CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Do Left-handers and Left-footers Have a Competitive Advantage in Sports?

Bachelor's thesis

Author: Aner Hadžić

Study program: Economics and Finance

Supervisor: prof. PhDr. Ladislav Krištoufek, Ph.D.

Year of defense: 2023

Declaration of Authorship

I hereby declare that I wrote my thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

I grant a permission to Charles University to reproduce and to distribute copies of this thesis in whole or in part. Moreover, I agree with the thesis being used for study and scientific purposes.

Prague, May 2, 2023

Aner Hadzic

Bibliographic Note
Hadžić, Aner: *Do Left-handers and Left-footers Have a Competitive Advantage in Sports?*. Bachelor's thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2023, pages 116. Advisor: prof. PhDr. Ladislav Krištoufek, Ph.D.

Abstract

Left-sided athletes are often perceived as better performing as they can leverage their minority status within the sports world. While various specific left-sided athletes, such as Lionel Messi and Rafael Nadal, perform at the very top of their disciplines, these might be simply non-representative outliers. The current thesis puts the hypothesis of left-sided over-performance to test via a battery of tests and regressions. My thesis thoroughly analyses the prevalence and the performance of left-handed/left-footed athletes across 5 different sports. As majority of the current studies are focusing only on a few performance metrics in the given sport, my work broadens the knowledge on the topic since it compares the performance of left-sided and right-sided athletes in many categories in order to cover a great portion of the in-game action. Furthermore, this thesis also expands the current understanding of the (potential) left-sided advantage in direct encounters between both teams and individuals, achieving so by implementing predictive Bradley-Terry models that are based on past matches. The overall results are rather surprising: in the majority of the performance comparisons between left-handers/left-footers and right-handers/right-footers, no significant difference between the two groups was detected. Moreover, no important impact of left-handedness/left-footedness on direct contests was found. However, in 4 out of the 5 examined sports there was a significant overrepresentation of left-sided athletes. The results could serve well to coaches or scouts who are looking to seize a prospective advantage of the knowledge of the left-sided performance characteristics.

JEL Classification	C01, C12, C21, C51, Z20
Keywords	handedness, footedness, performance, overrepresentation, sport, athletes

Title	Do Left-handers and Left-footers Have a Competitive Advantage in Sports?

Abstrakt

Levorucí/levonozí sportovci jsou často vnímáni jako výkonnější, protože mohou využít jejich menšinového postavení ve světě sportu. Ačkoliv určití leváci, jako Lionel Messi a Rafael Nadal, jsou na vrcholové úrovni ve svých konkrétních disciplínách, může se jednat pouze o nereprezentativní výjimky. Moje práce pečlivě analyzuje prevalenci a výkonnost levorukých/levonohých sportovců napříč 5 různými sporty. Protože většina současných studií se zaměřuje pouze na několik výkonnostních metrik v daném sportu, moje práce rozšiřuje znalost na dané téma, neboť srovnává výkony leváků a praváků v mnoha kategoriích, aby pokryla značnou část zápasových akcí. Dále tato práce také rozšiřuje současné chápání (potenciální) levácké výhody v přímých střetnutích mezi týmy i jednotlivci, čehož dosahuje pomocí implementace prediktivních Bradley-Terry modelů, které jsou založeny na zápasech z minulosti. Celkové výslekdy jsou poněkud překvapující: ve většině výkonnostních srovnání mezi leváky a praváky nebyl odhalen výrazný rozdíl mezi těmito dvěmi skupinami. Navíc nebyl nalezen ani významný vliv levorukosti/levonohosti na přímé souboje. Nicméně ve 4 z 5 zkoumaných sportů bylo objeveno nadměrné zastoupení leváků. Výsledky mohou doběe posloužit trenérům nebo skautům, kteří chtějí využít potenciální výhody znalosti výkonnostních charakteristik leváků.

Klasifikace JEL C01, C12, C21, C51, Z20

Klíčová slova ručnost, nohatost, výkonnost, nadměrné zastoupení, sport, sportovci

Název práce Mají leváci komparativní výhodu ve sportu?

Acknowledgments

I would like to express my sincere gratitude to my supervisor prof. PhDr. Ladislav Krištoufek, Ph.D.for his help and guidance throughout my writting of the thesis. I would also like to thank my family and friends, especially from IES FSV UK, who supported me throughout my studies.

Typeset in FSV LaTeX template with many thanks to prof. Zuzana Havrankova and prof. Tomas Havranek of Institute of Economic Studies, Faculty of Social Sciences, Charles University.

Contents

List of Tables

List of Figures

Acronyms

NBA National Basketball Association

MLB Major League Baseball

FIFA Federation Internationale de Football Association

EHF The European Handball Federation

ITTF The International Table Tennis Federation

ATP Association of Tennis Professionals

Chapter 1

Introduction

For many centuries there has been a relatively stable prevalence of left-handers in the population (around 10%) (Scharoun and Bryden (2014)). There are many explanations for this peculiar phenomenon, the most convincing one being the desired balance between cooperation and competition among humans. Since our species is to a large extent dependent on cooperation, we naturally leaned more towards one side, that being the right one. However, there has been a persistent combative aspect to our existence, from wars and sword fights to nowadays more common sporting competitions. Multiple studies prove this point, with left-handers (left-footers) being often largely overrepresented in interactive sports compared to the population rate (Live Science (2012)). The cooperative impact on right-sided prevalence can paradoxically also be found in the world of sports. Golf, for example, is a non-interactive game and the player's performance, therefore, is not influenced by other players. On the other hand, it is easier for right-handed players to share the same clubs or learn from another right-handed player, hence cooperation plays a prominent role. To prove our point, only 4% of top golfers are left-handed (Live Science (2012)).

As it turns out, my interest in the left-sided sporting performance is not unique. Many relevant sources in the past have analysed the possible relationship between laterality and sporting achievement. Usually, they would study either the possibility of overrepresentation of lefties in a certain sport (e.g. Loffing (2017)) or better performance of lefties in the sport or in its specific skills (e.g. Laxdal et al (2022)). In my thesis, I will examine both of the aforementioned phenomena. Furthermore, I will extend the current knowledge by comparing the performance of lefties and righties across many different performance metrics, to cover a large portion of the in-game actions as well as to inspect if the left-sided advantage is more prominent in a certain type of action. In addition, I will analyse direct encounters of both individuals and teams to explore the potential benefit of being a leftie (or having more lefties on the team) in head-to-head contests. Specifically, the binomial test is used to detect a prospective overrepresentation of lefties in tennis, football, basketball, handball and table tennis. Afterwards, an OLS model (or FGLS if issues with heteroscedasticity are present) is constructed to analyse the significance and impact of handedness/footedness on a chosen dependent variable (performance metric). Findings from the model are then supported by a series of Mann-Whitney U tests (or T-tests) which inspect if lefties are performing significantly better than righties in specific performance metrics in football, tennis, basketball and handball. For tennis, table tennis and football, a Bradley-Terry model is run to determine if being a leftie is an advantage in direct contests between teams and players.

The results of my analysis suggest that the leftie advantage manifests itself through a significant overrepresentation in all of the sports analysed (with the exception of basketball). However, significantly better performance of lefties compared to righties was not found in almost any specific performance metric. Since my analysis is focused on the top performers in a certain field, a plausible explanation may be that left-handers (left-footers) have a greater chance of becoming an elite performer at a specific skill, but once at the elite level, the leftie advantage diminishes.

The thesis first discusses the already existing literature regarding the performance of lefties in sports. Chosen hypotheses from this section are tested in later parts of the thesis. After the literature review, I present my data, explain the individual variables, inspect their descriptive statistics and analyse the correlation matrices in order to build models where multicollinearity is not an issue. Following the data description, I state the hypotheses for each sport, introduce my approach to examining them, explain the used methods and clarify that all the needed assumptions are met. In the next section, results are presented and discussed along with supplementary tables. The stated hypotheses are revisited and it is shown if my findings are in support of them or not. Lastly, the thesis is concluded and the overall inference is summarized with the mention of possible ways to follow up on my work.

Studying the effects of laterality on sporting success provides a useful guideline for coaches, scouts, or anyone who is interested in improving their team and gaining an edge over their rivals through the correct approach of dealing with and utilizing the leftie advantage.

Chapter 2

Literature review

2.1 Explanation of the main concepts

According to Loffing et al (2016) there are two main hypotheses that may explain the intriguing trend of left-handed/left-footed overrepresentation in sports.

Firstly, there is the innate superiority hypothesis, which states that due to a variety of processes associated with left-handedness athletes have an inborn advantage in sporting activities. The left side of our body is controlled by the right hemisphere of the brain, which is also responsible for visuospatial and spatiotemporal skills. Left-sided people may therefore be more efficient when performing manual and motoric tasks. Another reason for the innate superiority is the proposed lack of lateralization (McLean and Ciurczak (1982)) and greater proficiency in handling motoric skills with the non-dominant hand for left-handers compared to their right-handed counterparts. This lack of lateralization should give them an advantage when developing bimanual skills. The last argument that advocates for the innate superiority is the potential difference in hormonal configuration between left-handers and right-handers. Faurie et al (2011) propose that testosterone levels might be higher for left-handers compared to right-handers, which is generally associated with higher aggressivity and therefore is considered a potential advantage in competitive fighting.

Another explanation for the left-sided advantage lies in the fact that left-dominant people are scarce among the general population, which in turn makes them more unpredictable in sporting contests. This hypothesis is often described as negative frequency dependence. The negative frequency dependence hypothesis, therefore, suggests that left-handers should have a competitive advantage only in duel-like interactive contests where athletes are directly influencing their opponents. The lower familiarity of left-dominant movements which stems from the rarity of left-handers makes it more difficult for athletes to react in the most suitable way. This hypothesis also states that left-dominant people should not have any competitive advantage in noninteractive sports such as golf or gymnastics. It also predicts that the left-sided advantage should decrease with the rise of the rate of left-dominant people in population. Hence we could hypothesise that the representation of left-handed athletes in interactive contests should reach its optimal rate, which should be higher than in the general population, and approximately maintain that level further on.

Another reason for the overrepresentation of lefties is linked to the tactical element of team sports. In many team sports (football, handball, etc.) there are a few positions designated specifically for left-dominant people (left-back in football, right-backcourt in handball) and other positions where it is advantageous to have a variety of both left-dominant and right-dominant players (wingers in football). This tactical need for left-dominant players is usually much greater than the common occurrence of lefties in the general population.

2.2 Literature concerning handedness in team sports

As Loffing et al (2016) suggest, left-footed footballers are overrepresented and their occurrence among professionals is around 20%. Interestingly enough, the percentage of left-footed footballers gradually drops with the decrease in performance level as semi-professionals and amateurs are left-footed less frequently. Findings from Akpinar and Bicer (2014) reinforce the assumption that left-footed players occur with higher frequency in top-level football, with data from the 2013/2014 season confirming that Real Madrid, Arsenal, Besiktas, AC Milan and Barcelona all had more than the aforementioned 10% (population rate) of left-footed players in their preferred starting line-up. Regarding possible performance advantages, Bozkurt and Kucuk (2018) have compared under-15 left-footed and right-footed footballers from various football clubs in Istanbul in 4 different technical skills (dribbling, shooting, long passing and juggling). There was no statistically significant difference occurring between left-footed and right-footed players in any of the aforementioned skills. The findings of the experiment favor the more frequent theory, which suggests that any possible performance advantage of left-handed athletes would stem from their scarcity rather than their innate predisposition.

Another remarkable finding by Loffing et al (2016) is that the frequency of use of the dominant foot in football-related behavior varies from discipline to discipline with set-pieces (free-kicks, penalties, goal-kicks, corners), dribbling and passing having the highest dominant foot use at about 85%. On the other hand, shooting, first touch, clearances and tackles are slightly less lateralized actions, with dominant foot use only at around 70%. High lateralization for in-game actions may suggest two findings: firstly, footedness plays a crucial role for football players as the majority of the time they will utilize their dominant foot even in a situation where it would be easier to switch to the non-dominant foot. Secondly, even after football-specific training and rigorous practice helping footballers to acquire a high skill level for both feet, in-game research shows us that the preference for dominant foot use prevails. This phenomenon proves that footballers are given roles such that it will be tactically advantageous to play with their stronger foot.

An interesting analysis regarding left-handed advantage in handball has been conducted by Laxdal et al (2022). They collected data on 7-meter shots from 4 European Championships in the time period between 2016 and 2022. The data contained 1625 7-meter shots taken by 185 different players across 229 games. The relationship was analyzed by a Bayesian two-level analysis. The outcomes of the 7-meter shots from level one were nested within the shooters in level two where handedness was a covariate. The results of the study tell us that 75.3% of all the 7-meter shots taken resulted in a goal. 54.8% of total shots were thrown by right-handed players, with the left-handed players being responsible for the remaining 45.2%. Of the 158 different 7-meter executors, 98 (62%) were right-handed and 60 (38%) were left-handed. The study demonstrates that left-handers are clearly overrepresented among the designated 7-meter takers compared to the general population (38% vs 11.6%) and the proportion of left-handers on any given team (38% vs 25%). According to the analysis, handedness had no significant relationship with regular scoring. The authors argue that their analysis serves as a solid argument for the negative frequency dependence as left-handers are overrepresented among world-class 7-meter executors. Interestingly enough, the actual success rate of the left-handers and right-handers in the study does not differ significantly, suggesting that being left-handed does not give players a competitive edge among world-class 7-meter takers, it rather increases one's chances of becoming a world-class taker. Laxdal et al also propose that there might be an innate advantage to being a left-handed 7-meter taker as top-level handball goalkeepers had constant exposure to left-handed shooters in practice throughout their careers and therefore should have increased perceptual familiarity with them. However, Schorer et al (2012) argue against the aforementioned notion, suggesting that exclusive confrontation with left- or right-handed penalty takers in handball during practice enhanced the hand-specific prediction of penalty shots for novice handball goalkeepers. In other words, goalkeepers who practiced with the left-handed group improved their success against left-handed penalties and vice-versa for the right-handed group.

Regarding overall performance, Loffing et al (2015) found that left-handed players are regularly overrepresented among top goalscorers in international tournaments, indicating that the left-handed advantage may apply to a larger array of skills in handball.

Further proof of the persisting struggles when anticipating left-handed movements may be found in a study led by Loffing et al (2011) where the focus is on anticipation of the direction of spikes in volleyball. In addition to handedness, skill level and temporal occlusion were also taken into account when analyzing the effect on prediction success. 3 left-handed and 3-right-handed volleyball players were assembled to perform spikes from 3 different positions (Positions 2,4 and 3) with each of the positions having 2 possible ball trajectories. 18 skilled and 18 novice players were supposed to predict the correct direction of the spikes from footage of the 6 test players. The study revealed two important results: firstly, as expected, the predictions of right-handed spikes were notably more accurate than the left-handed ones. Secondly, there was no significant interaction between attacker's handedness and skill level, meaning that the difficulty of predicting left-handed spikes was not affected by the performance level of the given player.

A paper by Lawler and Lawler (2011) examined a large sample of basketball players in the NBA to study a potential relationship between handedness and performance. They found that left-handed players performed better than their right-handed counterparts in rebounds, assists, points per game and field-goal percentage, thus they assumed that left-handers could have a negative frequency dependence advantage over right-handers. Furthermore, a study conducted by Stockel and Vater (2014) analysed the relationship between everyday life and basketball-specific hand preference and found those two to be highly correlated. However, both of them were poorly correlated to self-reported measures of basketball-specific practice, which tends to lessen the lateralization of players. This finding is confirmed by another study by Stockel and Weigelt (2012) which states that professional basketball players use their nondominant hand only for 27.7% of all ball contact. This may suggest that hand preference is a robust characteristic of basketball players and therefore handedness plays a major role in in-game movements and the playing style of basketballers.

Multiple different studies came to the conclusion that left-handed prevalence among basketball players is similar to the one among the general population, which potentially argues against any significant negative frequency dependence advantage. (Loffing et al (2016))

Stockel and Vater (2014) revealed that players demonstrated the highest dominant hand bias in long-distance shooting. On the other hand, layups, dribbling skills and short passes had the lowest deviance from equal hand use. Patterns of hand preference did not differ between playing positions. The evidence of their study also points out that when the pressure from time and opponent is lowered players tend to revert to their dominant hand. On the other hand, when players have to adapt quickly to their circumstances on the court they usually use the hand that is more convenient in the given situation. This phenomenon could indicate that the value of nondominant hand use is growing together with the competition level, as higher competitions usually comprise high time pressure. In support of this argument are Stockel and Weigelt (2012) who found that the proficiency to use both hands is increasing with the rising performance level, therefore being a professional basketball player may be facilitated by the ability to use the nondominant hand.

According to a study conducted by Brooks et al (2004) left-handed batters provide an advantage for their team in cricket. Team success was positively related to the left-handed innings rate by the team, which would suggest that left-handed hits are more difficult to cope with.

2.3 Literature concerning handedness in one-on-one sports

A significant overrepresentation of left-sided athletes is present at the elite level of duel-like interactive individual sports (fencing, boxing, table tennis) and in team sports with high importance of one-on-one interactions (Loffing et al (2016)). Since the same cannot be said about non-interactive sports such as darts, golf, snooker, or bowling, they interpret that sort of overrepresentation as indirect evidence for an interactive advantage. The logical implication of that theory would be a case of the negative frequency dependence effect emerging, as any hypothetical innate advantage would naturally translate into noninteractive contests.

Loffing et al (2016) divide interactive sports into direct and indirect. Unlike in indirect sports, in direct interactive sports, athletes can directly physically manipulate each other. (Grouios et al (2000)).

As Grouios et al (2000) affirm in their research on ”class A” athletes in northern Greece, left-sidedness is more frequent in direct interactive sports compared to the indirect ones. Moreover, recent studies suggest that among the class of structurally related interactive racket sports (badminton, tennis, table tennis, etc.) the percentage of left-sided athletes increases with the decreasing time available for players to react, suggesting that left-sided athletes are more frequent in sports with greater time pressure. (Loffing et al (2016)) This notion is then confirmed by a later study from the same author. (Loffing (2017)) It deals with time pressure as a proposed factor that intensifies the effect of the negative frequency dependence advantage. He defined time pressure as the "mean time interval between the actions of two interacting players in male competition". Data on 6 different sports with similar performance and strategic demands (ball is hit/thrown; bat/racket is used for intercepting the ball etc.) were chosen in order for them to be comparable, with the significant difference between each of them being the time pressure. The sports included are tennis, table tennis, squash, badminton, cricket and baseball. Data were analysed in the 2009-2014 time frame. For each racket sport, the top 100 players of all the year-end rankings were analysed, for cricket, the top 100 Test bowlers were investigated and for baseball, ranking data for the top 78 to 94 (differing by season) pitchers in Major League Baseball were assessed. To analyse the relationship between time constraint and the left-sided advantage, Loffing correlated the time pressure in each sport with the frequencies of left-handers observed in the elite rankings. The time pressure was found to be highest in baseball and lowest in squash.

One-tailed binomial tests reported a left-hander overrepresentation in baseball, cricket and table tennis, but not in the other three sports. As expected, left-handedness increased with the rise of time pressure. There was no clear overrepresentation among the three sports that were rated as low time pressure, which offers an assumption that the negative frequency dependence advantage is present only in high time pressure circumstances.

Analysis of year-end world rankings in men's tennis, ranging from 1973 to 2011, shows a linear decrease of left-handed players in the top 10 over time and an increase followed by a decrease in the top 100 players. These trends may propose that the left-handed advantage was more prominent in the past when it was more difficult to have a detailed preparation for each opponent. The unfamiliarity of the left-handed strokes would then make them even harder to predict for the opposing players. In the aforementioned timeframe, there have been 16 different year-end world no. 1s, out of which 3 were left-handed. The left-handed number 1s spent in total 11 out of the 39 years as the world's best player.(Loffing, Hagemann, et al (2012))

The upper charts in the image below depict the year-end world no.1s in male and female tennis (with the number of years they spent in that position) in the period of 1973-2011. The bottom tables show the frequencies of left-handed tennis players among the top 10 and the top 100 players in the timeframe of 1973-2011.

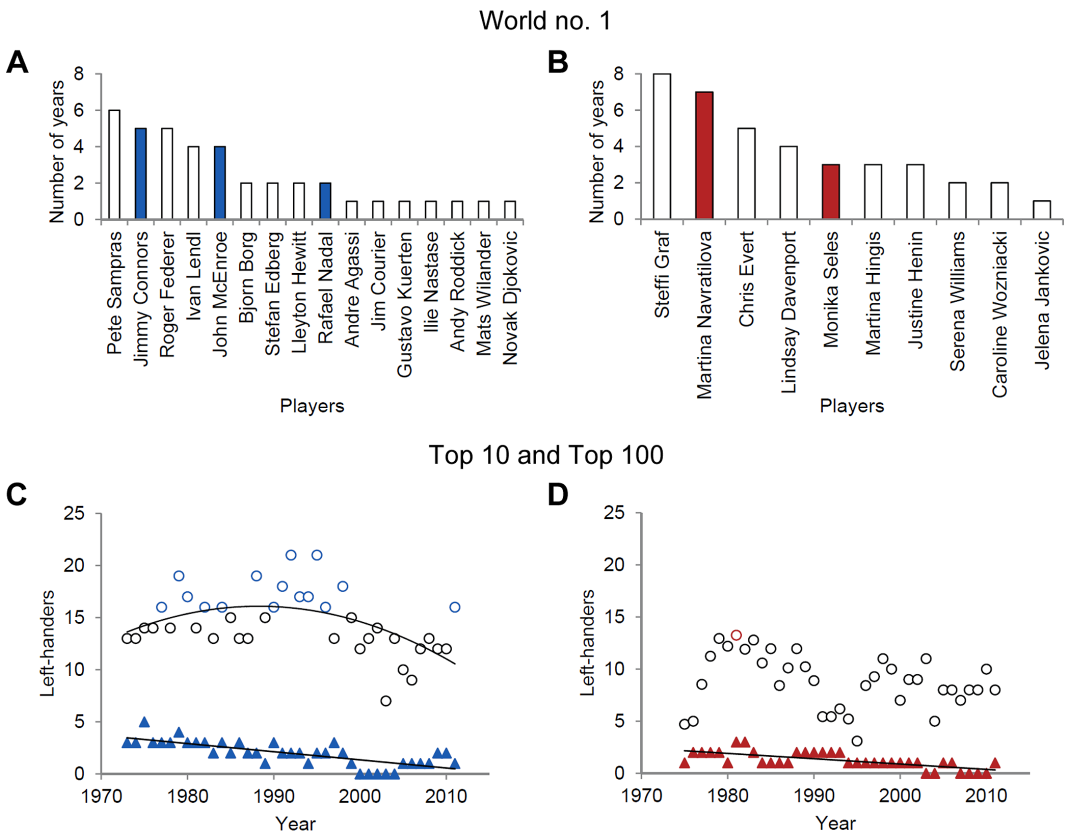

Figure 2.1: Left-handed frequencies among top 10 and top 100 ATP players; year-end no.1s (by Loffing, Hagemann, et al (2012))

In the same article an analysis of left-handed tennis players at Grand Slam tournaments was also conducted. The data were collected from all the Grand Slam tournaments from 1968 until 2011 (since the beginning of the Open era). The analysis suggests that the representation of left-handers at the Grand Slam tournaments was increasing until the early 1990s and then slowly started decreasing. This trend is consistent with the development of the ATP top 100 year-end rankings over time, which makes sense as the participants at the Grand Slam tournaments are mainly the top-ranked players. Similarly, the number of left-handers in Grand Slam finals reached its peak in the early years of the open era and has since decreased. Taking these statistics together, it can be assumed that left-handers had an advantage for some period of time in the past. However, that advantage has diminished over time due to the higher professionalism and more technologically advanced tactical and practice tools that enable players to prepare in detail for each individual opponent they are about to face. These factors may create an effective way for tennis players to deal with the negative frequency dependence advantage that left-handers would assumably have. In support of the latter notion is a paper by Loffing and Schorer (2021) which proposes that the left-handed advantage among tennis players is present at the junior level but not at the senior level. Their possible explanation for this observation is the fact that senior players have a better position for detailed preparation against each individual left-handed opponent.

Furthermore, Loffing, Hagemann, et al (2012) have analysed the left-handed advantage among amateur tennis players. They conducted a study on registered tennis players in the German Westphalian Tennis Association (WTV) in the summer season of 2008. Each individual player was assigned a performance level mark ranging from 1 (highest) to 23 (lowest) based on their playing results in the previous season. Only 6.82% out of the total 2185 players in the sample were left-handed, which would argue against any sort of overrepresentation at the amateur level. Crucially, however, the left-handed frequency was logarithmically increasing with higher performance levels among the amateurs. We can interpret this result as a left-handed advantage among amateurs due to negative frequency dependence, especially with their incidence in the sample being even lower than among the general population. Therefore, it may be hypothesised that the negative frequency dependence advantage is more evident at the amateur level because of the players' lack of access to elite-level preparation tools (video analysis, performance statistics, etc.). The amateurs cannot prepare as precisely for each upcoming opponent and therefore may struggle with the less familiar left-handed strokes.

The graphs below show the relationship between left-handed frequencies and performance level for male and female tennis players at the amateur level.

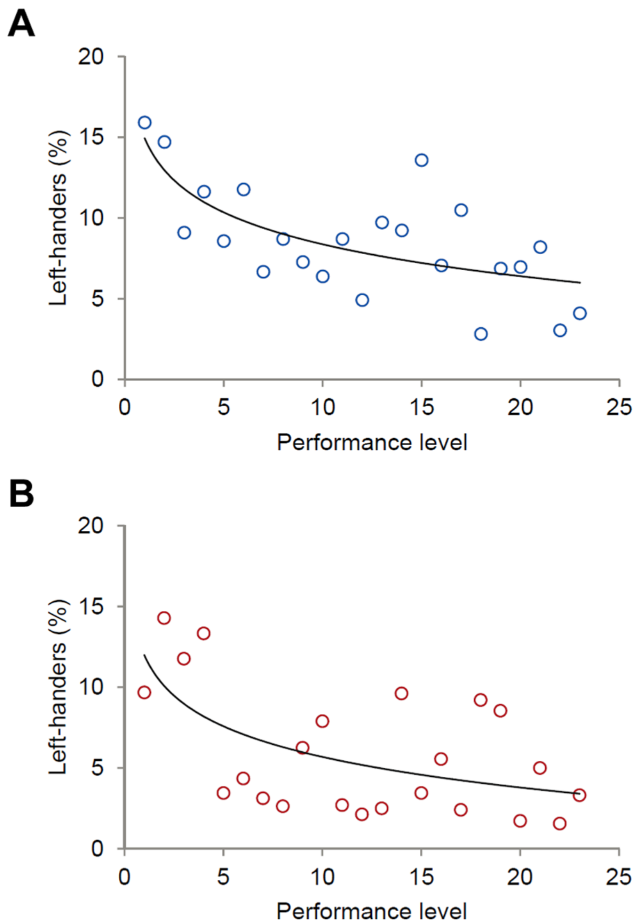

Figure 2.2: Left-handed performance across amateur performance levels (by Loffing, Hagemann, et al (2012))

Another argument in support of the left-handed advantage is provided by a study conducted by del Corral and Prieto-Rodriguez (2010) that analysed match data from Grand Slam tournaments in the period of 2005-2008. They discovered that once quality differences were controlled for, the probability of success for male right-handed players against left-handers lowered by about 5.9%.

In the same study, simulations based on world ranking data in the period 2005-2008 conducted by del Corral and Prieto-Rodriguez (2010) tell us that left-handed tennis players with lower ranking have a higher probability of winning against higher-ranked right-handers. These results would contradict the previously mentioned stance that left-handed advantage among top-level tennis players has diminished with the introduction of comprehensive preparation tools. As a demonstration of difficulties caused by the rarity of left-handed tennis players when predicting their moves, I present the analysis conducted by Hagemann (2009) where he asked a total of 108 tennis players to predict the outcome of left and right-handed tennis strokes shown in videos. As a measure of control for the effect of handedness, he selected 54 right-handers and 54 left-handers to participate in the analysis. In accordance with the negative frequency dependence hypothesis, the accuracy of predictions was lower against left-handed players than against right-handed players, irrespective of the handedness of the players that were predicting the outcomes. Tennis players often try to stroke the ball in the direction of the opponent's backhand as forehand is usually the more preferable stroke. As such, Loffing, Hagemann and Strauss (2010) computed a computer-based experiment from which they found that tennis players across all performance levels stroked more balls to the backhand side of right-handed players compared to left-handers' backhand side. An assumption can be made that through constant exposure to right-handed players, tennis players of all performance levels automatised directing the ball to their backhand side. As such, left-handed players may potentially benefit from this behavior as, in theory at least, they receive most of the balls on their preferred forehand side.

To really underline the sole effect of being left-handed, Fagan et al (2019) created a latent ability model to differentiate between natural talent and the advantage that stems from being left-handed. Their methods were applied to four different sports with left-handed overrepresentation and the findings are truly compelling. For each sport, they defined variables Drop_alone and Drop_all. Drop_alone denotes the drop in rank for a left-handed player when only he gives up his left-handed advantage. Drop_all depicts the drop in the rank for a left-hander if all the left-handers give up their left-handed advantage. Naturally, both variables show the largest drop in the rank for table tennis, where the left-handed overrepresentation is the most pronounced. Conversely, the lowest drop-off values can be recorded for tennis. As discussed previously in my thesis, the high drop scores in table tennis may be explained by the critical influence of time pressure in this sport, as opposed to the comparably lower time pressure that is present in tennis.

Finally, it is important to consider the lateral advantage for different divisions and disciplines in interactive sports. Interactive sports that are divided into varying disciplines or different weight divisions often have differing rates of left-sided frequency. In boxing in particular, the frequency of the so-called ”southpaw” athletes that were listed in The Ring magazine’s annual ratings from 1998 to 2012, had a high variation across the 17 weight divisions, with junior middleweight having the largest left-handed representation of 42.3%. Conversely, the heavyweight category had the lowest left-handed representation with only 14%. (Loffing and Hagemann (2015)).

Chapter 3

Data and Methodology

3.1 Data description

The data used in my thesis are divided by two important factors: Firstly, it should come as no surprise that each sport has its own dataset. More interestingly, however, 3 sports in my analysis have more than one dataset, namely football, tennis and table tennis.
The reason for this division is my intention to analyse regressions on a broader sample as well as study the impact of left-handedness/left-footedness in direct sporting confrontations on the expected probability of winning success on a smaller sample. Both of these methods involve a factor of left-handedness/left-footedness being incorporated into the model in order to explore the possible laterality effect. In this section, I will describe the data separately for each sport and discuss my selection methods.

It is important to mention that due to the nature of my analysis I opted for analysing each sport for only one season/tournament, the reason being that there is a great fluctuation of players in the majority of the competitions in my sample. The best example of this trend is football, where every summer clubs spend enormous amounts of money on new signings while offloading players that are deemed surplus to requirement. In sports where I selected players based on ranking, namely tennis and table tennis, the fluctuations of the top 100 players each year are so significant that it would not be reasonable to use data for multiple seasons. Lastly, international tournaments that take place every few years (FIFA World Cup, EHF European Championship, etc.) never feature the exact same teams, many players end their international careers while others newly arrive on the scene. Therefore the composition of players participating in these events changes dramatically each time. For all the reasons mentioned above, I decided on cross-sectional data for all of the sports reviewed in my work.

3.1.1 Data for Tennis

As mentioned before, I have created 2 separate datasets for tennis in order to analyse 2 different models and the impact of left-handedness in each of them. I will start with the dataset used for the Ordinary Least Squares Regression. I gathered data on the top 100 ATP male tennis players according to the year-end world ranking in 2021. The data were collected from a specialized tennis statistics website ultimatetennisstatistics.com.

As my dependent variable, I chose *ATP points*. *ATP points* for each year-end rankings are determined by all the official results achieved during the previous year. I believe that this metric is the perfect way to measure the success of a tennis player in a given season as all professional tennis players are ranked in the ATP rankings based on their *ATP points*.

I selected in total 6 independent variables. While deciding which independent variables to choose my main objective was to look for a set of variables that together would cover the majority of in-game situations in tennis. The most important independent variable for my thesis is *leftright*, which denotes 1 if the given player is left-handed and 0 if he is right-handed. The same dummy variable approach to determine handedness/footedness is used for the rest of the sports. I also wanted to analyse the possible effect of a player's physique on his performance, so I added a variable *height* to measure each player's height in centimeters. Moving on to the actual performance metrics, I carefully selected variables in order to capture a broad variety of actions in tennis. To understand the impact of serve skills on the overall success of a player I decided to include 2 variables: *1st serve won %* depicts the ratio of in-game points won from the given player's 1st serve. This variable will give us a good idea about how consistent the player is in serving, as good servers will oftentimes get themselves into a good position for winning a point with a quality serve. The other serve-related variable is *Ace%*. *Ace%* indicates the percentage of serves that are directly converted into points, without the opponent stroking the ball once. Unlike the aforementioned *1st serve won %* variable, this variable paints a picture of how often a player is capable of producing an almost unstoppable serve. I find the importance of having both variables in the fact that they picture the ability to serve in differing ways. One describes the consistency level of quality serves, i.e., how often does a player serve well enough to gain an in-game advantage as a result. The other one depicts how often is the player able to stand out and produce a top serve. Besides serving itself, a crucial part of a tennis performance is the ability to react well to the opponent's serves and play the ball back in the most beneficial way. For that reason, I included a variable *return points won %*. This variable measures the percentage of points won when receiving a serve from the opponent, which should approximately tell us how often a player is capable of reacting well enough on the incoming serve. To cover a player's capacity of hitting volleys and reacting near the net I included my last variable which is *Net points won %*. It illustrates the percentage of points a player has won when playing near the net. As volleys and other techniques used in this area are incredibly difficult to intercept, the percentages for top players are naturally high.

The summary statistics provide us with a clear picture of the nature of our variables. It is evident that the dependent variable *ATP points* has a large range and therefore we can assume that the performance differences are quite large even for the top 100 players. Our main variable of interest *leftright* shows us via its mean value that there are 18 left-handed players in the dataset, which is greater than the proposed population rate of around 10%. In terms of *height*, the mean and median values are very similar, around 186 centimeters, which is distinctly more than the average male height in the general population (around 175 cm). While the minimum value in the sample of 170 centimeters is still relatively common, the maximum value of 211 centimeters is extremely rare among humans. All the height information put together shows that top male tennis players are, on average, pronouncedly taller than the average male. We could therefore hypothesise that there is some performance advantage stemming from being tall. Moving on to the performance metrics, the conclusively lowest success rates are in the *Ace%* variable. This makes sense as the serve needs to be extraordinarily good to have a realistic chance of becoming an ace. The mean value of ace success rate is at 7% while the maximum is at a whopping 24%, confirming the general opinion that some players are so-called "Ace specialists". When any in-game points won from the player's 1st serve are concerned, the success rate grows drastically, with both the mean and median values being around 71%. As the minimum value of 60% and the maximum value of 81.10% are both relatively near the mean value, we can see that there is not too much variation regarding this variable. It is also worth mentioning that the values for all the players in the sample are larger than 50%, meaning that serving is an advantage for any top male player in tennis. This notion is only confirmed by the success rates of points won from returns, as all of the values in the dataset for *return points won %* are below 50%. Interestingly though, just by looking at the summary statistics, we can clearly notice that *return points won %* has by far the lowest range of values, with mean and median being almost identical at around 37.2% and the minimum and maximum value being at 28.30% and 44.40% respectively. There is a slight issue with the *Net points won %* variable, as there are 20 missing values out of 100 observations in total. Since the values are missing completely at random (MCAR) and 80% of observations are complete, the statistical analysis will remain unbiased and I ,therefore, decided to use it for the performance comparison between lefties and righties. However, I did not include the variable in the model, as the missing values may significantly affect the estimate of the variable. There is quite a large variance for *Net points won %*, with the mean and median being around 65% and the sample ranging from 40% (minimum) up to 80% (maximum).

	1st serve won %	return points won %	height	ATP points
Min.	60.0000	28.3000	170.0000	759.0000
1st Qu.	68.0800	36.0000	183.0000	860.0000
Median	70.9500	37.2000	185.0000	1147.0000
Mean	71.1600	37.2900	186.9000	1778.0000
3rd Qu.	74.7000	39.0200	193.0000	1938.0000
Max.	81.1000	44.4000	211.0000	11540.0000

Table 3.1: Summary Statistics for tennis (a)

	Ace %	Net points won %	**leftright**
Min.	1.0000	40.0000	0.0000
1st Qu.	4.6250	60.6700	0.0000
Median	6.6000	65.4000	0.0000
Mean	7.7010	64.1400	0.1800
3rd Qu.	9.9000	68.3300	0.0000
Max.	24.1000	80.0000	1.0000

Table 3.2: Summary Statistics for tennis (b)

To avoid multicollinearity issues in my model, I constructed a correlation matrix of all the variables mentioned above. The largest correlation is present between all the combinations of *Ace%*, *1st serve won %* and *height*. This comes as no surprise since the percentage of successful aces and the percentage of points won after the first serve both measure a player's ability to serve well. It is also a common assumption that taller players usually serve better as they have higher reach and can therefore generate more downward directing power. Since I want my models to be the Best linear unbiased estimators (BLUE) I opted to avoid violating the "No multicollinearity" assumption by constructing 3 different models, each containing one of the 3 highly correlated variables, and decided to present the best fitting one (which turned out to be the one with the *1st serve won %* variable).
I decided to keep all the independent variables (except the incomplete variable *Net points won %*) as none of them are as lowly correlated with the dependent variable to the extent that it should be deemed insignificant for the analysis. The variable *leftright* possesses the lowest correlation with the dependent variable, but since it is the main variable of our interest I kept it in the dataset.

Table 3.3: Correlation matrix for tennis

	ATP pts	1st serve w %	return points w %	height	Ace%	**leftright**
ATP pts	1.0000	0.3929	0.3422	0.2428	0.1880	-0.1101
1st serve w %	0.3929	1.0000	-0.4477	0.7235	0.8510	-0.2368
return points w %	0.3422	-0.4477	1.0000	-0.4366	-0.6380	0.0687
height	0.2428	0.7235	-0.4366	1.0000	0.7869	-0.2270
Ace%	0.1880	0.8510	-0.6380	0.7869	1.0000	-0.2095
leftright	-0.1101	-0.2368	0.0687	-0.2270	-0.2095	1.0000

The second dataset gathered for tennis is a list of total of 190 head-to-head records of the top 20 best-ranked players according to the year-end ranking in 2021. Since the year-end world rankings are calculated based on performance from the previous year I only included matches that took part in 2021. It is important to point out that not all players played against each other.
The structure of the dataset is relatively simple: it consists of four columns, in the first two columns there are the names of the players, in the third column is the number of wins of the first player against the second player and vice versa for the fourth column. The data have been modified to this form in order to be suitable for the Bradley-Terry model. In addition to this dataset, I also used the *leftright* variable which enabled me to use the predictive extension of the

Bradley-Terry model. The Bradley-Terry model along with its extension will be explained in detail in the Methodology section.

3.1.2 Data for Football

The second sport I analysed in my work is football. In terms of data, there are in total 6 datasets regarding football. The 6 datasets can be divided into 2 categories: Firstly, I gathered 5 datasets that were later utilized for the Ordinary Least Squares Regression, namely offensive player statistics for the top 100 goalscorers in each of the top 5 European leagues (Premier League, Serie A, La Liga, Bundesliga and Ligue 1). All of these datasets were collected from a football data site Whoscored.com and the data used in my thesis are from the 2021-2022 season. The footedness information was gathered from Sofifa.com (a webpage focused on the football video game FIFA).

As the dependent variable, I decided to use *GpG*, which stands for goals per game. This variable was not originally available on the site, but since I had both the *Goals* variable as well as *Apps* (Appearances) I simply divided goals by the number of appearances.
The same procedure was done with the variable *Assists* to obtain *ApG* (assists per game). As in all the sports examined in the thesis, the main variable of our interest is the one determining the footedness of the player. In this case it is the *Foot* and, similarly to tennis, assigns 1 if a player is left-footed and 0 if he is right-footed. Another independent variable is *SpG* which tells us how many shots per game the given player had on average. Next variable in the dataset is *Drb*, which denotes the number of times a player has, on average, beaten an opponent while in possession of the ball in a single game, i.e. dribbles per game. Continuing with the independent variables, the next one is *Fouled*, which depicts the number of times a player is fouled per game. Usually, players with high dribbling ability are fouled more often as it is difficult to stop them legally. Connected to the dribbling ability and ball control skills is also the variable *UnsTch* which shows the number of unsuccessful first touches a player has per game (i.e. how often does a player lose possession after his first touch). Variable *Disp* is another measure of a player's technical ability and measures the number of times a player loses the ball per game while dribbling. It is crucial to mention that even though the last two variables are presented as negative, the best dribblers in the world usually have among the highest

values in the aforementioned variables. It is simply because they are prone to try dribbling even in very tight and risky situations. The last variable in the dataset is *Off* which denotes how many times per game a player is found offside.

When looking at the descriptive statistics, we can see that the values are rather similar for some of the variables across all of the top 5 leagues. Starting with the dependent variable *GpG*, the minimum value for all the leagues is below 0.01 goals per game (the lowest being in La Liga (0.06452)) and the maximum value is in the range of 0.6 - 1.1 goals per game across the leagues (Bundesliga having the highest value at 1.0294 goals per game). For all the competitions the mean and median values are around 0.2 goals per game. Moving onto the variable of our interest, the mean values of the *Foot* variable show us the percentage of left-footed players among top goalscorers in each of the top 5 leagues. Out of the respective top 100 goalscorers, there are 29 left-footers in La Liga, 25 in the Premier League and 23 in each of the remaining leagues (Serie A, Bundesliga, Ligue 1). All of these values depict an apparent overrepresentation of left footers. As far as shots per game (*SpG*) are concerned, the minimum value across the leagues is in the range of 0.2 - 0.4 shots per game and the maximum value is around 4, with a slight outlier being the Bundesliga (maximum of 4.7 shots per game). The mean and median values in all of the leagues are around 1.5, with Premier League having the highest values (mean = 1.658; median = 1.6). Overall, we can conclude that the top scorers across the leagues are getting into shooting positions at a similar rate. A closer look at the chance creation statistic *KeyP* (key passes per game) reveals that the top scorers again performed very similarly across the leagues. The minimum value for all 5 samples was around 0 (with Bundesliga and La Liga actually having a player with 0 key passes in the sample). The maximum value was located around 3 key passes per game (the best creator of chances was from Ligue 1 with 3.2 key passes per game). The mean and median values in all the samples were between 0.7 and 1.1 key passes per game. Considering the best and the worst creators we can see that there is quite a lot of variety among the best scorers in terms of creating chances. To examine not only the number of chances created but also the quality of the chances, it can be useful to take a closer look at assists per game (*ApG*). The worst assist maker in each of the top 5 leagues had no assists over the course of the season. The best assist makers in Serie A, La Liga and Premier League had close to 0.4 assists per game, while in Ligue 1 the maximum value for *ApG* was 0.5385 and in Bundesliga it was even higher at 0.5625 assists per game. In all of the samples, the average assist maker provided around 0.1 assists per game. The difference between the mean (resp. median) values and the maximum values in each sample may suggest that in every league there are a few players who dramatically overperform the rest in the assists department. Dribbles per game (*Drb*) statistics show us that the worst dribblers in all the leagues completed a little more than 0 dribbles per game (in some samples even 0). The best dribblers across the leagues performed around 3 dribbles per game, except for the Premier League (4.3) and the Bundesliga (2.3). The Premier slightly outperforms the other leagues in terms of dribbling, with the mean and median values for the English top flight being 0.934 and 0.9 respectively. The other 4 leagues have both mean and median around 0.8 dribbles per game, with the median always being a little lower than the mean. The English dominance in the dribbling department is in accordance with the popular belief that the Premier League is the most dynamic league in Europe. The descriptive statistics also tell us that the top scorers are *Fouled* at a similar rate in all the aforementioned leagues. The least fouled players across the leagues are fouled in the range of 0.1 - 0.2 fouls per game, with the Premier League's least fouled player not having been fouled at all. The most fouled players in all the samples were fouled close to 3 times per game. The mean and median values for players being fouled per game are ranging from 0.8 to 1.15 fouls per game. An important aspect of every offensive player is not being caught offside during the attacking phase. For that reason, it is worth analysing the offside rates (*Off*) in our samples. The most cautious player in each league has not been caught offside all season. The least cautious players across the leagues were caught offside around once per game. The average players in the samples were caught offside around 0.2 - 0.25 times per game. The safest players with the ball have not been (*Disp*) dispossessed at all. The most reckless players when dribbling in Ligue 1, La Liga and Premier League were dispossessed around 3 times per game, whereas for Bundesliga and Serie A the maximum values are close to 2.2 times per game. As the Premier League is deemed the fastest, most dynamic league, it comes as no surprise that the players from the Premier League sample lose the ball while dribbling, on average, more often than in the other 4 leagues (around once per game). In the rest of the leagues, the average player loses the ball while dribbling somewhere between 0.8 and 0.95 times per game. Another way of losing the ball is by having a bad first touch (*UnsTch*). The most reliable players when receiving the ball across the leagues have between 0 and 0.5 unsuccessful touches per game. In Premier League, Serie A and Ligue1 the least reliable players when receiving the ball have around 4 unsuccessful touches per game, while in Bundesliga the maximum value is at 4.5 unsuccessful touches per game and in La Liga, it is at whopping 5.9 unsuccessful touches per game. The players in each sample have, on average, very similar first touch success with the mean and median values ranging from 1.5 to 1.8 unsuccessful touches per game.

A glance at the 5 correlation matrices (each for one league) would tell us that the majority of the relationships are very similar across the leagues. The *GpG* variable has a sufficient relationship with all the independent variables to not exclude any of them from the model, with the only exception being the *Foot* variable. But since we are interested mainly in the footedness effect on goalscoring I naturally kept the variable in the model. In the Premier League and the Ligue 1 dataset, the variables *Drb* and *Disp* have a dangerously high correlation coefficient. This comes as no surprise since players who dribble often tend to lose the ball more often simply for the sheer volume of their take-ons. The *Disp* variable also has an alarmingly high relationship with the *UnsTch* variable in each of the datasets. Therefore, I decided to omit the *Disp* variable from all the models. The variables *Keyp* and *ApG* have a dangerously strong relationship in all the datasets except the Premier League one. But since I need only one chance-creating metric in my model I decided to omit the *ApG* variable in all 5 respective models. Other than that there are no problematically high correlations in any of the datasets.

As there are a lot of tables for descriptive statistics and correlation matrices for football, I decided to put them in the Appendix for better clarity.

The second category includes a dataset consisting of all of the match results from the 2022 FIFA World Cup in Qatar. The data were collected from sports statistics and results site Livesport. Likewise as in tennis, not all the teams participating have played against each other. Since some playoff matches end in extra time or even by penalty shootout I want to clarify that in my dataset I count overtime victories as regular wins, to simplify the procedure. The structure is exactly the same as for tennis: there are four columns, the first two describing the two teams that played against each other, the third column assigns the number of wins of the first team against the second team and the last column is vice versa. There is also an extra list of all the teams from the tournament and their average number of left-footed players per game (the number of left-footed players that actively took part in the game).

The footedness information of players for the second dataset was also obtained from Sofifa.com.

3.1.3 Data for Handball

Another team sport that I analysed is handball. Specifically, I looked at data from the 2022 EHF European Championship. The data were gathered from the official tournament webpage ehfeuro.eurohandball.com. My dataset contains selected performance metrics on the top 111 goalscorers at the tournament.

The dependent variable in my model is *GpG* which, exactly as in football, stands for goals per game. Unlike in football, this variable is contained in the original dataset.
For determining handedness I created a variable *Hand* which assigns 1 if a player is left-handed, otherwise, it assigns 0. As the information on players' handedness is not contained in any of the datasets that I searched in, I watched each of the 111 players in footage freely accessible on a streaming platform Youtube to determine the handedness of every single player. Another variable in my dataset is *Games* which simply tells us the number of games a player has played at the championship. As far as shooting metrics are concerned, I have decided to include two variables related to shooting: Firstly, there is *SpG* which denotes the number of shots a given player had on average in a game. I created this variable by dividing *Shots* by *Games*. The second shooting-related variable is *Efficiency* which shows what percentage of a player's shots results in goals. The final independent variable in my dataset is *ApG* which stands for assists per game. This variable was added to explore a player's offensive contribution outside of scoring.

Looking at the descriptive statistics, the dependent variable *GpG* has its mean and median values very close to each other, with the former being 3.582 and the latter being 3.14. The variance is quite large for this variable, with the minimum value at 1.44 and the maximum at 9. Reading into this information we can expect the majority of the players in our dataset to have scored less than 4 goals per game with the most productive players significantly outperforming the rest. It just goes to show that even at top-level handball there are great differences in terms of goalscoring ability. Looking at the most important variable for my analysis, the mean value of the *Hand* variable points out that 34.23% of players in the sample are left-handed, which is equal to 38 players. As expected, the *Games* variable has a higher variance due to the fact that some teams are eliminated at earlier stages than others. The median value (7) is slightly larger than the mean (6.532) meaning that there are more players with noticeably fewer games than the average. It comes as no surprise that the minimum value is 3 games since every participant plays 3 games in the preliminary round. The maximum value is 9 as teams that managed to get to the playoff semifinals would have played that amount of games. The *SpG* variable continues the trend of a large range, with the minimum and the maximum being very far from each other, with the former at 1.444 and the latter at 12.4. This would make sense as some positions in handball are purposely situated on the court in a way that enables them to get more shots on target. As handball is a relatively productive sport when it comes to goalscoring, it is expected that the efficiency of shots would be quite high. The statistics for *Efficiency* are consistent with the aforementioned assumption as both the mean and the median are above the 50% mark (mean at 67.44%; median at 68%). A minimum value of 38.9% just goes to show that even the least efficient players in the dataset are scoring more than a third of their shots. Some players in the dataset have scored from every shot they directed towards the goal, making the maximum efficiency value 100%. As far as creating goals is concerned, naturally, some players had not registered a single assist throughout the tournament, making the minimum value for assists per game (*ApG*) equal to 0. The mean (1.686) and the median (1.33) are both around the 1.5 assists per game value, meaning that the average providers in the championship will be assisting at a rate close to these values. The best provider in the dataset has an assisting rate of 6.57 assists per game, drastically exceeding the average.

	Hand	Games	Efficiency (%)	ApG	GpG	SpG
Min.	0.0000	3.0000	38.9000	0.0000	1.4400	1.4440
1st Qu.	0.0000	5.0000	59.5500	0.4300	2.2550	3.4640
Median	0.000	7.0000	68.0000	1.3300	3.1400	5.0000
Mean	0.3423	6.5320	67.4400	1.6860	3.5820	5.3520
3rd Qu.	1.0000	8.0000	75.6500	2.5700	4.7250	6.6190
Max.	1.0000	9.0000	100.0000	6.5700	9.0000	12.4000

Table 3.4: Summary Statistics for handball

By looking at the correlation matrix we can see that all the independent variables have a relevant relationship with the dependent variable *GpG*. The only exception to this trend is the handedness variable, but since it is the object of interest I decided to keep it in the regression. Overall, it can be stated that no 2 independent variables have dangerously high correlation coefficient between them in absolute value, meaning that there is no issue regarding multicollinearity in the dataset. Therefore, I decided to include all the performance-related variables (on per game basis) along with the handedness variable in the model.

Table 3.5: Correlation matrix for handball

	Hand	Games	Shots	Goals	Efficiency	ApG	GpG	SpG
Hand	1.0000	0.1040	0.0889	0.1079	0.0382	-0.2403	0.0463	0.0235
Games	0.1040	1.0000	0.3664	0.3201	-0.0993	-0.1695	-0.4986	-0.4721
Shots	0.0889	0.3664	1.0000	0.9278	-0.2021	0.2721	0.5160	0.6059
Goals	0.1079	0.3201	0.9278	1.0000	0.1518	0.1501	0.6235	0.5768
Efficiency	0.0382	-0.0993	-0.2021	0.1518	1.0000	-0.3129	0.2509	-0.1054
ApG	-0.2403	-0.1695	0.2721	0.1501	-0.3129	1.0000	0.2440	0.3695
GpG	0.0463	-0.4986	0.5160	0.6235	0.2509	0.2440	1.0000	0.9250
SpG	0.0235	-0.4721	0.6059	0.5768	-0.1054	0.3695	0.9250	1.0000

3.1.4 Data for Basketball

The last team sport that is a subject of my thesis is basketball. For basketball I collected a wide range of performance-related data on all the NBA players (605 in total) from the 2021/2022 season. The data were taken from the official NBA webpage (nba.com). The handedness information on each player was gathered from a basketball data specialized webpage basketball-reference.com. The dataset contains many different variables, but since many of them are either irrelevant to a large extent or a variation of a different, more telling variable, I decided not to consider these sorts of variables for the analysis. To even out the distorting effect of a superior number of games played I decided to convert all the variables used in the work into per-game form.

The dependent variable is PPG which stands for points per game.
Moving on to the independent variables, $FGMPG$ is an abbreviation for field goals made per game and it indicates how many times the player shoots the ball through the basket per game. $FGAPG$ tells us how many field goals the given player attempts per game. The ratio of the latter two variables is FG_pct (field goal percentage). My dataset also contains a variable $FTMPG$ (free throws made per game) which indicates how many free throws the player scores per game. This variable may serve as a good proxy of the player's ability when not directly pressured by his opponents. To determine how often the player retrieves the ball after a missed field goal I have included $REBPG$ (rebounds per game). Variables $OREB$ and $DREB$ help us distinguish between offensive and defensive rebounds. As another measure of offensive capability, I added the variable $ASTPG$ which shows the number of assists the player registered per game. To analyse how sloppy players can be with the ball, $TOVPG$ (turnovers per game) signals how often the player loses a ball before his team attempts a shot. The last two independent variables serve for estimating the player's defensive ability. The first of them is $STLPG$ (steals per game) indicating how often the defensive player causes a turnover by legally taking the ball from the opponent. The other defensive metric is $BLKPG$ (blocks per game) which displays how often the player deflects an attempted shot at his team's basket, i.e. blocks.

To get a general idea about the nature of the aforementioned variables I created the summary statistics chart. The dependent variable PPG (points per game) has a mean of 6.919 and a median of 8.236, indicating that there are more extremes at the bottom half of the distribution, i.e. the distribution is negatively skewed. As some players did not score any points in the season the minimum value is 0. The maximum value is 30.574. Since we know that some players did not score any points it is logical that the minimum value of $FGMPG$ (field goals made per game) is also 0. The mean and median values are near each other, with the form at 3.035 and the latter at 2.565. The maximum value for $FGMPG$ is 11.429, which would be considered an outlier. All the values for $FGAPG$ are naturally higher than for $FGMPG$ as a player needs to attempt a field goal in order to score. Specifically, the mean and median values are 6.72 and 5.5 respectively. The minimum value is 0, meaning that some players have not registered a single shot from open play in the whole season. The maximum value is 21.804 which would again be considered an outlier. As some players have not scored a single point the FG_pct (field goal percentage) has its minimum at 0. Both the mean (43.19%) and the median (44.1%) are below the 50% mark. This finding indicates that the average players in the dataset miss more than half of their shots. The maximum value of the variable is 100% meaning that at least one player in the dataset has scored all of his shots in the season. The $FTMPG$ values show that the average players in the 2021/2022 NBA season had around 1 free throw scored per game (mean = 1.2513; median = 0.9375). The best free-thrower of the season had an average of 9.6176 free throws scored per game. The minimum value for free throws was 0 as not every player got to shoot one. As far as rebounds per game ($REBPG$) are concerned, average players in the dataset have around 3 rebounds per game, which is depicted by the mean (3.445) and the median (3). As some players have not registered a single rebound all season the minimum value is 0. The maximum value for rebounds per game is 14.667. When distinguishing between offensive ($OREBPG$) and defensive ($DREBPG$) rebounds interesting findings can be discovered. It is not surprising that both variables have their minimum values equal to 0 (exactly as rebounds per game). However, mean ($DREBPG$ 2.621; $OREBPG$ 0.8242) , median ($DREBPG$ 2.357; $OREBPG$ 0.5968) and maximum ($DREBPG$ 11.015; $OREBPG$ 4.5921) values are all significantly larger for defensive rebounds. It can therefore be concluded that defensive players catch the ball more often after a missed shot. Compared to points per game, the statistics for assists per game are relatively lower ($ASTPG$, with a mean of 1.866 and a median of 1.211). The best assister in the dataset has recorded 10.8 assists per game and the lowest value is 0 as some players have not assisted at all. Considering how often the player loses the ball, the mean (1.0141) and median (0.8085) values for turnovers per game ($TOVPG$) are both around 1. Some players have not lost possession all season, making the minimum value equal to 0. The most unreliable player in possession has 4.4923 turnovers per game. Looking closer at the defensive metrics, namely steals per game ($STLPG$) and blocks per game ($BLKPG$), gripping results can be observed. The minimum value for both is 0 suggesting that some players have not contributed defensively at all. As far as steals per game are concerned, the mean (0.6069) and median (0.5417) values are both slightly above 0.5. The best stealer in the NBA for the 2021/2022 season had an average of 2.333 steals per game. Continuing with blocks per game, the mean and median values are relatively close to each other considering the variance, with the former being at 0.3692 and the latter at 0.2778. The best blocker of the season had an average of 2.8095 blocks per game. Finally, the variable of interest *Hand* indicates via its mean value that 9.091% of all the NBA players were left-handed, which equals 55 players.

Table 3.6: Summary statistics for basketball

	Min.	1st Qu.	Median	Mean	Max.
PPG	0.0000	3.5140	6.9190	8.2360	30.5740
Hand	0.0000	0.0000	0.0000	0.0909	1.0000
FGMPG	0.0000	1.3090	2.5650	3.0350	11.4290
FGAPG	0.0000	3.2260	5.5000	6.7200	21.8040
FG_pct	0.0000	39.3000	44.1000	43.1900	100.0000
DREBPG	0.0000	1.2860	2.3570	2.6210	11.0150
REBPG	0.0000	1.7500	3.0000	3.4450	14.6670
OREBPG	0.0000	0.3333	0.5968	0.8242	4.5921
ASTPG	0.0000	0.6000	1.2110	1.8660	10.8000
TOVPG	0.0000	0.4694	0.8085	1.0141	4.4923
STLPG	0.0000	0.3125	0.5417	0.6069	2.3333
BLKPG	0.0000	0.1176	0.2778	0.3692	2.8095
FTMPG	0.0000	0.4000	0.9375	1.2513	9.6176

A glance at the correlation matrix suggests that all the independent variables (except the *Hand* variable) have a high enough correlation with the dependent variable *PPG* to be kept in the model. As in the previous sports, the *Hand* variable is also kept in the model due to it being the main focus of the analysis. There can be found a few examples of dangerously high correlation coefficients in the dataset. Because I want to avoid multicollinearity issues a thorough inspection of the matter was necessary. Both field goals made *FGMPG* and field goals attempted *FGAPG* were highly correlated with defensive rebounds *DREBPG*, assists *ASTPG*, turnovers *TOVPG* and free throws made per game *FTMPG*, I decided against including these two variables in the final model. To incorporate a set of variables that would cover a wide range of in-game basketball actions I decided to add both offensive *OREBPG* and defensive *DREBPG* rebounds in the model. To assess the offensive capabilities of a player when not scoring points I also included the assists variable *ASTPG*. The accuracy of the player's shot should logically influence his scoring rates, therefore the variable *FG_pct* field goal percentage was included in the model. As it is a unique way of scoring points I also added the free throws variable in the model. Finally, to cover the defensive actions of a player and interpret its effects on the scoring rates I decided that the model should contain *STLPG* and *BLKPG* variables.

I decided not to include the table for the correlation matrix for basketball since there are too many variables and the table therefore would not fit nicely in the text.

3.1.5 Data for Table Tennis

The table tennis data are again divided into two separate categories: The first one is a list of the top 100 ranked players according to ITTF's year-end ranking from 2022. The data were gathered from the official ITTF site ittf.com. Since specific performance metrics are generally not analysed and therefore are hardly accessible, I decided to simply inspect the frequency of left-handers among the top 100 players as well as their average ranking. The handedness information of players was collected from tabletennis.guide.

The second dataset is a list of head-to-head encounters between the top 20 players in 2022. The dataset was assembled from a table tennis data site tabletennis-reference.com. It has a very similar shape to the one used for tennis, with the first two columns denoting the names of the players. On this occasion, however, the last two columns are indicating the number of sets won by the given player. I decided on this change since there are not as many table tennis games in a year and the dataset including strictly games would contain very little information.

3.2 Hypotheses for each sport

In this section, I will briefly explain all the hypotheses I found relevant and therefore decided to test in my thesis. For each sport, the overrepresentation hypothesis will be tested along with other hypotheses specific to the given sport.

3.2.1 Hypotheses regarding Tennis

In tennis, there are several intriguing questions I wanted to answer via my methods. Firstly, a possible overrepresentation of left-handers among the top 100 ranked players compared to the general population was tested via one-sided binomial test. This test was conducted to analyse if my findings would support the results from the analysis by Loffing, Hagemann et al (2012) which stated that the left-handed advantage in tennis has diminished over time and nowadays is not present. This finding would also be consistent with the hypothesis proposed by Loffing (2017) stating that the left-handed advantage is not present in tennis as tennis is not regarded as a "high time pressure sport". Secondly, through a linear regression, I tested if there is a significant performance effect caused by left-handedness, specifically if being left-handed, on average, increases the player's win record, which is mirrored in his ATP points score for the given season. A more specific tool for analysing the (potential) left-hander advantage is used afterwards.
Since the *ATP points* are not normally distributed I could not use an ordinary t-test. Instead, I used a one-tailed Mann-Whitney U test to test if left-handers have a higher probability of having more *ATP points* than right-handers (by comparing their respective medians). The assumptions of the Mann-Whitney U test are met since both groups are independent random samples, the dependent variable *ATP points* is continuous and the independent variable *leftright* consists of two categorical, independent groups (left-handers and right-handers). The Mann-Whitney U test was also used to determine a possible left-handed ad-

vantage in different performance metrics such as net points, returns or aces. To test if there is a left-handed performance advantage regarding serves I used the t-test since the *1st serve won %* variable is normally distributed. All the other t-test assumptions are also met (lefties and righties have the same variance, the variable is continuous and the two groups are independent and normally distributed).

Because I wanted to test the hypothesis made by del Corral and Prieto-Rodriguez (2010), which declares that left-handed players have a higher chance of beating higher-ranked right-handers in direct contests, the Bradley-Terry model was used among the top 20 players to check if their ability scores obtained from comparing their head-to-head encounters would differ in rank with the player's standard ATP rankings. Concretely, the possible upward/downward movements on the ranking scale of left-handed players were investigated. Afterwards, a predictive extension of this Bradley-Terry model was run to analyse the effect of the *leftright* variable on the ability coefficients of the players. All the Bradley-Terry assumptions were met, hence there was no significant problem with the method.

3.2.2 Hypotheses regarding Football

To test the findings from Loffing et al (2016), which suggest that left-footers are clearly overrepresented in football with a frequency of 20%, I ran a one-sided binomial test to see if there was a significant left-footed overrepresentation in my sample.

A linear model was also conducted to analyse a potentially significant effect of left-footedness on goalscoring rates. Testing the left-footed impact on a specific performance metric (such as goalscoring) would extend the current knowledge of the laterality effects in football.

Afterwards, to determine if left-footers are better goalscorers, I compared the average goalscoring success between left-footers and right-footers via a one-tailed Mann-Whitney U test. The test's assumptions are met because the GpG variable is continuous, the *Foot* variable is categorical and is made of two independent, randomly sampled groups. All the aforementioned ideas and hypotheses were tested on all 5 of the top 5 league samples. The same procedure was also conducted for other performance metrics (ApG, Drb, etc.).

The predictive Bradley-Terry model, with footedness being the predictor variable, was run on the match results from the 2022 FIFA World Cup to examine the impact of having more left-footed players on the results (ability scores). The result could indicate if teams benefit from having more left-footed players on the pitch.

3.2.3 Hypotheses regarding Handball

Results from an analysis conducted by Loffing et al (2015) were put to test in my work. The results claimed that left-handed handball players are regularly overrepresented among top goalscorers in international tournaments. To check the validity of the aforementioned claim I ran a one-sided binomial test to see if there was any significant left-handed overrepresentation in my sample. It would also serve as an extension to the study performed by Laxdal et al (2022) which revealed that left-handers are overrepresented among top 7-meter shooters.
Furthermore, to test if left-handers are, on average, ranked higher among top goalscorers I ran a one-tailed Mann-Whitney U test. The assumptions for the test are met as the *GpG* variable is continuous, the *Hand* variable is categorical and consists of two independent groups that are sampled randomly. The same approach was used to test left-handers' performance in assists and shots per game. For shooting efficiency, I opted for testing the (potential) left-handed advantage via t-test since the *Efficiency* variable is normally distributed. All the other t-test assumptions hold as well because lefties and righties are both normally distributed and independent of each other, the two groups have approximately the same variance and the *Efficiency* variable is continuous.
Finally, a linear regression model was computed to estimate the effect of left-handedness on goalscoring rates. This model could contribute to the currently limited knowledge of handedness impact on goalscoring in handball.

3.2.4 Hypotheses regarding Basketball

To test the hypothesis by Lawler and Lawler (2011) which declares that left-handed basketball players performed better at rebounds, assists, points per game and field-goal percentage, I constructed the Mann-Whitney U test for all of the four variables mentioned above with handedness being the independent variable. The same test was also run for other performance metrics to explore different possibilities of the left-handed advantage. For all the tests, the Mann-Whitney assumptions are met as all the dependent variables are continuous, handedness is categorical and both lefties and righties are randomly sampled.
As a next step, I ran the one-sided binomial test to check if there is a significant

overrepresentation of left-handers in the NBA, which would refute the findings discovered by Loffing et al (2016).
Finally, a linear model with the dependent variable being points per game *PPG* was created. One of the independent variables in the model was *Hand* as the purpose of the model is to estimate a possible significant effect of left-handedness on the scoring rate.

3.2.5 Hypotheses regarding Table Tennis

As table tennis is regarded as a "high time pressure sport" I decided to investigate the hypothesis made by Loffing (2017) which states that left-handers are overrepresented in high time pressure sports. To achieve that, I used a one-sided binomial test on my sample of the top 100 table tennis players from 2022.
In addition, I compared the mean rankings of lefties and righties to check if left-handers are, on average, performing better than right-handers.
FInally, as I was curious about the impact of left-handedness in direct confrontations among the top 20 players I created a Bradley-Terry model that estimated the ability score for the top 20 players. I then compared the rankings based on the ability scores with the official year-end rankings to see if left-handedness is an advantageous factor when facing the very best players. This model was later extended to its predictive version to estimate the actual effect of left-handedness on the winning success.

3.3 Methodology

In this section I will go over the methods that I used in my thesis as well as all the sports and datasets the aforementioned methods were applied on. For each specific model or test the assumptions of the particular method were checked. In case of an assumption being violated an appropriate measure is applied to deal with the violation.

3.3.1 Ordinary Least Squares and Feasible Generalised Least Squares

The first method used to analyse the relationships between the dependent and independent variables is the Ordinary Least Squares method. I want my mod-

els to be the best linear unbiased estimators (BLUE). To achieve that, the model needs to satisfy the Multiple linear regression assumptions. (Wooldridge (2016)) In the data description part of the thesis I already discussed possible multicollinearity issues among variables and I constructed the model accordingly. For each model, I also calculated the variance inflation factors and all of them were significantly below the critical value of 10. Therefore the "No Multicollinearity" assumption holds for all the models. Since my thesis is focused on analysing top athletes in their respective sports, the sampling of my data was not done completely at random. I always selected a certain variable as a proxy of performance level to determine the top N number of players in that particular sport. However, even by sampling the data in the aforementioned way the variation for both dependent and independent variables is usually large. Therefore we can assume that the samples are representative of the focus group and there is no bias present. The parameters for all the models are linear. This assumption was also checked by plotting residuals against fitted values. (Bartell (2019)) Since the relationship for the majority of the respective plots is mostly linear, the assumption is confirmed. In the few cases where the relationship is not linear, I inspected the relationship between the dependent variable and the independent variable using a simple scatter plot and adjusted the functional form accordingly. The "Zero Conditional Mean" assumption was tested by plotting the residuals in a normal Q-Q plot. (Bartell (2019)) For all my models the residuals approximately followed the normal distribution, therefore the assumption holds. Finally, to test possible heteroscedasticity in my models I applied the White test. The OLS model for Premier League (football) proved to be homoscedastic. For all the other models, the p-value shown by the White test was lower than 0.05, signalling that heteroscedasticity is present.
To deal with this issue, I replaced the Ordinary Least Squares models with the Feasible Generalised Least Squares (FGLS). I opted for the FGLS method since the heteroscedasticity robust standard errors method requires a greater amount of observations. The FGLS method deals with heteroscedasticity when the form of heteroscedasticity is unknown. (Wooldridge (2016)) The FGLS estimator can be obtained via the following procedure:

After running the original OLS regression, we obtain the model residuals. These residuals are then squared and the squared residuals are put into a natural logarithm. Afterwards, these logarithms are regressed on the independent variables from the original OLS model. We obtain the fitted values from the latter regression and exponentiate them. We then divide 1 with the aforementioned exponentiated fitted values to create the weights for the upcoming step. Lastly, we run the original linear regression with the newly created weights from the previous step, which together give us the FGLS regression. (Wooldridge (2016))

Tennis FGLS model

For tennis, the "Linearity in parameters" assumption did not hold for the original model, so I looked at the scatter plots to accommodate the functional forms according to the relationships between the independent variables and the dependent variable. Via this method, I added two more variables to the model: *servesquared* (*1stservewonpct* to the power of 2) and *returnssquared* (*returnptswonpct* to the power of 2). The model ,therefore, is not BLUE, it is BUE (best linear estimator).

The FGLS model for tennis can be written in the following form:
$ATP\_p_i = \beta_0+\beta_1 leftright_i+\beta_2 1stservewonpct_i+\beta_3 1stservewonpct_i+\beta_4 returnptswonpct_i+ \beta_5 returnssquared_i + \epsilon_i$

Football Top 5 Leagues model

All the football-related models (OLS for Premier League and FGLS for the rest) contain the same set of variables, so generally, the model is written in the following form:
$GpG_i = \beta_0 + \beta_1 SpG_i + \beta_2 Drb_i + \beta_3 Foot_i + \beta_4 KeyP_i + \beta_5 Off_i + \beta_6 UnsTch_i + \beta_7 Fouled_i + \epsilon_i$

Basketball FGLS model

The FGLS model for basketball is written in the following form:
$PPG_i = \beta_0+\beta_1 Hand_i+\beta_2 DREBPG_i+\beta_3 OREBPG_i+\beta_4 ASTPG_i+\beta_5 STLPG_i+ \beta_6 BLKPG_i + \epsilon_i$

Handball FGLS model

The FGLS model for handball is written in the following form:
$GpG_i = \beta_0 + \beta_1 Hand_i + \beta_2 SpG_i + \beta_3 ApG_i + \beta_4 Efficiency_i + \epsilon_i$

3.3.2 Binomial test

A rather simple method, the binomial test is a crucial part of my analysis. Concretely, thanks to the one-sided binomial test I can check if the left-handed frequency in my sample is significantly larger than the estimated global population rate of 10%. The null hypothesis of the test is that both rates are the same, with the alternative being that left-handers are overrepresented in my sample. The test works as follows:

Let n denote the sample size, x is the number of successes in the sample (i.e. the number of lefties) and p is the assumed population rate (10%). Then the one-sided binomial test calculates the p-value in the following way:

$$\sum_{t=0}^{x} \binom{n}{x} p^x (1-p)^{n-x}$$

If the p-value is less than 0.05 we reject the null hypothesis, otherwise we fail to reject the null hypothesis, meaning that there is no significant difference between the sample ratio of lefties and the aforementioned 10%.
For all the datasets the assumptions are met as there is a binary outcome (either the rate is or is not 10%), the observations are always independent and there is a fixed number of observations for each dataset. (McClenaghan) (University of Texas)

3.3.3 Mann-Whitney U test

Since the majority of the performance variables I chose to investigate the left-handed/footed advantage are not normally distributed, I decided for the one-tailed Mann-Whitney U test to check if lefties perform, on average, better than righties in the specific performance metrics. Specifically, the Mann-Whitney U test tests if the two populations (lefties and righties) have the same median. As the alternative hypothesis, I set left-handers having a greater median, meaning that they perform better on average. The test statistic is calculated as follows:

All the observations from both groups are ranked based on their score in the given performance metric. Each observation is then assigned a number of points based on how many observations from the other group are ranked better. After that, we sum up all the points for observations for one group (lefties) and do the same for the other group (righties). Then the two sums are compared and the smaller one is assigned as U. Let N_L denote the number of lefties in the sample and N_R the number of righties. Then the test statistic is in the following form:

$$z = \frac{U - \frac{N_L N_R}{2}}{\sqrt{\frac{N_L N_R (N_L + N_R + 1)}{12}}}$$

If the z-statistic is greater than 1.65 the null hypothesis is rejected. As discussed in the Data Description part of the thesis, the necessary assumptions for the Mann-Whitney U test hold for every single test computed. (Statistics Lectures)

3.3.4 Two Sample T-test

Since few performance metrics are normally distributed (namely *1st serve won %* in tennis and shot *Efficiency* in handball) the one-sided two sample t-test was used to analyse if the mean value for lefties is greater than for righties. If it was, it would suggest that lefties are performing, on average, better than righties in the aforementioned metrics. The t-test assumptions are met in both cases as the two samples are independent, normally distributed, the variable of interest (*1st serve won %* or *Efficiency*) is continuous and the two groups have the same variance. The test statistic is calculated in the following way:

$$\frac{\overline{X}_L - \overline{X}_R}{\sqrt{\frac{\sigma_L^2}{N_L} + \frac{\sigma_R^2}{N_R}}}$$

where $\overline{X}_L$ denotes the mean for the left-handed population, $\overline{X}_R$ denotes the mean for the right-handed population, σ_L^2 is the variance of the left-handed group, σ_L^2 is the variance of the right-handed group, N_L is the number of players in the left-handed group and N_R is the number of players in the right-handed group. The null hypothesis is that left-handers and right-handers have the same means, with the alternative being that left-handers have the greater mean of the two groups. The null hypothesis is rejected if the following inequality holds:

$$t > t_{1-\alpha,\nu}$$

where t stands for the test statistic, α denotes the significance level (5% in my thesis) and ν indicates the number of degrees of freedom, which is dependent on the number of observations. (Fernandez (2020))

3.3.5 Bradley-Terry model

The Bradley-Terry model is a logistic model that deals with pairwise comparisons. It is based on the maximum likelihood estimation procedure. It takes each possible pair of two objects from a sample of size n and compares them based on their results from the dataset. Let i, j be two players from the sample. Then the probability that i beats j is in the following form:

$P(ibeatsj) = \pi_{ij}$

Since the Bradley-Terry model does not handle ties, the probability that j beats i is:

$P(jbeatsi) = \pi_{ji} = 1 - \pi_{ij}$

The model assumes that for each object i from the sample there is a coefficient β_i that shows the object's (a player, a team, etc.) capability of competing against other objects from the sample. This coefficient is also called the ability score. The ability scores are calculated using the pairwise probabilities in the following way:

$\log(\frac{P(ibeatsj)}{P(jbeatsi)}) = \beta_i - \beta_j$

To precisely calculate the betas, we need to set one of them as a reference (for example β_a). Then the reference beta would be equal to 0.

$\beta_a = 0$

The rest of the betas are calculated in relation to the reference beta. (Hanek (2010)) (Turner and Firth (2012))

I decided for the general Bradley-Terry model to see if left-handed players (in tennis and table tennis) would overperform their year-end ranking when only results between the top 20 players are taken into consideration. If they did overperform, we could hypothesise that left-handedness is a performance advantage among the absolute elite.
The Bradley-Terry assumptions were checked according to the diagnostic frame-

work by Wu et al (2022). Firstly, the model was tested for overdispersion by dividing the residual deviance by the residual degrees of freedom. For the tennis model the test statistic is close to 1, meaning that there is no overdispersion issue in the models. The other 2 Bradley-Terry models in my thesis have somewhat worse overdispersion value. Since there are no subjects in my model (individuals making the comparisons) I only need to deal with the object assumptions. The two object assumptions are the following:

- Independence assumption.
- Functional assumption.

The Independence assumption holds if all comparisons are statistically independent. Since I am dealing with independent matches, this assumption holds. The Functinal assumption states that "the relationship between the results of comparisons and the object scores is correctly specified" (Wu et al (2022)). Violating the functional assumptions means that the model may suffer from a lack of fit.

Both assumptions are checked via the following diagnostic plots:

- Boxplot of object residuals: the two assumptions hold when the distribution of the object residuals is symmetric and does not contain many outliers.
- Normal Q-Q plot of object residuals: The points need to be randomly scattered around the 45-degree line for the two assumptions to be met.
- Object residuals plotted against the Bradley-Terry ability scores: The object residuals need to be randomly scattered around the horizontal 0 line for the assumptions to hold.

Since the results of the diagnostic measures are different across the three Bradley-Terry models I used in this thesis, I decided to comment on each of them specifically in the Results part of the thesis.

Besides the general Bradley-Terry model, I also used one of its extensions: the Bradley-Terry model with predictors. In this case, the predictors are players-specific explanatory variables. Since I am interested in the effect of left-handedness/left-footedness on the winning success, my only predictor variable is the one concerning handedness/footedness in the given sport. Specifically, for tennis and table tennis, the predictive variable is a binary handedness variable. For football (World Cup) the predictive variable is the average number of left-footed players per game.

Chapter 4

Results and Discussion

In this part of the thesis, I will present several results from each sport via tables. The tables will be thoroughly explained to give the reader a clear idea of the results and their meaning.

As there is a large number of tables for Mann-Whitney U tests (and T-tests) I decided to put them in the Appendix so the Results section is more clear.

4.1 Results regarding tennis

4.1.1 Binomial test

After counting the number of left-handed players (18) in the sample of 100 players I conducted a one-sided binomial test to check if left-handers are significantly overrepresented compared to the general population rate of approximately 10%. As the p-value was lower than 0.05 I can safely say that left-handers were overrepresented among the top 100 ATP players. This result contradicts Loffing (2017) who suggested that left-handed overrepresentation in tennis is not present due to it not being a "high time pressure sport".

Table 4.1: Results of the binomial test for tennis

Parameter	**Value**
Number of successes	18
Number of trials	100
Sample estimate of success probability	0.18
P-value	0.0100
Confidence interval (95%)	[0.1197, 1.0000]
Alternative hypothesis	True probability of success is greater than 0.1

4.1.2 Feasible Generalised Least Squares model

In the Data description part of the thesis, I explained my choice of the dependent and independent variables for the model. The original model was a simple Ordinary Least Squares regression. However, since non-linear (specifically quadratic) relationships were discovered with the dependent variable (namely *1st serve won %* and *return points won %*) I created the squared versions of these variables and added them in the model (*servesquared* and *returnssquared*). This model was then, due to heteroscedasticity, reconstructed into Feasible Generalised Least Squares model. We can see that only *return points won %* and *returnssquared* are statistically significant predictors of the *ATP points* variable. This means that in our model, unlike the ability to return the ball well, serving success is not significantly affecting the overall performance level of a top tennis player. Precisely, the estimates for *return points won %* and *returnssquared* are -1344.856 and 21.152 respectively. This means that with the increasing success rate of return points, the ATP points should be surprisingly decreasing up to a certain point and afterwards they should start increasing. By taking the first derivative and setting it equal to 0 I found that the breaking point between decreasing and increasing ATP points is the return success rate of 31.79028 percentage points. The importance of returns and unimportance of serves may be explained by the assumption that at the top level, matches are usually won by winning the games where opponents serve because the majority of the top-level tennis players have a very good serve themselves and therefore can rely on winning the games where they serve. More importantly, however, handedness does not significantly affect a player's performance. This finding is in accordance with the analysis conducted by Loffing, Hagemann, et al (2012) which states that the left-handed advantage in tennis has diminished over time.

The table below shows the estimates of the FGLS regression on the tennis dataset with *ATP points* as the dependent variable.

Table 4.2: Coefficients of the FGLS model for tennis

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	28684.1900	22608.5910	1.2690	0.2077	
leftright	-75.7430	194.3910	-0.3900	0.6977	
1st serve won %	-320.0130	641.6810	-0.4990	0.6192	
servesquared	3.2130	4.5620	0.7040	0.4830	
return points won %	-1344.8560	464.7290	-2.8940	0.0047	**
returnssquared	21.1520	6.5080	3.2500	0.0016	**
*p<0.5	$**p < 0.01$	***p<0.001			
Observations	100				
R^2	0.3938				
Adjusted R^2	0.3616				

4.1.3 Mann-Whitney U tests and T-test

To further inspect the hypothesis by Loffing, Hagemann, et al (2012) I conducted a series of one-sided Mann-Whitney U tests on tennis performance metrics to see if left-handers have a greater median in these variables than right-handers. After looking at the FGLS model it comes as no surprise that the null hypothesis is not rejected for *ATP points* (since the p-value is more than 0.05), meaning that lefties do not have significantly greater median value and therefore do not perform better overall. The same result was discovered after performing the Mann-Whitney U test on more detailed performance metrics, specifically *Ace%*, *Net points won%* and *return points won%*. For *1st serve won%* a t-test was used due to the normality of the variable. The t-test revealed that lefties do not have a significantly greater mean than righties (as the p-value was again more than 0.05) and therefore do not have a higher serving success. Overall, we can conclude that among the top 100 ATP players being left-handed does not provide any significant performance advantage, confirming the hypothesis by Loffing, Hagemann, et al (2012) which states that lefties no longer have an advantage in top-level tennis.

4.1.4 Bradley-Terry model

Firstly, a general Bradley-Terry model was constructed on the head-to-head results between the top 20 ranked players in the ATP year-end ranking from 2021. When the players are ranked based on the Bradley-Terry ability scores (calculated from their direct encounters) it can be seen that all 3 left-handed players in the top 20 have changed their ranking compared to the classic ATP ranking. Nadal has improved by 4 places, going from 6th to 2nd. Norrie has declined by 2 places, falling from 12th to 14th. Lastly, Shapovalov has also ranked worse in the ability scores ranking, dropping from 14th to 17th. These findings are rather mixed and no conclusions can be drawn from them. It is also important to note that the ability score for Nadal is not statistically significant, so the previous analysis serves only as an illustrative preview of a possible handedness effect.

To thoroughly inspect the influence of handedness on ability scores an extension of the Bradley-Terry model with predictors was applied. As the predictive variable, I chose the handedness variable *leftright*. The results show that the handedness effect is not statistically significant and even if it was, the estimate is very small and therefore the impact would be negligible. Altogether, it can be concluded that among the top 20 players, being left-handed does not give a player an advantage in direct encounters against other players. My findings are strongly contradicting those of del Corral and Prieto-Rodriguez (2010) who declare that left-handed players have a higher chance of beating higher-ranked right-handers in direct contests.

There were no major issues concerning the model assumptions as no overdispersion was detected, the independence assumption holds because the matches are independent of each other and the functional assumption holds as the object residuals are scattered around the horizontal 0 line when plotted against the ability scores.

The table below shows the estimates of the ability scores for the Bradley-Terry used on head-to-head contests between the top 20 tennis players in 2022.

Table 4.3: Bradley-Terry tennis results

	Estimate	Std. Error	Pr(>|t|)
..Djokovic	0.0000	0.0000	0.0000***
..**Nadal**	-0.3425	0.7599	0.6522
..Zverev	-0.4919	0.5926	0.4064
..Medvedev	-0.8443	0.6214	0.1742
..Tsitsipas	-0.9326	0.6356	0.1423
..Berrettini	-1.4696	0.7177	0.0406*
..Karatsev	-1.6281	0.7507	0.0301*
..Rublev	-1.7941	0.7221	0.0130*
..Hurkacz	-1.9160	0.7106	0.0070**
..Busta	-1.9415	0.8668	0.0251*
..Sinner	-1.9478	0.7402	0.0085**
..Aliassime	-2.0202	0.7329	0.0058**
..Ruud	-2.0549	0.7246	0.0046**
..**Norrie**	-2.2105	0.7552	0.0034**
..Garin	-2.6761	1.0171	0.0085**
..Federer	-2.7445	1.3659	0.0445*
..**Shapovalov**	-2.7642	0.8095	0.0006***
..Bautista Agut	-2.8889	0.8144	0.0004***
..Thiem	-3.3049	1.3040	0.0113*
..Schwartzmann	-4.4196	1.1884	0.0002***
..**leftright**	-0.0392	0.2801	0.8890
*p<0.5	$**p < 0.01$	***p<0.001	

4.1.5 Discussion

By analysing the left-handed performance of top tennis players in various different ways I came to the conclusion that left-handers do not perform better than right-handers. My work broadens the current knowledge of left-handed performance in tennis as, unlike the majority of the previous studies, I analysed individual performance metrics such as serving, returning, etc (lefties have not overperformed righties in any of them). This is further proven by the finding that lefties do not having a significant advantage in direct matches at the elite

level. However, my results argue against a renowned study by Loffing (2017) which states that lefties are no longer overrepresented at top-level tennis because the binomial test conducted in my thesis suggests otherwise.
Overall, based on my analysis I would hypothesise that left-handers do not have any significant advantage over right-handers at the top level. Nevertheless, I would suggest that being left-handed provides a player a better opportunity of becoming a top-level tennis player, as the overrepresentation is significant. In other words, getting to the top may be easier for lefties, but once at the top, the advantage vanishes.

4.2 Results regarding football - Premier League

4.2.1 Binomial test

There were in total 25 left-footed players in the sample of 100. To compare if this rate is significantly higher than the general population rate a one-sided binomial test was run. As the p-value of the test was lower than 0.05 I can conclude that left-footed players are significantly overrepresented among the top 100 goalscorers in the Premier League. This is in accordance with the findings from Loffing et al (2016) who argue for clear left-footed overrepresentation in football.

Table 4.4: Results of the binomial test for Premier League

Parameter	**Value**
Number of successes	25
Number of trials	100
Sample estimate of success probability	0.25
P-value	1.307e-05
Confidence interval (95%)	[0.1802, 1.0000]
Alternative hypothesis	True probability of success is greater than 0.1

4.2.2 Ordinary Least Squares model

As there was no problem with heteroscedasticity, the original OLS model was kept. The p-values of the estimates indicate that only shots per game (*SpG*) and offsides per game (*Off*) are statistically significant when predicting goals per game (*GpG*). Both shots per game (0.11234) and offsides (0.14150) have a positive effect on goalscoring. It is quite straightforward why shots per game would increase a scoring rate of a player since more shots present more opportunities to score goals. However, the effect of offsides per game on goalscoring is much less obvious. I would hypothesise that it may be a proxy variable for getting into chances since players who make a lot of runs (and therefore get into goalscoring positions) would usually also be caught offside more often. Crucially, the *Foot* variable is not statistically significant, meaning that footedness does not predict the goalscoring rate well in the model. Therefore left-footed players would not have an advantage in goalscoring in our sample. The rest of the independent variables are also statistically insignificant (*UnsTch*, *Fouled*, *KeyP*, *Drb*).

The table below shows the estimates of the OLS regression on the Premier League dataset with *GpG* as the dependent variable.

Table 4.5: Coefficients of the OLS model for Premier League

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.0360	0.0250	1.4430	0.1524	
SpG	0.1123	0.0180	6.2480	0.0000	***
Drb	0.0006	0.0203	0.0300	0.9764	
Foot	-0.0126	0.0189	-0.6650	0.5074	
KeyP	0.0328	0.0209	1.5680	0.1202	
Off	0.1415	0.0481	2.9400	0.0041	**
UnsTch	-0.0240	0.0175	-1.3780	0.1717	
Fouled	-0.0101	0.0191	-0.5260	0.6002	
*p<0.5	$**p < 0.01$	***p<0.001			
Observations	100				
R^2	0.6126				
Adjusted R^2	0.5832				

4.2.3 Mann-Whitney U tests

To analyse the possible left-footed advantage in football more deeply, a series of Mann-Whitney U tests was run on many different performance attributes to cover a wide range of in-game actions. Similarly as in the linear model, the Mann-Whitney U test showed us that lefties were not overperforming at goals per game (p-value greater than 0.05). The p-value was higher than the significance level in all the other performance metrics, meaning that left-footed players have not performed significantly better in assists per game, dribbles per game, shots per game or key passes per game. Considering all the tests together, I can conclude that left-footed footballers do not have a performance advantage among the top 100 goalscorers in the Premier League.

4.3 Results regarding football - La Liga

4.3.1 Binomial test

To see if the results presented by Loffing et al (2016) are supported by data from the Spanish top tier, I examined a potential left-footed overrepresentation among the top 100 goalscorers in La Liga using the one-sided binomial test. As there were 29 lefties among the top 100 goalscorers and the population rate of lefties is approximately 10%, it is no surprise that the p-value of the test was below the 0.05 mark, suggesting a significant overrepresentation of left-footed players among La Liga's top 100 goalscorers.

Table 4.6: Results of the binomial test for La Liga

Parameter	**Value**
Number of successes	29
Number of trials	100
Sample estimate of success probability	0.29
P-value	9.444e-08
Confidence interval (95%)	[0.2159, 1.0000]
Alternative hypothesis	True probability of success is greater than 0.1

4.3.2 Feasible Generalised Least Squares model

The original Ordinary Least Squares model had to be transformed into the Feasible Generalised Least Squares model in order to deal with the present heteroscedasticity. A glance at the estimates and their p-values would tell us that the shots per game variable (SpG) is (again) statistically significant when predicting the goalscoring rate. The offsides per game variable (*Off*) and the dribbles per game variable (*Drb*) are both statistically significant as well.. The shots per game variable has a positive estimate of 0.098240 which makes sense as more shots would naturally help players score more goals. Offsides per game also have a positive estimate (0.100813) as players who are caught offside more often tend to make more runs and, therefore, get into more goalscoring opportunities. Dribbles per game, surprisingly, have an estimated negative impact (-0.036026) on goalscoring. Naturally, more successful dribbles should lead to a player getting into more dangerous areas and therefore having a greater chance to score. However, since we are comparing the best goalscorers, it may be explained by the fact that majority of them have a very good shooting ability and oftentimes dribbling more could potentially lead to a delay of a good shooting opportunity, sometimes even ruining it. However, it is important to note that the estimate is rather small and the relationship may be unique for the given set of players. Our variable of interest *Foot* is not statistically significant, which can be interpreted as footedness (particularly being left-footed) not having a significant impact when predicting goalscoring success. The remaining variables are statistically insignificant.

The table below shows the estimates of the FGLS regression on the La Liga dataset with *GpG* as the dependent variable.

Table 4.7: Coefficients of the FGLS model for La Liga

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.0514	0.0169	3.0490	0.0030	**
SpG	0.0982	0.0169	5.8190	8.55e-08	***
Drb	-0.0360	0.0166	-2.1730	0.0323	*
Foot	-0.0017	0.0150	-0.1160	0.9082	
KeyP	0.0155	0.0201	0.7710	0.4426	
Off	0.1008	0.0491	2.0530	0.0429	*
UnsTch	0.0087	0.0155	0.5610	0.5759	
Fouled	-0.0055	0.0179	-0.3070	0.7595	
*p<0.5	$**p < 0.01$	***p<0.001			
Observations	100				
R^2	0.5113				
Adjusted R^2	0.4741				

4.3.3 Mann-Whitney U tests

A one-sided Mann-Whitney U test confirmed the result from the FGLS model that left-handers do not have an advantage when it comes to goalscoring (p-value was above 0.05). The same result (p-value higher than 0.05) was obtained when testing (via Mann-Whitney U test) the potential better performance of left-footers in dribbles per game, key passes per game, shots per game and assists per game. Therefore we can conclude that lefties do not perform better than righties in any of the aforementioned attributes among the top 100 La Liga goalscorers.

4.4 Results regarding football - Bundesliga

4.4.1 Binomial test

Once again I wanted to explore the hypothesis by Loffing et al (2016), this time on a sample of the top 100 goalscorers from the German Bundesliga. There were in total 23 left-footed players in the sample. When comparing this rate to the proposed global population rate of 10%, the p-value of the binomial test was way below the 0.05 mark, meaning that there is a significant overrepresentation of left-footed players among top goalscorers in Germany. Therefore the findings from Loffing F. et al (2016) are supported once again.

Table 4.8: Results of the binomial test for Bundesliga

Parameter	**Value**
Number of successes	23
Number of trials	100
Sample estimate of success probability	0.23
P-value	0.0001
Confidence interval (95%)	[0.1626, 1.0000]
Alternative hypothesis	True probability of success is greater than 0.1

4.4.2 Feasible Generalised Least Squares model

After transforming the original OLS model (due to heteroscedasticity) into a FGLS model, I took a look at the estimates to examine the effects of independent variables on the dependent variable (*GpG*). As in the previous two models, the shots per game (*SpG*) estimate is statistically significant and positive (0.10972), as more shots would naturally increase a player's chance to score more goals. The variable of interest *Foot* is yet again statistically insignificant, which can be understood as footedness having no important impact on goalscoring in the model, hence being left-footed is not an important factor when goalscoring is considered. *Drb*, *KeyP*, *Fouled*, *UnsTch* and *Off* are all statistically insignificant as well.

The table below shows the estimates of the FGLS regression on the Bundesliga dataset with *GpG* as the dependent variable.

Table 4.9: Coefficients of the FGLS model for Bundesliga

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.0715	0.019218	3.7180	0.0003	***
SpG	0.1097	0.0181	6.0700	2.82e-08	***
Drb	-0.0173	0.0196	-0.8820	0.38011	
Foot	-0.0097	0.0179	-0.5430	0.5882	
KeyP	-0.0137	0.0103	-1.3340	0.1854	
Off	0.0508	0.0354	1.4350	0.1546	
UnsTch	0.0149	0.0152	0.9820	0.3284	
Fouled	-0.0238	0.0142	-1.6770	0.0969	
*p<0.5	$**p < 0.01$	***p<0.001			
Observations	100				
R^2	0.5186				
Adjusted R^2	0.482				

4.4.3 Mann-Whitney U tests

A one-sided Mann-Whitney U test comparing the performance of left-footers and right-footers in terms of goalscoring (GpG) only confirmed the result of the FGLS model, namely that being left-footed does not come with any sort of goalscoring advantage (p-value above the significance value). No significant overperformance of left-footers was found when running the Mann-Whitney U test on different performance metrics, specifically dribbles per game, shots per game, key passes per game and assists per game. Overall, the analysis tells us that left-footers are not performing better than their right-footed counterparts in the aforementioned offensive attributes among the top 100 goalscorers in Bundesliga.

4.5 Results regarding football - Serie A

4.5.1 Binomial test

As per usual, a one-sided binomial test was conducted to test the hypothesis by Loffing et al (2016) which argues for left-footed overrepresentation in football. In the sample of the top 100 Serie A scorers I counted 23 left-footed players in total. This rate was compared via the test with an approximate global population rate of 10%. Since the p-value of the test was below the significance level (0.05) I conclude that left-footers are overrepresented among the top 100 goalscorers in Serie A.

Table 4.10: Results of the binomial test for Serie A

Parameter	**Value**
Number of successes	23
Number of trials	100
Sample estimate of success probability	0.23
P-value	0.0001
Confidence interval (95%)	[0.1626, 1.0000]
Alternative hypothesis	True probability of success is greater than 0.1

4.5.2 Feasible Generalised Least Squares

The Feasible Generalised Least Squares model was used instead of the common OLS model for heteroscedasticity reasons. The model summary shows us that there are 3 statistically significant estimates when predicting goals per game (GpG). Similarly to the previous FGLS football models, the variable shots per game (SpG) is statistically significant and has a positive estimate (0.12243) as shooting and goalscoring are naturally strongly related. The offsides per game variable (Off) is also statistically significant with a positive coefficient of 0.07563. It was previously discussed that the offside per game variable may serve as a proxy to the number of runs made, which would logically lead to more goalscoring opportunities for the given player. The last statistically significant variable is the variable fouled per game ($Fouled$) with a negative coefficient of -0.02837. Its effect on the dependent variable is a bit unclear. However, I would hypothesise that the negative impact may be explained by the fact that players that are less fouled usually tend to withstand punishable offences more. As this often happens near the opposition goal, they, therefore, give themselves more opportunities to score. Most importantly, the footedness variable ($Foot$) is statistically insignificant in this model, meaning that left-footedness does not guarantee players any additional success when attempting to score. The variables ($UnsTch$), (Drb) and ($KeyP$) are all statistically insignificant as well.

The table below shows the estimates of the FGLS regression on the Serie A dataset with GpG as the dependent variable.

Table 4.11: Coefficients of the FGLS model for Serie A

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.0423	0.019	2.231	0.0281	*
SpG	0.1224	0.0153	7.979	3.98e-12	***
Drb	-0.0162	0.0188	-0.864	0.3896	
Foot	0.0114	0.0138	0.823	0.4128	
KeyP	-0.0204	0.0132	-1.551	0.1243	
Off	0.0756	0.0346	2.185	0.0314	*
UnsTch	0.0251	0.0160	1.568	0.1203	
Fouled	-0.0284	0.0142	-1.997	0.0488	*
*p<0.5	$**p < 0.01$	***p<0.001			
Observations	100				
R^2	0.6994				
Adjusted R^2	0.6765				

4.5.3 Mann-Whitney U tests

The notion stemming from the FGLS model that left-footers do not have a goalscoring advantage is further affirmed by a one-sided Mann-Whitney U test, where lefties and righties had their medians compared in terms of GpG. Lefties did not perform significantly better (p-value over 0.05). Afterwards, the same test was run on different performance metrics, namely dribbles per game, shots per game, assists per game and key passes per game. Left-footed overperformance was not found in any of the aforementioned categories, concluding that lefties do not have a performance advantage in these offensive skills among the best 100 goalscorers in Italy.

4.6 Results regarding football - Ligue 1

4.6.1 Binomial test

To test a possible overrepresentation of lefties among the top 100 goalscorers in Ligue 1 a one-sided binomial test was conducted. There were altogether 23 left-footed footballers in the sample. Comparing this ratio with the general population rate of approximately 10% gave us a p-value pronouncedly lower than the 0.05 mark, signalling a significant overrepresentation.

Table 4.12: Results of the binomial test for Ligue 1

Parameter	**Value**
Number of successes	23
Number of trials	100
Sample estimate of success probability	0.23
P-value	0.0001
Confidence interval (95%)	[0.1626, 1.0000]
Alternative hypothesis	True probability of success is greater than 0.1

4.6.2 Feasible Generalised Least Squares

An FGLS model was chosen to deal with heteroscedasticity that was present in the original OLS model. Looking closely at the model summary, we can see that there are two statistically significant estimates of the independent variables, namely shots per game (*SpG*) and dribbles per game (*Drb*). Shots per game have an estimated positive impact (0.11356) on goalscoring, which has been a consistent occurrence across the models. Dribbles per game have an estimated negative impact (-0.04262) on goalscoring which could be explained by the assumption that overdribbling in dangerous areas may be detrimental to a player's goalscoring record as he could lose the ball instead of taking a shot. The footedness (*Foot*) variable is yet again statistically insignificant, so being left-footed does not make any difference when predicting goalscoring success. The rest of the independent variables (*Off*, *KeyP*, *UnsTch*, *Fouled*) are not statistically significant.

The table below shows the estimates of the FGLS regression on the Ligue 1 dataset with *GpG* as the dependent variable.

Table 4.13: Coefficients of the FGLS model for Ligue 1

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.0497	0.0174	2.8490	0.0054	**
SpG	0.1136	0.0185	6.1250	2.21e-08	***
Drb	-0.0426	0.0181	-2.3500	0.0209	*
Foot	-0.0164	0.0172	-0.9540	0.3427	
KeyP	0.0145	0.0096	1.5210	0.1316	
Off	0.0257	0.0586	0.4390	0.6620	
UnsTch	0.0278	0.0184	1.5110	0.1342	
Fouled	-0.0176	0.0195	-0.9010	0.3699	
*p<0.5	$**p < 0.01$	***p<0.001			
Observations	100				
R^2	0.5457				
Adjusted R^2	0.5112				

4.6.3 Mann-Whitney U tests

A one-sided Mann-Whitney U test comparing lefties and righties in terms of goalscoring is in support of the finding from the FGLS model that lefties do not have a goalscoring advantage (as the p-value was drastically above 0.05). In terms of dribbles per game, shots per game and key passes per game, the exact same conclusion was reached. However, when comparing the median values of the two groups in assists per game, the null hypothesis was rejected as the p-value dropped below the significance level. It means that left-footed footballers significantly overperformed their right-footed counterparts in terms of assisting. As the same cannot be said about their performance in the key passes department, it could be hypothesised that lefties and righties were creating chances at a similar rate, but the chances created by lefties were probably better as significantly more of them were converted into goals.

4.7 Results regarding football - FIFA World Cup 2022

A predictive Bradley-Terry model was constructed to test the effect of the average number of left-footed players per game on game-winning success. The data used were all the games played at the 2022 FIFA World Cup in Qatar. The estimate for the variable denoting the average number of lefties per game ($AvgL$)

is not statistically significant, meaning that having more left-footed players on the pitch had not significantly increased a team's chance of winning a game.
It is important to mention that the nature of a World Cup fixture list is such that there are large differences in the number of games played between the teams. For example, a team that was eliminated in the group stage would play only three times whereas a finalist would have accumulated 7 games in total. Add to it the fact that each team plays against a different set of opponents and it comes as no surprise that the estimated ability scores may not be perfectly accurate. The fact that none of the estimates for the team abilities are statistically significant goes in support of the previous claim. Therefore the estimated impact of the number of lefties per game should serve as an illustrative depiction of the possible left-footed advantage rather than a reliable tool used to measure it precisely.
As there are great differences between the number of games played between teams, the result obtained from the overdispersion test was quite concerning. A value of 1.372616 was derived by dividing the residual deviance by the residual degrees of freedom, which is relatively far from the benchmark value of 1. The independence assumption holds as no two games in the sample have a direct impact on each other. The object residuals are more or less scattered around the horizontal 0 line (except for a few outliers), so the functional assumption should be met as well. Overall, the model has a lot of drawbacks due to the inconsistency in the number of games played and could perform better by increasing the number of opponents the early exiting teams would play. Nonetheless, it still has its use as an illustrative measure.

The table below shows the estimates of the ability scores for the Bradley-Terry used on matches played at the 2022 FIFA World Cup.

Table 4.14: Bradley-Terry FIFA World Cup 2022 results

	Estimate	Pr(>|z|)
..England	5.693e+01	0.9970
..Senegal	3.778e+01	0.9970
..USA	3.992e+01	0.9990
..Argentina	7.918e+01	0.9970
..Denmark	5.714e+01	0.9980
..Mexico	7.837e+01	0.9970
..France	7.793e+01	0.9970
..Morocco	4.106e+01	0.9990
..Germany	2.141e+01	0.9990
..Spain	2.141e+01	0.9990
..Belgium	2.204e+01	0.9990
..Switzerland	2.039e+01	0.9990
..Uruguay	2.071e+01	0.9990
..Portugal	2.142e+01	0.9990
..Brazil	2.112e+01	0.9990
..Wales	1.407e+00	1.0000
..Netherlands	5.899e+01	0.9980
..Tunisia	7.772e+01	0.9970
..Poland	7.744e+01	0.9970
..Japan	2.194e+01	0.9990
..Croatia	5.956e+01	0.9980
..Cameroon	2.076e+01	0.9990
..South Korea	2.010e+01	0.9990
..Ecuador	1.889e+01	0.9980
..IR Iran	2.085e+01	0.9990
..Australia	7.750e+01	0.9970
..Saudi Arabia	7.755e+01	0.9970
..Canada	3.011e+00	1.0000
..Costa Rica	2.088e+01	0.9990
..Ghana	2.000e+01	0.9990
..Serbia	1.764e-01	1.0000
..**AvgL**	0.04861	0.7610

Limitations

In addition, I would like to point out some limitations regarding this model. Mainly, the Bradley-Terry model does not deal with tied results. Since the majority of the group stage games have not ended as ties and the play-off games cannot end as ties (because one team has to progress), I decided to ignore the 10 tied games that occurred at the group stage. Since there were 64 games in total, this approach should not drastically affect the final estimates.

4.7.1 Discussion

In each of the top 5 leagues, a significant left-footed overrepresentation was detected among the top 100 goalscorers. On the other hand, lefties did not overperform righties in any of the analysed performance metrics across all the top 5 leagues, with the only exception being assisting in the French Ligue 1, suggesting that they are not performing better than righties when top goalscorers are concerned.

Altogether I would conclude that based on my analysis left-footed players do not have a relevant advantage among top offensive players. However, similarly to tennis, it seems that left-footedness makes it more likely for a player to become elite in the offensive department of the game. Goalscoring was taken as a proxy variable of the overall offensive ability as I was interested in comparing the performance of the two groups in the offensive part of the game. Therefore my analysis is only limited to the offensive performance of the players and could be followed up in the future by research that would take into consideration the defensive part of the game as well.

As far as the effect of left-footedness on the team performance is concerned, the Bradley-Terry model on World Cup results clearly shows that there is no competitive advantage stemming from having more left-footed players on the pitch. However, this model is very limited for the fact that the teams differ a lot in terms of the number of games played as well as the diversity of opponents. More complete data (such as a full league season) where each team plays a high number of games and all teams play each other twice could be analysed in the future to give us a firmer conclusion on the left-footed impact on team achievement.

To sum up, my thesis broadens the current knowledge on the topic as no relevant source in the past has compared the left-footed and right-footed performance of top-level players in such specific performance metrics. In addition, there is no previous study that deals with the impact of having more left-footers on the team.

4.8 Results regarding handball

4.8.1 Binomial test

To test the stance of Loffing et al (2015) who claimed that left-handers are regularly overrepresented among top scorers in international tournaments in handball, I decided to run a one-sided binomial test to see if left-handers are overrepresented in the sample of the top 111 goalscorers at the 2022 EHF European Championship. 38 left-handed players were found in the aforementioned sample. Putting this ratio in comparison to the general population rate of left-handers (approximately 10%) gave us a p-value very close to 0, meaning that the null hypothesis was rejected and therefore left-handers are significantly overrepresented in the sample. This result also indirectly supports the finding by Laxdal et al (2022) which declared a left-handed overrepresentation among 7-meter shooters in handball.

Table 4.15: Results of the binomial test for handball

Parameter	**Value**
Number of successes	38
Number of trials	111
Sample estimate of success probability	0.3423
P-value	4.338e-12
Confidence interval (95%)	[0.2677, 1.0000]
Alternative hypothesis	True probability of success is greater than 0.1

4.8.2 Feasible Generalised Least Squares

As heteroscedasticity was found in the original OLS model, the Feasible Generalised Least Squares model was used to deal with the issue. When exploring the model in detail we can see that the variables shots per game (SpG) and assists per game (ApG) are both statistically significant. The estimate for shots per game is quite large and positive (0.69069) indicating that the more a player shoots, the more goals he is likely to score. The estimate for assists per game is negative (-0.15336) and could be understood as that players in handball have certain roles. Playmakers are expected to assist more than wingers or centres. On the other hand, they do not get into as many goalscoring opportunities, which would explain the negative relationship. However, for the thesis, the most relevant is the estimate of the handedness variable (*Hand*) which is statistically insignificant. Therefore being left-handed does not have a major impact on predicting the goalscoring rate of a player in the model.

The table below shows the estimates of the FGLS regression on the handball dataset with *GpG* as the dependent variable.

Table 4.16: Coefficients of the FGLS model for handball

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.1624	0.1280	1.2690	0.2072	
Hand	-0.0743	0.0962	-0.7720	0.4418	
SpG	0.6907	0.0310	22.2890	< 2e-16	***
ApG	-0.1534	0.0419	-3.6590	0.0004	***
*p<0.5	$**p < 0.01$	***p<0.001			
Observations	111				
R^2	0.838				
Adjusted R^2	0.8335				

4.8.3 Mann-Whitney U tests and T-test

The notion from the previous model that lefties do not have an advantage in terms of goalscoring was subsequently confirmed by a one-sided Mann-Whitney U test. In the test, I compared the median values of lefties and righties in terms of goals per game and since the p-value was high above the 0.05 mark, the null hypothesis was not rejected and therefore it can be concluded that lefties were not overperforming in goalscoring in the sample. The same procedure was conducted for assists per game and shots per game with the same result (no significant left-handed overperformance). The *Efficiency* variable (which measures the success rate of shots) was analysed via t-test since the variable is normally distributed and all the other t-test assumptions also hold. The comparison of left-handed and right-handed mean values led to the conclusion that lefties did not perform significantly better than righties when shooting efficiency was concerned. Overall, left-handers in the sample did not show any sign of performance advantage concerning the offensive attributes tested.

4.8.4 Discussion

A one-sided binomial test confirmed the result by Loffing et al (2015) that lefties are overrepresented among top scorers at international handball tournaments. However, neither the model nor the performance comparisons of lefties and righties via the Mann-Whitney U test (or the T-test) have shown any sign of left-handed overperformance in offensive handball attributes. My results would suggest that left-handedness is helpful when trying to become a world-class offensive handball player. However, once at the very best level, handedness does not play a significant role in terms of success among the elite. This conclusion is very similar to the one by Laxdal et al (2022) who propose the same hypothesis regarding the 7-meter shots in handball. My work has extended the contemporary knowledge by also reviewing the left-handed performance in terms of assisting, the efficiency of shots and the number of shots, whereas the past studies focused mainly on goalscoring. The contribution can be enhanced by future studies that would focus more on the defensive side of handball, which is omitted in my analysis.

4.9 Results regarding basketball

4.9.1 Binomial test

The stance that left-handed prevalence among basketball players is similar to the general population rate (Loffing et al (2016)) was tested by a one-sided binomial test. The sample of all the NBA players from the 2021/2022 season contains 605 players in total, among which 55 are left-handed. This left-handedness ratio was compared to the presumed rate in the general population (around 10%). The p-value of the test was above the 0.05 threshold, meaning that there is no significant overrepresentation of left-handers in the NBA, which affirms the original notion.

Table 4.17: Results of the binomial test for basketball

Parameter	**Value**
Number of successes	55
Number of trials	605
Sample estimate of success probability	0.0909
P-value	0.7902
Confidence interval (95%)	[0.0724, 1.0000]
Alternative hypothesis	True probability of success is greater than 0.1

4.9.2 Feasible Generalised Least Squares model

To treat the heteroscedasticity issue in the original model, a Feasible Generalised Least Squares model was developed. Looking at the model we can see that four of the independent variables are statistically significant when predicting the dependent variable points per game (PPG). Firstly, the assists per game variable ($ASTPG$) has a positive estimate of 1.53231. This can be understood as that offensively productive players in basketball are oftentimes good at providing for others as well as scoring points themselves. The steals per game variable ($STLPG$) also has a positive estimate (1.91991), which is not surprising since winning possession, especially in the opponent's half, increases the team's chance to score, which directly increases the given player's chance to score a point. Furthermore, both defensive ($DREBPG$) and offensive ($OREBPG$) rebounds per game variables are statistically significant. The defensive rebound variable has a positive estimate of 1.72671, suggesting that a lot of points are scored after a counterattack when the defending player gains possession via rebound. The offensive rebound variable, however, has a negative estimate of -0.51244, which is intriguing since common sense would suggest that gaining possession near the opponent's basket after a missed shot would present a good opportunity for scoring a point. The variable of interest $Hand$ is not statistically significant, therefore being left-handed does not affect a player's chances of becoming a prolific points scorer. Finally, the blocks per game variable ($BLKPG$) is not statistically significant in the model.

The table below shows the estimates of the FGLS regression on the basketball dataset with PPG as the dependent variable.

Table 4.18: Coefficients of the FGLS model for basketball

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-0.0083	0.1486	-0.0560	0.9557	
Hand	-0.0531	0.3015	-0.1760	0.8602	
ASTPG	1.5323	0.1298	11.8040	< 2e-16	***
DREBPG	1.7267	0.1249	13.8220	< 2e-16	***
OREBPG	-0.5124	0.2116	-2.4210	0.0158	*
BLKPG	0.5703	0.4077	1.3990	0.1624	
STLPG	1.9199	0.3987	4.8150	1.87e-06	***
*p<0.5	$**p < 0.01$	***p<0.001			
Observations	605				
R^2	0.7369				
Adjusted R^2	0.7342				

4.9.3 Mann-Whitney U tests

The hypothesis by Lawler and Lawler (2011), who claim that left-handed basketball players performed better at rebounds, assists, points per game and field-goal percentage, was tested by constructing a one-sided Mann-Whitney U test and comparing the median values of left-handers and right-handers in the aforementioned attributes. A few additional attributes were also tested to broaden the current knowledge of left-handed performance in basketball. For all four variables mentioned in the Lawler and Lawler study the tests showed that there is no significant overperformance of lefties in any of these skills (*PPG*), *REBG*, *FG pct*, *ASTPG*). Lefties were not more successful at rebounds even after dividing them into defensive and offensive rebounds. Even when direct pressure of opposing players is discarded, lefties did not perform better, which is illustrated by them not having significantly more success from free throws (*FTMPG*). Left-handers also have not overperformed in *FGAPG* meaning that they are not shooting more often than right-handers. In terms of defensive attributes, left-handers are not, on average, more successful at steals per game (*STLPG*). However, when blocks per game (*BLKPG*) are concerned, left-handers did perform significantly better than right-handers. Overall, lefties proved to overperform righties only in blocking, while no other metric showed any significant performance difference between the two groups.

4.9.4 Discussion

No left-handed overrepresentation was found after running the one-sided binomial test, affirming the stance by Loffing et al (2016) that the left-handed prevalence among basketball players is similar to the one in the general population. The FGLS model did not reveal any important impact of handedness on the rate of scoring points. Lefties also did not overperform in any of the offensive or defensive metrics according to the Mann-Whitney U tests conducted, which directly argues against claims by Lawler and Lawler (2011). The only exception to this trend was blocking, where lefties performed significantly better than righties. However, since blocking is an activity that usually requires both hands to an equal extent, it is difficult to uphold the standpoint that left-handers would be in advantage. I would rather suppose that in the particular analysed season (2021/2022) there were simply more left-handers that were focused on blocking. This hypothesis would need to be tested across more seasons.

Altogether, the evidence suggests that there is no advantage to being left-handed in the NBA. My thesis contributes to the current knowledge by analysing the left-handed ability in many specific in-game attributes (both offensive and defensive), which has not been done by a relevant source in the past. It could be further perfected by conducting a similar analysis on the Euroleague and then comparing the results.

4.10 Results regarding table tennis

4.10.1 Binomial test

A one-sided binomial test was applied to check if there is a significant overrepresentation of left-handers among the top 100 ranked table tennis players in 2022. There are altogether 26 lefties in the sample. Comparing this ratio to the popualtion rate of 10% showed that overrepresentation was present in the sample (p-value lower than 0.05 mark). Thus, my analysis affirms the results by Loffing (2017) who claims that left-handers are overrepresented in table tennis since it is a ”high time pressure sport”.

Table 4.19: Results of the binomial test for table tennis

Parameter	Value
Number of successes	26
Number of trials	100
Sample estimate of success probability	0.26
P-value	4.1e-06
Confidence interval (95%)	[0.1890, 1.0000]
Alternative hypothesis	True probability of success is greater than 0.1

4.10.2 Bradley-Terry model

As I wanted to check if left-handed table tennis players have a competitive advantage in direct encounters among the very best players, I constructed a Bradley-Terry model and compared the ability score estimates given by the model with the official year-end rankings to see if lefties would improve their rank when only head-to-head contests are concerned. By looking at the left-handed players in the two rankings we can see that Wang improved by two places when only ability scores based on direct encounters are concerned (from 4th to 2nd). Lin Yun-ju did not move at all from one ranking to another. Boll's rank got worse when only head-to-head encounters are taken into the record, with his ability score being the 19th best (compared to 12th in the official ranking). Lin Gaoyuan improved his rank by 6 places in the head-to-head rankings, going from 13th up to 7th. Lim's rise is even more impressive, with him climbing a staggering 11 places from 16th up to 5th. Lastly, Karlsson's improvement in the ability scores table is also noteworthy, jumping from 19th all the way up to 9th. Overall, we can see the trend is such that lefties tend to perform more impressively in direct contests with the very best.

To see how significant left-handedness is as a factor in these encounters, a predictive extension of the Bradley-Terry model was used on the same head-to-head table with an additional dataset being a table containing the handedness information of each player. After running the predictive model it can be seen that the handedness variable (*Hand*) is statistically insignificant, meaning that the left-handed effect on success in direct matches between top-level players is not as important as it would have seemed from the original model.

It is important to note that overdispersion may be a problem in this model (the test statistic being 1.57745) as many players in the sample did not play each other at all in the season. On the other hand, some pairs of players have played each other on many occasions, so the differences in both the total matches played as well as the diversity of opponents are large. The independence assumption holds as every match is independent of all the other matches. The functional assumption also holds as all the points (except for two outliers) are scattered around the horizontal 0 line when plotting the object residuals against the ability scores.

The table below shows the estimates of the ability scores for the Bradley-Terry used on head-to-head contests between the top 20 table tennis players in 2022.

Table 4.20: Bradley-Terry table tennis results

	Estimate	Std. Error	Pr(>|t|)
..Harimoto	0.3503	0.5290	0.5078
..**Wang**	0.2258	0.4354	0.6041
..Zhedong	0.0000	0.0000	0.0000***
..Ma	-0.1006	0.4335	0.8164
..**Lim**	-0.2669	0.5219	0.6090
..Liang	-0.3803	0.5518	0.4907
..**Lin G**	-0.5065	0.5114	0.3220
..**Lin Y**	-0.5743	0.4618	0.2137
..**Karlsson**	-0.7648	0.7017	0.2757
..Filus	-0.7753	0.7357	0.2919
..Moregard	-0.8690	0.5478	0.1126
..Qiu	-0.9170	0.5525	0.0970
..Franziska	-0.9730	0.5092	0.0561
..Calderano	-0.9814	0.5564	0.0778
..Ovtcharov	-1.0232	0.5510	0.0633
..Aruna	-1.0372	0.6137	0.0910
..Chuang	-1.0578	0.5867	0.0714
..Jang	-1.5097	0.6202	0.0149*
..**Boll**	-1.5397	0.6648	0.0206*
..Jorgic	-1.6677	0.5654	0.0032**
..**Hand**	0.2400	0.2534	0.3440

*p<0.5 $**p < 0.01$ ***p<0.001

4.10.3 Comparison of the average ranking

Since the *Rank* variable is ordinal and not continuous, I could not use the T-test nor the Mann-Whitney U test for comparison. I decided to simply compare the average rank of lefties and righties. After doing so I found that there is no significant difference between their average ranks, with the left-handed average ranking being 52.73077, while the right-handed average ranking is slightly better at 49.71622. Overall, it can be concluded that no left-handed advantage was detected after the comparison.

Table 4.21: Mean rankings for left-handers and right-handers in table tennis

Left-handers mean ranking	52.7308
Right-handers mean ranking	49.7162

4.10.4 Discussion

The hypothesis by Loffing (2017) was proven right by the one-sided binomial test conducted on the top 100 table tennis players of 2022, as a significant overrepresentation of left-handers was found. When comparing the average rankings of lefties and righties, I have not found any perceptible difference between the two groups, suggesting that they perform on a similar level. Even though the original Bradley-Terry model discovered a trend of left-handers being ranked better (than in the official rankings) when only direct matches of the top 20 players are concerned, the predictive Bradley-Terry model proved the impact of left-handedness on success in the aforementioned direct contests to be insignificant. It is important to note that the data for the Bradley-Terry model contain only matches from the year 2022 and therefore are differing largely between players in terms of the total number of games and the variety of opposition.

Overall, the analysis advocates against the potential performance benefit stemming from being left-handed among top table tennis players. However, as there is a significant overrepresentation of lefties, it can be assumed that being left-handed helps players achieve elite-level ranking in table tennis.

The analysis regarding table tennis is not supported by any performance-related data as they are not easily accessible. Incorporating them in future research would improve the understanding of left-handed performance in table tennis provided by my thesis. My thesis enhances the current understanding of lefties in table tennis by analysing the impact of left-handedness in direct matches among the very best players, an approach not used in the past.

Chapter 5

Conclusion

In this thesis, we learned the specifics of the left-handers'/left-footers' performance in sports and the extent to which they are successful when compared to their right-handed/right-footed counterparts. Concretely, we found out that in 4 out of 5 sports that were subject to analysis of this work lefties were significantly overrepresented among a certain number of top players. However, when directly comparing the performance level of lefties and righties, we failed to detect any important difference. These findings put together could imply that left-handers have a competitive advantage when progressing to the elite level in the given sport. Nevertheless, once they reach that level, they no longer have an edge over righties. This may be explained by the fact that top-level athletes usually prepare specifically for each opponent and therefore any potential advantage stemming from the scarcity of lefties may diminish after an in-depth analysis of their game. When studying the impact of the left-handed (or left-footed) advantage in direct matches across 2 individual sports (tennis, table tennis) as well as in 1 team sport (football), no significant impact of left-handedness/left-footedness was discovered. This argues against the common notion that left-handers are at an advantage when facing opponents in direct encounters due to their scarcity and the unpredictability which results from it.

My thesis contributes to the contemporary knowledge of the topic to a large extent, as previously no renowned author inspected and compared the achievement of lefties and righties in a set of varying performance metrics that cover a great deal of in-game action across multiple sports. In addition, there are no previous studies that would examine the effect of left-handedness/left-footedness in direct encounters between individuals and teams. These findings may be of great importance mainly to coaches and scouts of professional sports teams. Coaches of smaller clubs could, for example, be aware of the fact that lefties seem likelier to reach the summit of the given sport and therefore could engage more lefties in their lineups in order to gain promotion to a higher division and develop future stars. Scouts may be well-served with the information that in top-level sports, handedness (or footedness) does not play an important role when predicting a performance success in a given sport, which contradicts the popular belief that lefties are generally more talented athletes. Finally, my thesis will be found interesting by many sports enthusiasts who enjoy investigating in great detail the sporting abilities of famous athletes and inspecting the underlying foundations on which their success rests.

Future studies could build on my work by analysing defensive performance metrics in team sports such as football or handball since I worked with offensive variables only. For basketball, a comparison with the Euroleague would show if the absence of the left-handed advantage is due to the nature of the NBA or if it is consistent across elite basketball competitions. The impact of having left-footers in a football team could be estimated more precisely if analysed on large, more consistent data where each team plays against each other the same amount of times (e.g. league season). Lastly, the left-handed effect on table tennis performance would be understood more in-depth if specific performance metrics were used for the analysis.

To sum up, this thesis provides many interesting findings regarding the performance of lefties across multiple sports. It argues for lefties having greater odds of becoming top-level athletes. However, after reaching that level, no important handedness nor footedness effect is present.

Chapter 6

Bibliography

- Akpinar, S. et al (2014). Left-handers/footers representation in Sports. Montenegrin Journal of Sports Science and Medicine, pp. 33-38
- Bartell, E. (2019). Code Through - Checking Assumptions for OLS. RPubs. Retrieved April 15, 2023, from https://www.rpubs.com/elliottb90/olsassumptions
- Bozkurt, S., Kucuk, V. (2018). Comparing of Technical Skills of Young Football Players According to Preferred Foot. International Journal of Human Movement and Sports Sciences, 6(1): 19-22
- Brooks, R., Bussiere, L. F., Jennions, M. D., Hunt, J. (2004). Sinister strategies succeed at the cricket World Cup. Proceedings of the Royal Society London.
- Del Corral, J., Prieto-Rodriguez, J. (2010). Are Differences in Ranks Good Predictors for Grand Slam Tennis Matches. International Journal of Forecasting, 26, 551-563
- Fagan, F., Haugh, M., Cooper, H. (2019). The advantage of lefties in one-on-one sports. Journal of Quantitative Analysis in Sports, vol. 15, no. 1, pp. 1-25
- Faurie, C., Llaurens, V., Alvergne, A., Goldberg, M., Zins, M., Raymond, M.(2011). Left-handedness and male-male competition: insights from fighting and hormonal data. Evolutionary Psychology
- Fernandez, J. (2020). The statistical analysis t-test explained for beginners and experts. Medium. Retrieved April 15,

2023, from https://towardsdatascience.com/the-statistical-analysis-t-test-explained-for-beginners-and-experts-fd0e358bbb62

- Grouios, G., Tsorbatzoudis, H., Alexandris, K., Barkoukis, V. (2000). Do left-handed competitors have an innate superiority in sports? Perceptual and Motor Skills, 90(3,Pt2), 1273-1282

- Hagemann, N. (2009). The advantage of being left-handed in interactive sports. Attention, Perception, and Psychophysics, 71(7), 1641-1648.

- Hanek, P. (2010). Bradley-Terry model. Bakalářská práce. Univerzita Karlova, Matematicko-fyzikální fakulta, Katedra pravděpodobnosti a matematické statistiky. Vedoucí práce Hudecová, Šárka.

- Lawler, T.P., Lawler, F.H. (2011). Left-handedness in professional basketball: prevalence, performance, and survival. Perceptual and Motor Skills

- Laxdal, A., Ivarsson, A., Thorgeirsson, S., Haugen, T. (2022). Born to Score? The Relationship between Left-Handedness and Success from the 7-Meter Line. Symmetry , 14, 2163

- Live Science Staff (2012). Study reveals why lefties are rare. LiveScience. Retrieved May 2, 2023, from https://www.livescience.com/19968-study-reveals-lefties-rare.html

- Loffing, F. (2017). Left-handedness and time pressure in elite interactive ball games. Biology Letters

- Loffing, F., Hagemann, N. (2015). Pushing through evolution? Incidence and fight records of left-oriented fighters in professional boxing history. Laterality: Asymmetries of Body, Brain and Cognition, 20(3), 270-286

- Loffing, F., Hagemann, N., Strauss, B. (2010). Automated processes in tennis: Do lefthanded players benefit from the tactical preferences of their opponents? Journal of Sports Sciences, 28(4), 435-443

- Loffing, F., Hagemann, N., Strauss, B. (2012). Left-handedness in professional and amateur tennis. PLoS One; 7

- Loffing, F., Hagemann, N., Strauss, B., MacMahon, C. (Eds.). (2016). Laterality in sports: Theories and applications. Academic Press, pp. 249-277; 309-328

- Loffing, F., Schorer, J. (2021). Handedness and Relative Age in International Elite Interactive Individual Sports Revisited. Frontiers in Sports and Active Living
- Loffing, F., Schorer, J., Hagemann, N., Baker, J. (2012). On the advantage of being left-handed in volleyball: further evidence of the specificity of skilled visual perception. Attention, Perception and Psychophysics
- Loffing, F., Solter, F., Hagemann, N., Strauss, B. (2015). Accuracy of Outcome Anticipation, But Not Gaze Behavior, Differs Against Left- and Right-Handed Penalties in Team-Handball Goalkeeping. Frontiers in Psychology
- McClenaghan, E. (n.d.). The binomial test. Informatics from Technology Networks. Retrieved April 15, 2023, from https://www.technologynetworks.com/informatics/articles/the-binomial-test-366022
- McLean, J. M., Ciurczak, F. M. (1982). Bimanual Dexterity in Major League Baseball Players: A Statistical Study. The New England Journal of Medicine, 307, 1278 Scharoun, S.M., Bryden, P.J.(2014). Hand preference, performance abilities, and hand selection in children. Frontiers in Psychology
- Schorer, J., Loffing, F., Hagemann, N., Baker, J. (2012). Human handedness in interactive situations: Negative perceptual frequency effects can be reversed!, Journal of Sports Sciences, 30:5, 507-513
- Statistics Lectures (n.d.). Mann-Whitney U. Retrieved April 15, 2023, from http://www.statisticslectures.com/topics/mannwhitneyu/
- Stockel, T., Vater, C. (2014). Hand preference patterns in expert basketball players: interrelations between basketball-specific and everyday life behavior. Human Movement Science, 38, 143-151
- Stockel, T., Weigelt, M. (2012). Plasticity of human handedness: decreased one-hand bias and inter-manual performance asymmetry in expert basketball players. Journal of Sports Sciences
- The University of Texas at Austin (n.d.). Binomial Test. Retrieved April 15, 2023, from http://sites.utexas.edu/sos/guided/inferential/categorical/univariate/binomial/

- Turner, H., Firth, D. (2012). Bradley-Terry Models in R: The BradleyTerry2 Package. Journal of Statistical Software, 48(9), 1-21

- Wooldridge, J. (2016): Introductory econometrics: a modern approach. Boston: Cengage Learning.

- Wu, W., Niezink, N., Junker, B. (2022) A Diagnostic Framework for the Bradley-Terry Model, Journal of the Royal Statistical Society Series A: Statistics in Society, Volume 185, Issue Supplement 2, pp S461-S484

Appendix A

Additional tables

A.1 Tables for football - correlation matrices ans summary statistics

	Foot	SpG	KeyP	Drb	UnsTch
Min.	0.0000	0.3000	0.1000	0.0000	0.1000
1st Qu.	0.0000	1.2000	0.7000	0.6000	1.2000
Median	0.0000	1.6000	1.0000	0.9000	1.7000
Mean	0.2500	1.6580	1.0360	0.9340	1.7380
3rd Qu.	0.0000	2.1000	1.4000	1.2000	2.2000
Max.	1.0000	4.0000	2.9000	4.3000	3.8000

Table A.1: Summary Statistics for Premier League (a)

	Fouled	Off	Disp	GpG	ApG
Min.	0.0000	0.0000	0.0000	0.0857	0.0000
1st Qu.	0.6000	0.1000	0.7000	0.1429	0.0476
Median	0.9000	0.2000	1.0000	0.2132	0.0976
Mean	0.9630	0.2530	1.0490	0.2381	0.1142
3rd Qu.	1.2250	0.4000	1.3000	0.2912	0.1667
Max.	3.1000	1.1000	3.3000	0.6571	0.3714

Table A.2: Summary Statistics for Premier League (b)

Table A.3: Correlation matrix for Premier League

	Foot	GpG	ApG	SpG	KeyP	Drb	Fouled	Off	Disp	UnsTch
Foot	1.0000	-0.0443	0.0346	-0.0322	0.1281	-0.0620	-0.1010	-0.0325	-0.0757	-0.0452
GpG	-0.0443	1.0000	0.3398	0.7544	0.3936	0.2286	0.1470	0.5265	0.3368	0.3464
ApG	0.0346	0.3398	1.0000	0.4797	0.6107	0.2879	0.0750	0.1180	0.2416	0.2384
SpG	-0.0322	0.7544	0.4797	1.0000	0.4896	0.3550	0.2545	0.5579	0.4827	0.5368
KeyP	0.1281	0.3936	0.6107	0.4896	1.0000	0.4240	0.3179	0.0165	0.3453	0.2632
Drb	-0.0620	0.2286	0.2879	0.3550	0.4240	1.0000	0.4834	0.2439	0.7481	0.6687
Fouled	-0.1010	0.1470	0.0750	0.2545	0.3179	0.4834	1.0000	0.2396	0.5349	0.5436
Off	-0.0325	0.5265	0.1180	0.5579	0.0165	0.2439	0.2396	1.0000	0.4342	0.5511
Disp	-0.0757	0.3368	0.2416	0.4827	0.3453	0.7481	0.5349	0.4342	1.0000	0.8429
UnsTch	-0.0452	0.3464	0.2384	0.5368	0.2632	0.6687	0.5436	0.5511	0.8429	1.0000

	Foot	SpG	KeyP	Drb	UnsTch
Min.	0.0000	0.2000	0.0000	0.0000	0.3000
1st Qu.	0.0000	0.9000	0.5000	0.4000	1.1750
Median	0.0000	1.3000	0.7000	0.8000	1.5000
Mean	0.2900	1.3780	0.8460	0.8620	1.6280
3rd Qu.	1.0000	1.7000	1.1000	1.2000	2.1000
Max.	1.0000	4.0000	3.1000	2.7000	5.9000

Table A.4: Summary Statistics for La Liga (a)

	Fouled	Off	Disp	GpG	ApG
Min.	0.2000	0.0000	0.0000	0.0645	0.0000
1st Qu.	0.7000	0.1000	0.6000	0.1140	0.0333
Median	1.0000	0.3000	0.9000	0.1719	0.0896
Mean	1.0970	0.2690	0.9260	0.2062	0.0973
3rd Qu.	1.4000	0.4000	1.2000	0.2500	0.1382
Max.	2.7000	1.0000	3.3000	0.8438	0.3750

Table A.5: Summary Statistics for La Liga (b)

Table A.6: Correlation matrix for La Liga

	Foot	GpG	ApG	SpG	KeyP	Drb	Fouled	Off	Disp	UnsTch
Foot	1.0000	-0.1030	-0.0613	-0.0504	0.0576	0.0141	-0.0494	-0.0402	0.0553	-0.0546
GpG	-0.1030	1.0000	0.3126	0.7448	0.1699	0.1506	0.0291	0.6081	0.1813	0.2646
ApG	-0.0613	0.3126	1.0000	0.3574	0.6985	0.5613	0.2819	0.2052	0.3408	0.2632
SpG	-0.0504	0.7448	0.3574	1.0000	0.2704	0.3247	0.2095	0.6881	0.3515	0.4722
KeyP	0.0576	0.1699	0.6985	0.2704	1.0000	0.6503	0.5221	0.0254	0.4861	0.2930
Drb	0.0141	0.1506	0.5613	0.3247	0.6503	1.0000	0.4951	0.1830	0.6952	0.5958
Fouled	-0.0494	0.0291	0.2819	0.2095	0.5221	0.4951	1.0000	-0.0513	0.5649	0.4832
Off	-0.0402	0.6081	0.2052	0.6881	0.0254	0.1830	-0.0513	1.0000	0.2018	0.3792
Disp	0.0553	0.1813	0.3408	0.3515	0.4861	0.6952	0.5649	0.2018	1.0000	0.7887
UnsTch	-0.0546	0.2646	0.2632	0.4722	0.2930	0.5958	0.4832	0.3792	0.7887	1.0000

	Foot	SpG	KeyP	Drb	UnsTch
Min.	0.0000	0.4000	0.0000	0.1000	0.5000
1st Qu.	0.0000	1.0000	0.5000	0.5000	1.1000
Median	0.0000	1.4000	0.8000	0.7000	1.6500
Mean	0.2300	1.4720	0.9480	0.8250	1.7390
3rd Qu.	0.0000	1.7250	1.2000	1.1000	2.2000
Max.	1.0000	4.7000	2.9000	2.3000	4.5000

Table A.7: Summary Statistics for Bundesliga (a)

	Fouled	Off	Disp	GpG	ApG
Min.	0.1000	0.0000	0.0000	0.0938	0.0000
1st Qu.	0.6000	0.1000	0.6000	0.1429	0.0430
Median	0.8000	0.2000	0.8000	0.1739	0.0876
Mean	0.9330	0.2440	0.8510	0.2290	0.1182
3rd Qu.	1.1000	0.4000	1.1000	0.2597	0.1521
Max.	2.8000	1.1000	2.1000	1.0294	0.5625

Table A.8: Summary Statistics for Bundesliga (b)

Table A.9: Correlation matrix for Bundesliga

	Foot	GpG	ApG	SpG	KeyP	Drb	Fouled	Off	Disp	UnsTch
Foot	1.0000	0.0619	0.0451	0.0353	0.0454	-0.0711	-0.0274	-0.0323	-0.1294	-0.0536
.GpG	0.0619	1.0000	0.1540	0.7641	0.1064	0.1371	0.0375	0.4641	0.2653	0.4030
ApG	0.0451	0.1540	1.0000	0.2761	0.8108	0.4062	0.1430	0.1483	0.2462	0.2330
SpG	0.0353	0.7641	0.2761	1.0000	0.3048	0.3060	0.2110	0.5157	0.4997	0.5508
KeyP	0.0454	0.1064	0.8108	0.3048	1.0000	0.4223	0.1970	0.1316	0.2865	0.2573
Drb	-0.0711	0.1371	0.4062	0.3060	0.4223	1.0000	0.2878	0.1423	0.6009	0.4781
Fouled	-0.0274	0.0375	0.1430	0.2110	0.1970	0.2878	1.0000	0.1595	0.5394	0.3761
Off	-0.0323	0.4641	0.1483	0.5157	0.1316	0.1423	0.1595	1.0000	0.3985	0.5728
Disp	-0.1294	0.2653	0.2462	0.4997	0.2865	0.6009	0.5394	0.3985	1.0000	0.7786
UnsTch	-0.0536	0.4030	0.2330	0.5508	0.2573	0.4781	0.3761	0.5728	0.7786	1.0000

	Foot	SpG	KeyP	Drb	UnsTch
Min.	0.0000	0.4000	0.1000	0.1000	0.2000
1st Qu.	0.0000	1.1000	0.6000	0.4000	1.1750
Median	0.0000	1.5000	0.9000	0.7000	1.6000
Mean	0.2300	1.6580	1.0140	0.8170	1.6780
3rd Qu.	0.0000	2.1000	1.3000	1.1000	2.2250
Max.	1.0000	3.8000	2.6000	2.9000	3.9000

Table A.10: Summary Statistics for Serie A (a)

	Fouled	Off	Disp	GpG	ApG
Min.	0.1000	0.0000	0.0000	0.0811	0.0000
1st Qu.	0.7000	0.1000	0.5000	0.1429	0.0417
Median	1.0000	0.2000	0.8000	0.2029	0.1035
Mean	1.1540	0.2640	0.8690	0.2453	0.1116
3rd Qu.	1.4000	0.4000	1.2000	0.3333	0.1615
Max.	2.9000	1.0000	2.3000	0.8710	0.4242

Table A.11: Summary Statistics for Serie A (b)

Table A.12: Correlation matrix for Serie A

	Foot	GpG	ApG	SpG	KeyP	Drb	Fouled	Off	Disp	UnsTch
Foot	1.0000	0.0372	0.0141	0.0325	0.0207	-0.0645	0.1296	-0.1031	-0.0387	-0.0876
GpG	0.0372	1.0000	0.0935	0.8161	0.1545	0.1424	0.2148	0.5317	0.5058	0.6129
ApG	0.0141	0.0935	1.0000	0.3256	0.7283	0.3563	0.1647	-0.0045	0.1667	0.1601
SpG	0.0325	0.8161	0.3256	1.0000	0.3686	0.3428	0.3492	0.4776	0.5719	0.6886
KeyP	0.0207	0.1545	0.7283	0.3686	1.0000	0.4444	0.0849	0.0251	0.1947	0.1838
Drb	-0.0645	0.1424	0.3563	0.3428	0.4444	1.0000	0.3325	0.0254	0.5352	0.4648
Fouled	0.1296	0.2148	0.1647	0.3492	0.0849	0.3325	1.0000	0.2509	0.3734	0.4914
Off	-0.1031	0.5317	-0.0045	0.4776	0.0251	0.0254	0.2509	1.0000	0.4632	0.6161
Disp	-0.0387	0.5058	0.1667	0.5719	0.1947	0.5352	0.3734	0.4632	1.0000	0.8159
UnsTch	-0.0876	0.6129	0.1601	0.6886	0.1838	0.4648	0.4914	0.6161	0.8159	1.0000

	Foot	SpG	KeyP	Drb	UnsTch
Min.	0.0000	0.4000	0.1000	0.0000	0.2000
1st Qu.	0.0000	0.9750	0.5000	0.3000	1.1000
Median	0.0000	1.5000	0.9000	0.8000	1.8000
Mean	0.2300	1.5390	0.9790	0.8970	1.8010
3rd Qu.	0.0000	2.0000	1.2000	1.1250	2.4000
Max.	1.0000	4.2000	3.2000	3.2000	3.8000

Table A.13: Summary Statistics for Ligue 1 (a)

	Fouled	Off	Disp	GpG	ApG
Min.	0.2000	0.0000	0.0000	0.0857	0.0000
1st Qu.	0.5750	0.1000	0.5000	0.1380	0.0548
Median	0.9000	0.2000	0.8000	0.2076	0.0938
Mean	0.9850	0.2420	0.9420	0.2397	0.1140
3rd Qu.	1.3000	0.4000	1.2250	0.2929	0.1490
Max.	3.2000	1.3000	3.0000	0.8000	0.5385

Table A.14: Summary Statistics for Ligue 1 (b)

Table A.15: Correlation matrix for Ligue 1

	Foot	GpG	ApG	SpG	KeyP	Drb	Fouled	Off	Disp	UnsTch
Foot	1.0000	-0.0910	0.2396	0.0576	0.0493	0.0023	-0.0430	-0.0696	0.0387	0.0586
GpG	-0.0910	1.0000	0.2773	0.6664	0.2993	0.2811	0.1852	0.4805	0.4068	0.4896
ApG	0.2396	0.2773	1.0000	0.4485	0.7779	0.4955	0.1939	0.1625	0.3415	0.2637
SpG	0.0576	0.6664	0.4485	1.0000	0.4397	0.5350	0.3806	0.4900	0.5756	0.6514
KeyP	0.0493	0.2993	0.7779	0.4397	1.0000	0.5614	0.3640	0.0518	0.4840	0.3347
Drb	0.0023	0.2811	0.4955	0.5350	0.5614	1.0000	0.6608	0.1196	0.7627	0.6384
Fouled	-0.0430	0.1852	0.1939	0.3806	0.3640	0.6608	1.0000	0.0682	0.6973	0.5882
Off	-0.0696	0.4805	0.1625	0.4900	0.0518	0.1196	0.0682	1.0000	0.2507	0.4617
Disp	0.0387	0.4068	0.3415	0.5756	0.4840	0.7627	0.6973	0.2507	1.0000	0.7993
UnsTch	0.0586	0.4896	0.2637	0.6514	0.3347	0.6384	0.5882	0.4617	0.7993	1.0000

A.2 Tables for football - Mann-Whitney U tests and T-tests

Table A.16: Results of the Mann-Whitney U test for goals per game in Premier League

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiespremGpGonly and rightiespremGpGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 864.5
P-value	0.7208
Alternative hypothesis	True location shift is greater than 0

Table A.17: Results of the Mann-Whitney U test for assists per game in Premier League

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiespremApGonly and rightiespremApGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 939
P-value	0.4968
Alternative hypothesis	True location shift is greater than 0

Table A.18: Results of the Mann-Whitney U test for dribbles per game in Premier League

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiespremDrbonly and rightiespremDrbonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 888
P-value	0.6551
Alternative hypothesis	True location shift is greater than 0

Table A.19: Results of the Mann-Whitney U test for shots per game in Premier League

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiespremSpGonly and rightiespremSpGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 860.5
P-value	0.7316
Alternative hypothesis	True location shift is greater than 0

Table A.20: Results of the Mann-Whitney U test for key passes per game in Premier League

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiespremKeyPonly and rightiespremKeyPonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 1116.5
P-value	0.07712
Alternative hypothesis	True location shift is greater than 0

Table A.21: Results of the Mann-Whitney U test for goals per game in La Liga

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftieslaligaGpGonly and rightieslaligaGpGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 951
P-value	0.7258
Alternative hypothesis	True location shift is greater than 0

Table A.22: Results of the Mann-Whitney U test for assists per game in La Liga

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftieslaligaApGonly and rightieslaligaApGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 985
P-value	0.6341
Alternative hypothesis	True location shift is greater than 0

Table A.23: Results of the Mann-Whitney U test for dribbles per game in La Liga

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftieslaligaDrbonly and rightieslaligaDrbonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 1092
P-value	0.3184
Alternative hypothesis	True location shift is greater than 0

Table A.24: Results of the Mann-Whitney U test for shots per game in La Liga

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftieslaligaSpGonly and rightieslaligaSpGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 990.5
P-value	0.6181
Alternative hypothesis	True location shift is greater than 0

Table A.25: Results of the Mann-Whitney U test for key passes per game in La Liga

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftieslaligaKeyPonly and rightieslaligaKeyPonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 1095
P-value	0.3102
Alternative hypothesis	True location shift is greater than 0

Table A.26: Results of the Mann-Whitney U test for goals per game in Bundesliga

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesbundesligaGpGonly and rightiesbundesligaGpGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 797
P-value	0.7671
Alternative hypothesis	True location shift is greater than 0

Table A.27: Results of the Mann-Whitney U test for assists per game in Bundesliga

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesbundesligaApGonly and rightiesbundesligaApGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 923
P-value	0.3808
Alternative hypothesis	True location shift is greater than 0

Table A.28: Results of the Mann-Whitney U test for dribbles per game in Bundesliga

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesbundesligaDrbonly and rightiesbundesligaDrbonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 791
P-value	0.7824
Alternative hypothesis	True location shift is greater than 0

Table A.29: Results of the Mann-Whitney U test for shots per game in Bundesliga

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftieslbundesligaSpGonly and rightiesbundesligaSpGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 890
P-value	0.4869
Alternative hypothesis	True location shift is greater than 0

Table A.30: Results of the Mann-Whitney U test for key passes per game in Bundesliga

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesbundesligaKeyPonly and rightiesbundesligaKeyPonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 870
P-value	0.5523
Alternative hypothesis	True location shift is greater than 0

Table A.31: Results of the Mann-Whitney U test for goals per game in Serie A

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesserieaGpGonly and rightiesserieaGpGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 922
P-value	0.384
Alternative hypothesis	True location shift is greater than 0

Table A.32: Results of the Mann-Whitney U test for assists per game in Serie A

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesserieaApGonly and rightiesserieaApGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 898.5
P-value	0.4592
Alternative hypothesis	True location shift is greater than 0

Table A.33: Results of the Mann-Whitney U test for dribbles per game in Serie A

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesserieaDrbonly and rightiesserieaDrbonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 795.5
P-value	0.7714
Alternative hypothesis	True location shift is greater than 0

Table A.34: Results of the Mann-Whitney U test for shots per game in Serie A

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesserieaSpGonly and rightiesserieaSpGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 914
P-value	0.4092
Alternative hypothesis	True location shift is greater than 0

Table A.35: Results of the Mann-Whitney U test for key passes per game in Serie A

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesserieaKeyPonly and rightiesserieaKeyPonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 942
P-value	0.3229
Alternative hypothesis	True location shift is greater than 0

Table A.36: Results of the Mann-Whitney U test for goals per game in Ligue 1

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesligue1GpGonly and rightiesligue1GpGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 852.5
P-value	0.6081
Alternative hypothesis	True location shift is greater than 0

Table A.37: Results of the Mann-Whitney U test for assists per game in Ligue 1

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesligue1ApGonly and rightiesligue1ApGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 1197.5
P-value	0.005317
Alternative hypothesis	True location shift is greater than 0

Table A.38: Results of the Mann-Whitney U test for dribbles per game in Ligue 1

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesligue1Drbonly and rightiesligue1Drbonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 919
P-value	0.3932
Alternative hypothesis	True location shift is greater than 0

Table A.39: Results of the Mann-Whitney U test for shots per game in Ligue 1

Parameter	**Value**
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesligue1SpGonly and rightiesligue1SpGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 973.5
P-value	0.2364
Alternative hypothesis	True location shift is greater than 0

Table A.40: Results of the Mann-Whitney U test for key passes per game in Ligue 1

Parameter	**Value**
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesligue1KeyPonly and rightiesligue1KeyPonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 1000
P-value	0.1745
Alternative hypothesis	True location shift is greater than 0

A.3 Tables for tennis

Table A.41: Results of the Mann-Whitney U test for ATP points

Parameter	**Value**
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiestennisATPonly and rightiestenisATPonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 543
P-value	0.9603
Alternative hypothesis	True location shift is greater than 0

Table A.42: Results of the Welch's t-test for 1st serves

Parameter	**Value**
Test type	Welch Two Sample t-test
Data sets	leftiestennisserveonly and rightiestenisserveonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	t = -2.2587
Degrees of freedom	df = 23.465
P-value	0.9833
Alternative hypothesis	True difference in means is greater than 0
95% confidence interval	(-4.946424, Inf)
Sample means	$\bar{x}_1 = 68.85000$, $\bar{x}_2 = 71.66341$

Table A.43: Results of the Mann-Whitney U test for returns

Parameter	**Value**
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiestennisreturnonly and rightiestenisreturnonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 802.5
P-value	0.2829
Alternative hypothesis	True location shift is greater than 0

Table A.44: Results of the Mann-Whitney U test for net points

Parameter	**Value**
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiestennisnetonly and rightiestenisnetonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 341.5
P-value	0.8167
Alternative hypothesis	True location shift is greater than 0

Table A.45: Results of the Mann-Whitney U test for aces

Parameter	**Value**
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiestennisaceonly and rightiestenisaceonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 500
P-value	0.9838
Alternative hypothesis	True location shift is greater than 0

A.4 Tables for handball

Table A.46: Results of the Mann-Whitney U test for goals per game in handball

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftieshandballGpGonly and rightieshandballGpGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 1355
P-value	0.5801
Alternative hypothesis	True location shift is greater than 0

Table A.47: Results of the Welch's t-test for efficiency of shots in handball

Parameter	Value
Test type	Welch Two Sample t-test
Data sets	leftieshandballEfficiencyonly and rightieshandballEfficiencyonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	t = 0.40956
Degrees of freedom	df = 80.559
P-value	0.3416
Alternative hypothesis	True difference in means is greater than 0
95% confidence interval	(-2.883242, Inf)
Sample means	$\bar{x}_1 = 68.06053$, $\bar{x}_2 = 67.11918$

Table A.48: Results of the Mann-Whitney U test for assists per game in handball

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftieshandballApGonly and rightieshandballApGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = W = 991
P-value	0.9932
Alternative hypothesis	True location shift is greater than 0

Table A.49: Results of the Mann-Whitney U test for shots per game in handball

Parameter	**Value**
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftieshandballSpGonly and rightieshandballSpGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 1320.5
P-value	0.6615
Alternative hypothesis	True location shift is greater than 0

A.5 Tables for basketball

Table A.50: Results of the Mann-Whitney U test for points per game in basketball

Parameter	**Value**
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesbasketPPGonly and rightiesbasketPPGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 15301
P-value	0.4435
Alternative hypothesis	True location shift is greater than 0

Table A.51: Results of the Mann-Whitney U test for field goal percentage in basketball

Parameter	**Value**
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesbasketFGpctonly and rightiesbasketFGpctonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 14751
P-value	0.6191
Alternative hypothesis	True location shift is greater than 0

Table A.52: Results of the Mann-Whitney U test for assists per game in basketball

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesbasketASTPGonly and rightiesbasketASTPGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 15804
P-value	0.2914
Alternative hypothesis	True location shift is greater than 0

Table A.53: Results of the Mann-Whitney U test for rebounds per game in basketball

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesbasketREBPGonly and rightiesbasketREBPGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 16443
P-value	0.1433
Alternative hypothesis	True location shift is greater than 0

Table A.54: Results of the Mann-Whitney U test for offensive rebounds per game in basketball

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesbasketOREBPGonly and rightiesbasketOREBPGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 16535
P-value	0.1271
Alternative hypothesis	True location shift is greater than 0

Table A.55: Results of the Mann-Whitney U test for defensive rebounds per game in basketball

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesbasketDREBPGonly and rightiesbasketDREBPGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 16027
P-value	0.2329
Alternative hypothesis	True location shift is greater than 0

Table A.56: Results of the Mann-Whitney U test for steals per game in basketball

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesbasketSTLPGonly and rightiesbasketSTLPGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 15984
P-value	0.2436
Alternative hypothesis	True location shift is greater than 0

Table A.57: Results of the Mann-Whitney U test for blocks per game in basketball

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesbasketBLKPGonly and rightiesbasketBLKPGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 17164
P-value	0.0493
Alternative hypothesis	True location shift is greater than 0

Table A.58: Results of the Mann-Whitney U test for free throws made per game in basketball

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesbasketFTMPGonly and rightiesbasketFTMPGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 14919
P-value	0.5664
Alternative hypothesis	True location shift is greater than 0

Table A.59: Results of the Mann-Whitney U test for field goals attempted per game in basketball

Parameter	Value
Test type	Wilcoxon rank sum test with continuity correction
Data sets	leftiesbasketFGAPGonly and rightiesbasketFGAPGonly
Sample size	n1 = [number of lefties], n2 = [number of righties]
Test statistic	W = 15576
P-value	0.3576
Alternative hypothesis	True location shift is greater than 0

Appendix B

Data sources

- https://www.whoscored.com/
- https://www.ultimatetennisstatistics.com
- https://sofifa.com/
- https://www.livesport.cz/
- https://ehfeuro.eurohandball.com/men/2022/player-statistics-details/?statistics=goalsfbclid=IwAR0PmPpFtWNWcrk8GEGWCO0DYJcgUSPnZ-DuIqRKRDiI8k52xXnYOHqbYbw
- https://www.basketball-reference.com/
- https://www.nba.com/stats/players/traditional?PerMode=Totalssort=PTSdir=-1Season=2021-22
- https://tabletennis.guide/
- https://www.ittf.com/rankings/