Luck vs. skill:
A statistical analysis of sports performances

David Hustadt
Matriculation number 6323981
david.hustadt@student.uni-tuebingen.de

Cornelius Wiehl
Matriculation number 6389275
cornelius.wiehl@student.uni-tuebingen.de

Abstract

The interplay of luck and skill in sports is a complicated and highly discussed topic. This paper aims to quantify and interpret the role of luck in various sport leagues. Using Classical Test Theory and the assumption that skill differences between teams become more apparent over the course of a season, we estimate luck contributions by comparing the variance of point distributions between actual and simulated results in a "luck-only" scenario. Results indicate significant variation across leagues and support previous findings by M.J. Mauboussin [2]. However, it is important to note that the results may not align with intuitive notions of luck and skill. Code can be found here: https://github.com/unicorne/luck_vs_skill

1 Introduction

In order to achieve success in a major sport, a combination of physical and mental talent and dedication to training are essential. However, luck also plays a role in determining success. The degree to which luck factors into success can vary between sports. This does not indicate that some sports require more skill than others. Rather, it reflects how accurately a sport is able to measure skill.

Determining which factors contribute to the effectiveness of measuring skill in a sport is a complex and difficult question that varies for each sport. In general, there is a certain degree of variability in how well a player can demonstrate their skill in a sport. For example, in basketball, a skilled player may have a higher field goal percentage over a large sample size, but a smaller sample size of throws may not accurately reflect their skill level. The number of opportunities a player or team has to showcase their skill, known as the sample size, is one of several factors that can impact the accuracy of measuring skill in a sport. Other factors such as the scoring system, officiating, team dynamics, and environmental conditions like weather and many others also play a role.

The more effectively a sport measures skill, the more confidently one can conclude that the winning player or team is more skilful than their opponent. In sports like chess where only a few factors can disrupt the skill of players there are less surprising outcomes compared to other sports. This allows the best players in the world like Magnus Carlsen to go unbeaten for 118 games [3]. In contrast to sports like chess, games like poker have a high degree of randomness where even the best players in the world can lose to amateurs.

The objective of this report is to classify various sports along this spectrum of luck and skill.

More precisely we are placing different sport leagues at this spectrum of luck and skill. The sport leagues are the five Major leagues of North America: The National Basketball Association (NBA), the National Hockey League (NHL), the National Football League (NFL), Major League Baseball (MLB) and Major League Soccer (MLS) and the top five leagues of European Football: Premier League (England), LaLiga (Spain), Serie A (Italy), Ligue1 (France) and the Bundesliga (Germany).

Project Report for *Data Literacy* 2022/23

1.1 Prior work

An attempt to place several sport leagues in the luck-skill continuum was done by M.J. Mauboussin [2]. The method used is related to the variance method described in 2. However the amount of data used was limited to five years per league and a detailed description of the method used is not available.

2 Method

The variance method used to measure the impact of luck in sports is based on Classical Test Theory which states:

$$ObservedOutcome(O) = Skill(S) + Luck(L) \tag{1}$$

Furthermore we assume the following axioms [1]:

$$\mathbb{E}(L) = 0 \tag{2}$$
$$\rho(L, S) = 0 \tag{3}$$
$$\rho(S_1, L_2) = 0 \tag{4}$$
$$\rho(L_1, L_2) = 0 \tag{5}$$

Axoim (2) means that the expectation of the luck component is zero. Axoim (3) that the skill and luck component are uncorrelated. Axoim (4) that the luck and skill component of two different events are uncorrelated and (5) that the luck components of two different events are uncorrelated. From (1) and (2) follows:

$$\mathbb{E}(O) = \mathbb{E}(S) \tag{6}$$

Which states that if the sample size is large enough the observed results measures the skill. Based on these assumptions we can conclude:

$$\sigma^2(S) = \sigma^2(S) + \sigma^2(L) \tag{7}$$

and therefore use reliability as a measure of how reliable the observed outcome measures the luck or skill component:

$$\rho^2_{OL} = \frac{\sigma^2_L}{\sigma^2_O} = 1 - \rho^2_{OS} \tag{8}$$

3 Experiment Setup

In order to calculate the luck and skill components of sports with (8) we need to obtain the variance of the observed outcome and the variance of the luck factor.

3.1 Sport leagues

A league system, as found in sports like basketball and baseball, offers advantages in determining variances due to the fact that teams play a similar number of games and face opponents of similar skill level, making comparisons across leagues more accurate.

3.1.1 Observed variance

For calculating the observed variance for a particular year we use the regular season results of different sport leagues. Because of changes and expansions of league systems we focus on data since 2000. In order to calculate the observed variance we look at the point distributions of the regular season and estimate the variance on that point distribution: For a league with T teams let $x = \{x_1, x_2, ...x_T\}$ be a set where x_i describes the achieved points of a team during the regular season. The variance of x is calculated with $\sigma^2(x) = \frac{1}{T}\sum_{i=1}^{T}(x_i - \bar{x})^2$ where $\bar{x}$ is the mean of x.

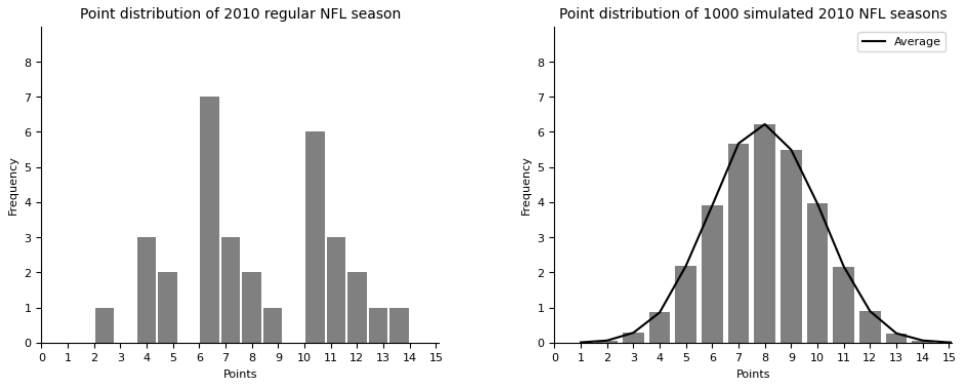

Figure 1: Observed results of the NFL regular season 2010 (left, variance: 8.69) and simulated results (right, variance: 3.98)

3.1.2 Luck variance

In order to estimate the variance of the luck component we are simulating an luck world where there is no skill difference among teams. Under this assumption, each team has a 50% chance of winning each game, resulting over time in an equal number of points for all teams and a variance of zero, in agreement with the axiom (2) that the expected value of the luck component is zero.

However, as sport teams do not play an infinite number of games during a season, some teams may have more luck than others. The variance in an all-luck world is influenced by factors such as the number of teams and games played in a season, as well as the schedule structure. For example, in most football leagues, each team plays each opponent twice, resulting in 38 games for a 20-team league. In contrast, the NFL has a different schedule structure, where teams play some opponents twice a year and others only every four years.

In order to account for these differences we used the actual schedules of each season to simulate each game in order to obtain a simulated regular season result for each season. The variance of the simulated point distribution is calculated according to (3.1.1).

3.1.3 Luck component

With equation (8) the luck component can be easily calculated for one season. For example from figure (1) we know the observed variance for the 2010 NFL season is 8.688. From (1) we know the simulated variance for that year is 3.987 which leads to a luck component of $\rho^2_{OL} = \frac{3.987}{8.688} = 0.459$. By taking the average over all available seasons we estimate the luck component for each sport. (See table 1)

3.2 Data acquisition

The data acquisition process differs for each sport. Most of the observed results were scraped from ESPN.com. To obtain the schedules for leagues different strategies were used. For some leagues like the European football leagues the schedule is the same for each year (except the order) and can therefore be reused. The same holds for various seasons of the MLB or NBA. However some seasons had to be excluded. Several seasons were cut short because of the Covid-19 pandemic and some seasons like the NBA 2011 lockout season were removed from the calculation as well. The NFL schedule was scrapped from ESPN.com and the MLS and NHL schedules were obtained using game data from Kaggle. Preprocessing was done specifically for each league. The detailed data acquisition process is available in the GitHub repository.

4 Results

The results for the five major North American sport leagues and the top five European football leagues are shown in 1. The order of the NBA, Premier League, MLB, NFL and NHL at the luck-

Table 1: Results variance method

League	Teams (avg.)	Seasons	Observed Variance	Luck Variance	Luck contribution
NBA	30.0	16	614.53	81.45	0.13
Premier League	20.0	19	300.57	86.03	0.29
MLB	30.0	20	137.23	40.65	0.3
SerieA	19.89	19	288.61	86.67	0.3
LaLiga	20.0	19	262.62	83.79	0.32
Bundesliga	18.0	19	202.27	76.39	0.38
NFL	32.0	19	9.62	4.02	0.42
NHL	30.14	14	187.31	79.39	0.42
Ligue1	20.0	19	199.39	86.13	0.43
MLS	17.39	18	93.05	73.22	0.79

skill continuum coincidences with the results from M.J. Mauboussin [2]. Except for the NHL, the calculated luck contribution is close to the results from M.J. Mauboussin [2].

5 Discussion

5.1 What is luck?

It is important to note that the concept of luck lacks a clear, definitive definition. In the context of our method, luck refers to all factors that are not related to skill. As discussed in the introduction, this includes a range of different factors. While some of these factors may be obvious, luck may also encompass other, more subtle elements. For instance, research suggests that the National Hockey League may show bias in favor of players born earlier in the year, which could introduce a selection bias [4]. The birthday of NHL players may not be intuitively part of the luck component.

5.2 What is skill?

In general skill in sports refers to the ability to perform physical or mental tasks related to a particular sport or activity. With our approach however, we aim to determine how effectively teams with varying degrees of skill are able to distinguish themselves from one another over the course of a season. The distribution of skill within a league is an essential element for that calculation. This means that the findings for professional leagues with a high overall skill level cannot be applied to amateur leagues.

5.3 Limitations of the variance method

It is debatable if the axioms assumed in 2 are meet in reality. For example axiom (3) assumes that the luck and skill component are uncorrelated. However it is possible that teams and fans are more motivated against stronger opponents introducing another component that is not part of the skill component, and therefore part of the luck component, that depends on skill.

5.4 Explanation of different luck contribution

The concept of luck in sports is a complex interplay of various factors, which poses significant challenges in terms of providing scientific explanations for causality. As assumed with equation (6) the longer a season is played the distinction between skill and luck should become more apparent. However, the presence of other factors may overshadow this effect.

Among other possible explanations for the results can be a selection bias. For example basketball rewards unusually tall players reducing the amount of available players, which in turn, decreases the skill differences within the league, and may result in a lower contribution of luck in the NBA. League structure is also a significant factor. The salary cap regulations vary across leagues, making it easier for teams with more financial resources to differentiate themselves. To make more accurate statements about the cause, further research using additional data is required.

References

[1] Manfred Amelang Lothar Schmidt-Atzert. *Psychologische Diagnostik*. Springer-Lehrbuch. Springer-Verlag, 2012.

[2] M.J. Mauboussin. *The Success Equation: Untangling Skill and Luck in Business, Sports, and Investing*. G - Reference,Information and Interdisciplinary Subjects Series. Harvard Business Review Press, 2012.

[3] Penguincw. Carlsen's unbeaten streak, 2020.

[4] Stephen Cobley Robert O. Deaner, Aaron Lowen. Born at the wrong time: Selection bias in the nhl draft. 2013.