interestingly, the draft system, a strong candidate for a factor of periodicity of performance, was first adopted by MLB in 1965. Therefore, the following hypothesis is made: *draft system decreases the periodicity of team performances*. This could be because the draft system distributes talents fairly and thereby maintains competitive balance among teams [10]. As a result, team's performance might become more consistent over time, instead of having periodic fluctuations.

To test this hypothesis, I compared the frequency spectrums of the sixteen older franchises before and after 1965, when the draft system was adopted. Win rates from 1901 to 1964 (64 years) were used for the frequency spectrums before the draft, and win rates from 1965 to 2018 (54 years) were used for those after the draft.

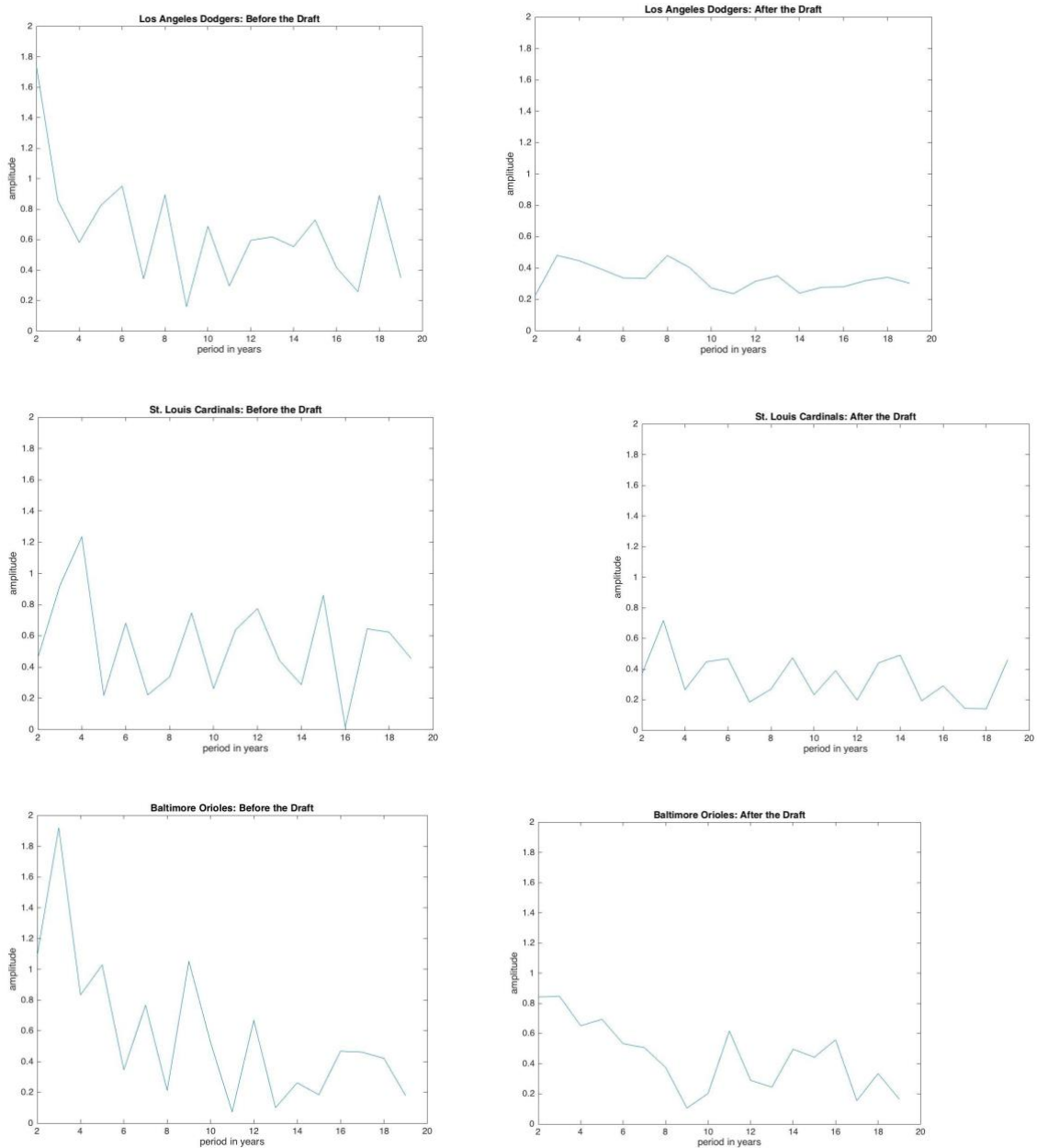


Figure 9. Frequency spectrums of franchises before (left) and after (right) the draft system was adopted. LA Dodgers, St. Louis Cardinals, and Baltimore Orioles, from top to bottom.

It is clearly visualized that the plots are ‘flatter’ after the draft system. Even though the plots of entire history of these franchises had clear peaks, none of the plots after the draft had points with amplitude higher than 1. Meanwhile, the plots before the draft had clear peaks. Therefore, also for the sixteen founding franchises of MLB, periodicity of their performances decreased after the draft system was adopted.

To quantify this finding, I measured the standard deviations of each of the aforementioned 16 teams before and after the draft.

	ATL	BAL	BOS	CHC	CHW	CIN	CLE	DET
Before	0.0920	0.0815	0.0952	0.0906	0.0798	0.0759	0.0679	0.0695
After	0.0808	0.0782	0.0505	0.0654	0.0600	0.0679	0.0703	0.0788
	LAD	MIN	NYN	OAK	PHI	PIT	SFG	STL
Before	0.0879	0.0884	0.0865	0.1195	0.0929	0.0924	0.0786	0.0958
After	0.0507	0.0683	0.0595	0.0788	0.0655	0.0727	0.0596	0.0547

Table 2. Standard deviations of the 16 franchises (that were founded in 1901 or 1903) before and after the draft system was adopted.

All but two franchises (Cleveland Indians and Detroit Tigers) had their standard deviations decrease after the draft system was implemented. Therefore, the team performances became more consistent and stable after the draft, instead of having periodic fluctuations as before the draft. This quantifies and proves the hypothesis that draft system decreases the periodicity of team performances.

4. Conclusion

FFT, a version of Fourier Transform, was used to produce frequency spectrum from time-domain data, and thereby measure the periodicity of team performances. The NBA franchises had a tendency of having strong periodicities of 3 or 4 years. For the NHL, three of the ‘Original Six’ franchises had a periodicity of 3 or 4 years, and else did not have a clear periodicity. Therefore, it is shown that significant number of franchises exhibit periodicities of certain intervals. For the MLB franchises, there is strong evidence that draft system decreases the periodicity of team performances.

It is unclear whether the hypothesis that draft system decreases the periodicities may apply to other leagues than the MLB. For instance, NBA has applied the draft system since the beginning of the league itself, but its franchises still show strong periodicities. Impact of the draft system in NHL cannot be reliably measured because its history before the draft is too short. Numerous factors, such as the contribution of an individual player to the team performance, may alter how the draft system impacts the periodicity for different sports. Moreover, implementation of the free agency system, salary cap, trades, international signings, tanking, changes in rules and the number of teams are just a few of the

factors aside from the draft system that may affect a team's performance and the periodicity of its win rate.

Reference List

(1) MLB Advanced Media, LP, 2019. Year-by-Year Results.

http://oakland.athletics.mlb.com/oak/history/year_by_year_results.jsp.

(2) MLB Advanced Media, LP, 2018. MLB Draft History. <http://mlb.mlb.com/mlb/history/draft/draft.jsp#>.

(3) NBA Media Ventures, LLC, 2019. All-Time NBA Draft History. <http://www.nba.com/history/draft>.

(4) Sports Reference LLC, 2019. NHL Entry and Amateur Draft History. www.hockey-reference.com/draft/.

(5) Raul Amoros, 2016. Top Professional Sports Leagues by Revenue. <https://howmuch.net/articles/sports-leagues-by-revenue>.

(6) The MathWorks, Inc., 2018. Fast Fourier Transform. <https://www.mathworks.com/help/matlab/ref/fft.html>.

(7) Hagit Shakay, 1995. The Fourier Transform - A Primer. <ftp://ftp.cs.brown.edu/pub/techreports/95/cs95-37.pdf>.

(8) Elster LLC, 2019. Fourier Transformation.

http://mriquestions.com/uploads/3/4/5/7/34572113/3311485_orig.gif.

(9) Aashirwad Viswanathan Anand, 2010. A Brief Study of Discrete and Fast Fourier Transforms.

www.math.uchicago.edu/~may/VIGRE/VIGRE2010/REUPapers/Anand.pdf.

(10) Jeffrey A. Rosenthal (1995). The Amateur Sports Draft: The Best Means to the End? Marquette Sports Law Review, vol. 6, no. 1.