EXAMINING COMPETITIVE BALANCE IN NORTH AMERICAN PROFESSIONAL SPORT USING GENERALIZABILITY THEORY: A COMPARISION OF THE BIG FOUR

by

Mitchell T. Woltring

A Dissertation Submitted in Partial Fulfillment of
the Requirements for the Degree of
Doctor of Philosophy in Human Performance

Middle Tennessee State University
August 2015

Dissertation Committee:

Dr. Joey Gray, Chair

Dr. Minsoo Kang

Dr. Benjamin Goss

ACKNOWLEDGEMENTS

There are many who deserve recognition for helping and supporting me to get to this point. First and foremost I thank my wonderful wife Sarah, without her support none of this would be possible. To my parents, thank you for instilling in me the work ethic and drive to achieve what I desire. A very special thanks to Dr. Joey Gray, Dr. Minsoo Kang, and Dr. Ben Goss, who without their help this dissertation would not have been possible. The expertise that each brought was unparalleled and irreplaceable. I would like to thank Dr. Colby Jubenville for asking me the critical questions with the initial one being, “do you want to be the man?”; effectively convincing me Middle Tennessee State University was the place for me. And finally, to everyone else who has helped shape me in to the person I am today, I am forever grateful.

ABSTRACT

The purpose of this study was to use Generalizability Theory to analyze levels of competitive balance in each of the four major professional sports leagues in North America (MLB, NBA, NFL, and NHL) to determine if Generalizability Theory has merit as a measure of competitive balance, if leagues are competitively balanced based on an absolute determination, and to what extent leagues are competitively balanced relative to the other leagues. The study analyzed a 10-year period for each of the four leagues from 2005-2014 and used game-by-game win/loss data to determine competitive balance. A single-facet, crossed design (Teams (T) x Games (G)) was applied for each of the leagues in each of the 10 observed seasons. G-Studies were performed to estimate the percentages of variance associated with each facet and their interaction. D-Studies were then performed to determine if leagues were competitively balanced based on an absolute decision, as well as how each league's level of competitive balance ranked relative to the other three leagues. The results from the G-Studies showed a majority of the variance for each league came from the interaction term. The D-Studies showed that the NBA was the least competitively balanced of the four leagues and was consequently the only of the four to exhibit an absolute measure of competitive imbalance. MLB was the third most competitively balanced league, while the NHL and NFL were the most competitively balanced leagues. The D-Study results also indicate that Generalizability theory has merit as a method for measuring competitive balance, as the ranking of leagues on levels of competitive balance from the study were comparable to the findings of existing literature utilizing accepted methods for measuring competitive balance.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

Sport brings out passion in people like few things. The level of passion may change from fan to fan, but one thing that all fans can agree on is that a lopsided contest is not fun for either side. Poorly matched teams act as a drain on fans' passion because it undermines the inherent excitement of a contest. For this reason, professional sport leagues, "are in the business of selling competition on the playing field" (Fort & Quirk, 1995, p. 1265). It is the uncertainty of the final outcomes, created by a match of skill during a contest, which drives the very nature of sport and fuels fan's interest by creating excitement. The uncertainty of outcomes in sports is commonly referred to as competitive balance. Competitive balance in a sports league is described as the distribution of wins in a league; and leagues that have a more even distribution of wins among its teams are said to be more competitively balanced (Larsen, Fenn, & Spenner, 2006). Contests become more predictable in a league lacking competitive balance, and predictability of outcomes disengages fans which negatively affects attendance (El-Hodiri & Quirk, 1971). Without fans, sports leagues are not able to thrive financially. As such, competitive balance is paramount to the core product of sport.

Life is not fair, and neither is sport. There are the haves and the have-nots. For every New York Yankee there is a Tampa Bay Ray, for every Green Bay Packer there is a Jacksonville Jaguar, for every Los Angeles Laker there is a Los Angeles Clipper, and for every Montreal Canadien there is a Winnipeg Jet. Some franchises are annual championship contenders while others are annual cellar dwellers. This is because in sport, for every winner there must be an opposing loser. Every team cannot succeed in the win-loss column simultaneously. In response to this, sport leagues have devised policies intended to protect the financial interests of each team, including the cellar dwellers; and in turn, the league as a whole.

In order to keep the league healthy each member must be assured of financial stability – regardless of actual team success. In turn, sport leagues have instituted policies to try and level the playing field and give every team a chance to succeed in any given season. These policies have been instituted in the name of competitive balance (Szymanski & Késenne, 2004). The notion of competitive balance as a core concern for sports leagues is rooted in the idea that fans purchase the product in advance not knowing the outcome. If a league moves towards the point where fans generally know who is going to win and who is going to lose, they will stop purchasing the product and league is said to be competitively *imbalanced* (Owen & King, 2015). Inversely, the more uncertainty that surrounds each contest in a league the more fans crave and purchase the product. The value of competitive balance cannot be understated: it fuels the suspense of unscripted drama as the core product of sport. Each and every team has a chance, and fans are driven to watch by the uncertainty of outcome. It is only through the collaborative effort of individual team owners that sport leagues can work to achieve a level of competitive balance that is desirable to the fans (Horowitz, 1997). As a business practice, a collaborative effort would seemingly counteract the fundamental competitive bedrock of sports.

Thus, professional sport leagues have often been referred to as cartels, a group of owners working together to maintain prices and restrict competition. Pantuosco and Stone (2007), in their competitive balance study on the NFL and NCAA, provided an analysis of the basic economic principles that guide cartel-like league actions. Specifically, they argue that the goal of the NFL and all professional sport leagues is to provide the fans with a product that will keep them engaged and happy, which leads to maximization of league profits. Uyar and Surdam (2012) echo this sentiment by stating, "The owners have long recognized that generating and maintaining fan interest is a key factor for their cartel's long-term survival and prosperity" (p. 480). Producing fan interest is a difficult concept for a cartel though, as it must be sustained at both the franchise and league level.

As a cartel, league members must balance competing with each other on the field while working together to strengthen the league as a whole. What comes to the forefront is a realization by all members that the enduring economic success of the league must come before the immediate interests of individual members. The idea of professional sport leagues operating as cartels eliminates the nostalgic idea of owners acting as "sportsmen" who were not necessarily interested in profits. Vrooman (2009) describes sportsmen owners as those who operate as win maximizers; they are willing to sacrifice profits in order to gain more wins. The opposite of sportsmen are those that operate teams strictly as a business or as profit maximizers; they are concerned with maximizing team and league profits.

As sports have become more profitable over time owners have increasingly moved away from being sportsmen owners to become profit maximizers, placing more of an emphasis on team profits than team wins; which has shifted the entire focus of league goals (Késenne & Pauwels, 2006). The long-held theory on this shift from sportsmen to businessmen holds that as leagues have become a cartel of profit maximizers they have, “worked for or permitted greater on-the-field equality” (Horowitz, 1997, p. 382). In this case, on-the-field equality is better known as competitive balance. The importance of the issue of competitive balance to professional sport cannot be understated. Horowitz (1997) said, “As a general principle *all* sports leagues, whether professional or amateur, profess a desire to ‘maintain competitive balance’” (p. 373). This desire for competitive balance stems “from an assumption that fans have a strong preference for uncertainty of outcomes” (Zimbalist, 2002, p. 112). The concern related to this is that in the absence of competitive balance fans grow tired with the product and no longer spend their time or money watching sports. Understandably, this is a fear for any owner or league official.

The fear of a competitively imbalanced league has led each of the four major North American professional sport leagues to address the subject through various measures (Dietl, Franck, Lang, & Rathke, 2010). The increasingly ardent push for competitive balance, or parity, among North American professional sport leagues has also led to an increase in the number of academic studies focusing on the subject (Crooker & Fenn, 2007). The abundance of academic studies relating to competitive balance in North American professional sport provides a wealth of knowledge to draw from in order to undertake this study. While this study will investigate competitive balance in North American professional sport, it will do so by applying a measurement theory, Generalizability Theory, to yield results that are comparable across each of the four (MLB, NBA, NFL, NHL) major professional North American sport leagues.

Generalizability Theory

The alternative method to for measuring competitive balance in the present study is Generalizability Theory. Generalizability Theory (G-Theory) is a measurement theory often used in educational and psychological research (Briesch, Swaminathan, Welsh, & Chafouleas, 2014) and is used to test reliability of estimates of assessments in these areas. The result produced by a Generalizability Theory study is a coefficient, and if this coefficient meets or exceeds a predetermined minimum level of reliability the researcher can be confident that there is a reliability of estimates. A reliability of estimates means that the researcher is confident that the results are repeatable if the test was run again, even if the conditions of the test were to change. As Briesch et al. (2014) describes it:

> Educators, for example, are not necessarily interested in the score that a student receives on a particular science assessment scored by a particular rater on a particular day but rather in obtaining an estimate of that student's ability in science. Similarly, psychologists are not interested in how a participant responds to a particular measure of self-esteem administered under a specific set of conditions on a particular day but rather in obtaining an estimate of that participant's general level of self-esteem (p. 15).

Because of this, Generalizability Theory is not typically used to test hypotheses, rather, it is used to identify relevant components of variance and use that information to, "design a measurement that minimizes error for a particular purpose (Shavelson & Webb, 1991, p. 83). This is done in two steps, a G-study and a D-study.

The G-study identifies the relevant components of variance. The information from the G-study can then be used to inform and develop relevant D-studies. The purpose of the D-study is to make a decision. In the previous example of student science assessment the G-study would identify how much each area (such as day, subject, rater, a particular test) contributes to the variance in assessing a student's ability in science. This information would then be used to design a D-study which would be used to make a decision. The decision made by the D-study would be whether or not a student meets the predetermined minimum level of science ability. Depending on the purpose of a study, a researcher may be interested in the results of both the G-study and D-study, or simply one or the other.

In addition to its use in education and psychology research, Generalizability Theory has been successfully used to test reliability of estimates in sport and exercise research. Generalizability Theory has been used with much success in the area of physical activity research (Kang, Bjornson, Barreira, Ragan, & Song 2014). Specifically, G-Theory has been used to test physical activity levels of adults with visual impairments and found that wearing a pedometer for six days produced reliable estimates of overall physical activity (Holbrook, Kang, & Morgan, 2013). Ishikawa, Kang, Bjornson, & Song, (2013) used Generalizability Theory to test the minimum number of days of step counts needed to obtain reliable estimates of ambulatory activity for children and adolescents with cerebral palsy. Based on age and function level they found that it takes anywhere from two to eight days of step count data to obtain reliable estimates of ambulatory physical activity. Additionally, Kang et al. (2014) used Generalizability Theory to determine the minimum number of days needed to obtain reliable estimates of physical activity based on step counts for children. In their study of children 2-15 years old they found that anywhere from 2-12 days of step count data was needed to obtain reliable estimates of physical activity, and that boys needed less days of monitoring, on average, than girls. On top of being used to test physical activity levels, Generalizability Theory has been used to test the validity and reliability of the devices used to measure physical activity (Strycker, Duncan, Chaumeton, Duncan, & Toobert, 2006; Holbrook, Barreira, & Kang, 2009; Kang, Bassett, Barreira, Tudor-Locke, Ainsworth, Reis, Strath, & Swartz, 2009). In each of these studies the focus was on determining how many days of monitoring were necessary for a particular step count device in order to obtain reliable estimates of physical activity. The key in each study is G-theory was used to obtain reliable estimates of a measurement.

In the context of competitive balance, a researcher would be most interested in the results of the D-study because it could be used to assess whether or not a sport league exhibits competitive balance. A sport league would not want to see a reliability of estimates because that would be synonymous with a certainty of outcomes, and competitive balance is based on an assumption that there is an uncertainty of outcomes. If at any point during a season the minimum level of reliability is met, it means that the particular season could be played over and over again, under varying conditions, and we would expect the final rankings of teams to remain the same. For a sport league, if the coefficient reaches or exceeds the prescribed minimum level of reliability at a point during the season it would mean that the data is providing, at that point, a reliable estimate of the final rankings for the entire season. For example, if a researcher were to find that at 120 games into a baseball season the coefficient exceeded the minimum level of reliability, they could be confident that at 120 games the season had achieved a reliable estimate of the final rankings. In this case, the season could have been stopped at 120 games and the researcher would be confident that the final rankings exhibited at the end of 162 games would be similar. In effect, it would render the remaining 42 games useless because we already knew what was going to happen after 120 games; indicating a lack of competitive balance. Because of this, testing the viability of Generalizability Theory as a measurement method for levels of competitive balance is appropriate. Therefore, in the setting of the present study Generalizability Theory will be used to provide an estimate of competitive balance in North American professional sport, as well as identify the point during a season at which we could reliably estimate the final rankings, if that point exists.

Scope of the Study

The following study will encompass a 10-year period from the 2004-2005 seasons through the 2014-2015 seasons of Major League Baseball (MLB), the National Basketball Association (NBA), the National Football League (NFL), and the National Hockey League (NHL). The study will use game-by-game win/loss results for each league to focus on competitive balance in each league. This study will focus on the competitive balance of each observed league as a whole to explore if teams are equally matched such that the outcome is uncertain. The coefficient produced by the Generalizability Theory D-Study will be used to measure the level of competitive balance in each league for each observed year. The coefficient ranges from 0-1, with 0 representing a league in perfect competitive balance and 1 representing a league that is perfectly imbalanced. If, at any point during a season, the coefficient exceeds the predetermined minimum level of reliability of 0.8 (as recommended by Briesch et al., 2014), it is determined that the league exhibits a certainty of outcomes, and is not competitively balanced on an absolute term. More specifically, the investigation will place an emphasis on comparing levels of competitive balance between different seasons within each league, and competitive balance amongst the four professional sport leagues. This allows the results to be compared on relative terms as well as the absolute decision of whether or not the minimum level of reliability was met. Whether the coefficient for a particular season of a particular league reaches the minimum level of reliability of 0.8 or not, it can be used to compare between other seasons and other leagues. For example, a coefficient of 0.5 would represent a higher level of competitive balance than a coefficient of 0.6. And while both would be considered surpassing the minimum level of reliability, a coefficient of 0.85 would represent more competitive balance than a coefficient of 0.95. Placing the emphasis on comparison within each league and between the four leagues will be a precursor for providing recommendations for improving competitive balance (if needed) in each league.

Statement of the Problem

There is an established record for examining competitive balance in professional sports. As sport literature has progressed, there are two conclusions that continue to surface. The first is competitive balance is important to professional sport leagues (El-Hodiri & Quirk, 1971; Fort & Quirk, 1995; Zimbalist, 2002; Larsen, Fenn & Spenner, 2006; Uyar & Surdam, 2012; Lenten, 2015). The second is there is no single, standardized method for measuring competitive balance (Sanderson, 2002; Zimbalist, 2002; Fort, 2003; Sanderson & Siegfried, 2003; Evans, 2014, Owen & King, 2015). The persistent predicament in competitive balance research is the need to test alternative methods, weigh them against the methods that have already been used, and eventually attempt to settle on a standardized method for measuring competitive balance. Without a standardized method of measurement, it is difficult for researchers and practitioners alike to agree on relative levels of competitive balance. Without such agreement, implementing new strategies for increasing competitive balance remain an application of educated guesses.

Importance of the Study

The measure of competitive balance in sports is of paramount concern to the four major professional sport leagues in North America, and therefore has directed a considerable amount of research by academics. This sentiment was asserted by Zimbalist (2002) when he said, "Competitive balance is like wealth. Everyone agrees it is a good thing to have, but no one knows how much one needs" (p. 111). What we do know is that professional sports leagues continue to push for as much competitive balance as possible. Generating the suspense of unscripted drama that fans crave speaks to the core product of sport; competitive balance defines the business of sport. As Lenten (2015) says, "Competitive balance is defined as the degree of parity in sports leagues, and helps describe the 'uncertainty of outcome' hypothesis – a cornerstone of understanding the idiosyncrasies of demand in sports economics" (p. 5). Absolute parity, or perfect competitive balance, is represented by a scenario of equality. In professional sports, perfect competitive balance would be characterized by each team in a league finishing with a .500 winning percentage. In such a scenario it would be impossible to determine which team would win any given game because the every team would be evenly matched as confirmed by the standings.

According to the uncertainty of outcomes hypothesis (Lenten, 2015), a league exhibiting perfect competitive balance would generate the highest level of fan engagement, so it is understandable that researchers would be immensely interested in the topic of competitive balance. Goosens (2006) asserts that increased interest in the discipline of team sports economics that led to the creation of many journals, most notably the *Journal of Sports Economics* in 2000. As such, a search of the terms 'competitive balance' and 'parity' in regards to sports will return an exorbitant number of studies published in the *Journal of Sports Economics* over the last 15 years. As a cornerstone for understanding sports economics, competitive balance and parity represent the basis for much of the research that is being conducted in the increasingly important discipline of sport economics. Competitive balance and parity represent a crossroads of significance to both the academic and professional sport communities.

Statement of Purpose

As the current literature reveals, the existing sport literature lacks an agreed upon method for measuring competitive balance. Thus, the purpose of the present study will be to employ Generalizability Theory to analyze relative levels of competitive balance in each of the four major professional sports leagues in North America (MLB, NBA, NFL, and NHL) from 2004-2014 to determine if Generalizability Theory has merit as a measure of competitive balance, if leagues are competitively balanced based on an absolute determination, and to what extent leagues are competitively balanced relative to the other leagues. Using Generalizability Theory will produce a concentration measure of competitive balance with the aim of obtaining results that can be comparable over a number of seasons and between each of the four leagues. By testing an alternative measurement theory for competitive balance the results can be compared to accepted measures of competitive balance to see if Generalizability Theory could produce useful measures of competitive balance. Increasing competitive balance in professional sports is understood as a basic economic tenet of the survival of the leagues; therefore, evidence-based proposals to measure and ultimately increase competitive balance are of the utmost importance.

Definition of Terms

Collective Bargaining Agreement (CBA). The agreements between the players and the team ownership in sports leagues which regulates how leagues will operate. Specific elements covered in CBA’s include, but are not limited to, salary cap structure, free agency restrictions, distribution of league revenues, the league draft, and disciplinary action. The CBA is the primary authority for everyone involved in league business (Forgues, 2012).

Competitive balance. The concept is focused on the expected closeness of contests in sport leagues. It is built on the uncertainty of outcomes hypothesis which holds that fans prefer contests between evenly matched teams. Competitive balance increases as an uncertainty of outcomes increases. As outcomes move towards certainty, the level of competitive balance decreases.

D-Study. One of the two studies used to carry out a Generalizability Theory assessment. The D-Study produces two coefficients which are used to make decisions regarding the purpose of the study. The generalizability coefficient is used to make relative decisions, such as a student's performance relative to that of their peers. The dependability coefficient is used to make absolute decisions, such as a whether or not a student received a passing or failing grade on an assessment.

Dependability coefficient. The coefficient produced by a D-Study which is used to make absolute decisions. It is used as a criterion-referenced decision. For example, a student either passed an assessment or they did not, or a league is either competitively balanced or they are not.

Facets. Facets in a Generalizability Theory study are the same as factors in an ANOVA study. They are any, "set of conditions under which measurements can be carried out" (Cardinet, Tourneur, & Allal, 1976, p. 122). Commonly used facets are subjects, days, raters, forms, methods, and occasions.

G-Study. One of the two studies used to carry out a Generalizability Theory assessment. The G-Study estimates the degree of variance which can be attributed to the facets, or variance components, of a study. The information gained from this can be used to design a measurement which reduces as much error as possible given the purpose of a particular study.

Generalizability coefficient. The coefficient produced by the D-Study which is used to make relative decisions. It is used to make decisions about a measure relative to another object's measure. For example, how a student performed on an assessment in relation to their classmates, or how a league's level of competitive balance compares to the competitive balance of another league.

Generalizability Theory. An approach to research assessment which offers the advantages of being able to simultaneously, "take into account the various sources of error that affect measurement rather than assuming a single source of measurement error" (Briesch, et al., 2014) as well as examine both the generalizability and the consistency of a measure. It allows a researcher to measure the same construct under many different conditions to get a more holistic assessment. For example, Generalizability Theory can provide a researcher with an estimate of a student's general ability in science rather than their score on a particular test on a particular day given by a particular teacher. Likewise, it can provide a researcher with an estimate of a league's general level of competitive balance rather than the competitive balance of a particular team on a particular day against a particular opponent.

Lockout. In sports a lockout occurs when management, or team ownership, initiates a work stoppage. A lockout is typically associated with a labor dispute, such as the process of negotiating a Collective Bargaining Agreement. This is not to be confused with a work stoppage initiated by the employees, which is called a strike.

Luxury tax. In sports a luxury tax is a mechanism to penalize teams who spend too much on player salary. It is levied in situations where leagues do not have a hard salary cap that prevents teams from spending over a certain amount. With a luxury tax, teams pay a predetermined amount to the league for every dollar they exceed the predetermined luxury tax threshold. Luxury taxes are used by both MLB and the NBA since they do not have hard salary caps.

Minimum level of reliability. This is represented by the desired level that must be met by a dependability or generalizability coefficient in order for a researcher to have confidence that the measure is reliable. It is the researcher's responsibility to set the minimum level of reliability, but most choose between 0.7 and 0.9, with 0.8 being widely used (Briesch et al., 2014). Typically, a higher minimum level of reliability, like 0.9, is used if the stakes for the decision are high, such as whether or not a child should be determined to have special needs.

Parity. Parity is term commonly in the place of competitive balance. A league is said to enjoy a high level of parity as the uncertainty of outcomes rises. As the outcome of contests become more certain, leagues exhibit lower levels of parity.

Reliability of estimates. A term used to describe the situation when a dependability or generalizability coefficient has exceeded the minimum level of reliability. When there is a reliability of estimates the researcher is confident that they would find similar results, even if the assessment were carried out under different conditions.

Revenue sharing. Revenue sharing in sports is a system of wealth distribution in which a predetermined amount (or percentage) of individual team revenues is gathered and then distributed to other teams in the league. It is viewed as a necessary component for maintaining competitive balance because it keeps an individual team from making too much more money than any other team. Each of the four North American professional sports leagues has a system of revenue sharing, but they vary greatly in how they operate.

Salary cap. A salary cap for sports leagues is an agreement which determines how much money an individual team can spend on player salaries. It can be levied on a per-player basis or a per-team basis. The NFL and NHL have a hard salary cap, meaning that teams absolutely cannot exceed the level agreed upon for player salaries. The NBA has a soft salary cap, meaning that teams can exceed the level agreed upon for player salaries as long as the reason for exceeding the salary cap falls into one of the many predetermined exceptions in the Collective Bargaining Agreement. MLB has no form of salary cap.

Uncertainty of outcomes hypothesis. This represents the basis of much of the competitive balance research in sports. It is based on the assumption that fans prefer to watch sport contests in which team playing abilities are more evenly matched. It holds that the more evenly matched teams are, the more uncertainty there is surrounding the outcome of the contest, which increases attendance (Knowles, Sherony, & Haupert, 1992).

CHAPTER II

LITERATURE REVIEW

Competitive Balance

Competitive balance is a concept that has occupied a place in the academic literature for over a half a century. Yet, it can be a difficult concept to appreciate because of the numerous methods used to define and measure competitive balance. Essentially, fans desire an uncertainty of outcomes in order to pique their interest. This uncertainty of outcomes may be represented on a game-to-game basis, over the course of an entire season, or even by looking at dynasties covering many years. Neale (1964) originally referred to this as the “League Standing Effect”. In relation to effects on attendance, he stated:

> There is excitement in the daily change in the standings or the daily change in possibilities of change in standings. The closer the standings, and within any range of standings the more frequently the standings change, the larger will be the gate receipts (p. 3).

Neale puts an emphasis on the closeness of the standings. If every team in a league had the same record, that league would have achieved the pinnacle of competitive balance – it would have absolute parity. There would be an uncertainty of outcome for every contest between all teams. Vrooman (2009) refers to this hypothetical situation as the ‘Theory of a Perfect Game’. He says:

> In theory the perfect game is a symbiotic contest between evenly matched opponents. The practical economic problem is that games in professional sports leagues are played between teams from imperfectly competitive markets that form imperfectly competitive natural cartels. Fortunately the natural duality of sports leagues implies that any single team is only as strong as its weakest opponent and the success of any league ultimately depends on the perfection of its games (p. 1).

As Vrooman alludes, we do not see the Perfect Game scenario in professional sport; and therefore, research on the subject focuses on answering why it does not occur. While strides have been made toward this end, a need still remains for a standardized method for measuring competitive balance among professional sports.

Academic literature on competitive balance finds it roots in Rottenberg's (1956) Coasian argument about Major League Baseball's labor market. His proposal, which was written 20 years before MLB players were granted free agency, stated that the reserve system and a free agency system would lead to the same distribution of talent across MLB teams. Key to Rottenberg's argument, especially to the players, was that free agency would help move player salaries toward their marginal revenue products, thus weakening the monopoly that MLB had exercised in the exploitation of their players. Additionally, he claimed that revenue sharing would have no positive effects on the distribution of talent, and that its effects could be seen in a more profound exploitation of the players.

Neale's (1964) League Standing Effect, or fans' desire for an uncertainty of outcomes has been empirically tested numerous times. Predictability in outcomes has been shown to decrease attendance (El-Hodiri & Quirk, 1971). Predictability in outcomes has even been shown to negatively affect dominant teams. Quirk and Fort (1992) found that the dynastic Cleveland Browns of 1946-1949 saw major decreases in attendance because of league issues with competitive balance, despite winning four-straight league championships.

The desire for uncertain outcomes and an avoidance of the above mentioned case have led sports leagues to institute policies such as salary caps and revenue sharing. Noll (1974) identified the intended effect of such policies going beyond simply avoiding bidding wars where teams with less money would not be able to participate. While that is one of the intended effects, the other is to reduce the overall discrepancy in quality between the top and bottom teams in a league. In their analysis of the demand for Major League Baseball, Knowles et al. (1992) also found uncertainty of outcome to be a significant predictor of fan attendance. Given these findings, it is no surprise that it is standard league policy to institute measures meant to increase competitive balance.

Without fan interest, sports leagues would find their stadiums emptying, their television contracts dwindling, and their place as a top-15 industry in the United States all but gone (Milano & Chelladurai, 2011). There are many issues that each professional sport league has to deal with, with many of the issues influencing others. On the importance of competitive balance, Sanderson (2002) says that for any sports league, maintaining an uncertainty of outcomes within the league falls at the top on a long list of issues that must be addressed. As such, we have seen a concerted effort from each of the commissioners of the four major professional sports in North America to place an emphasis on growing and/or maintaining competitive balance in their leagues.

Major League Baseball. In Major League Baseball (MLB), the commissioner's office has reports dating back to 1995 regarding a lack of competitive balance (Zimbalist, 2002). As a result of this, Commissioner Bud Selig commissioned the *Blue Ribbon Economics Committee* to provide a report on the state of the game as it related to competitive balance. In an interview with Richard Justice (2014), Selig recalled one of the first reports he received from Paul Volcker, former chairman of the Federal Reserve and a member of the *Blue Ribbon Economics Committee*. Volcker told Commissioner Selig, "You've got a problem, you've got 25 teams that can't really win" (p. 1). Two years later, in 2000, the committee published *The Report of the Independent Members of the Commissioner's Blue Ribbon Panel on Baseball Economics* (Levin, Mitchell, Volcker, & Will, 2000). This report detailed many of the shortcomings of Major League Baseball at the time, which focused on the five-year period from 1995-1999. Much of the shortcomings are summed up in a quite damning statement:

> MLB has an outdated economic structure that has created an unacceptable level of revenue disparity and competitive imbalance over the same period. The growing gap between the "have" and the "have not" clubs—which is to say the minority that have a realistic chance of succeeding in postseason play and the majority of clubs that have poor prospects of reaching the postseason—is a serious and imminent threat to the popularity, health, stability and growth of the game (p. 11).

Despite the lack of supporting evidence to corroborate their conclusions, including no comparisons to any periods previous to 1995-1999 (Eckard, 2001), Major League Baseball began enacting swift changes to adhere to the Panel's suggestions to increase competitive balance.

As a result of the Panel's report, improving competitive balance has been a cornerstone of MLB administration since then. When asked at the 2014 All-Star Game what he will be most proud of regarding his legacy, Commissioner Bud Selig said it will be baseball's new age of competitive balance (Felder, 2015). Selig went as far to declare that the reforms he oversaw related to competitive balance will be the defining legacy of his time as commissioner (Barrabi, 2014). The now retired commissioner is clearly staking his lasting reputation on baseball's continuing increase of competitive balance.

In an interview with Bob Nightengale (2014), Commissioner Selig spoke unabashedly about the steps he has taken as commissioner to change the game of baseball. On the topic of competitive balance he said:

> Isn't this wonderful? This is what we set out to do. We have, unquestionably, more competitive balance than at any time in history. We had to do a lot of things to achieve this, make changes in the economic system, and it has led us to where we are today (p. 1).

Selig goes on to talk about where the game was and what it has evolved to, by stating:

> We used to have teams year after year who couldn't make the playoffs, and I couldn't say it, but you knew they couldn't make the playoffs. Our job was to

> provide hope and faith in as many places as possible, and today, we have that. It really is remarkable. The fact that this is the best 10 years in attendance in history is no accident (p. 1).

Not only has competitive balance been a focus of Major League Baseball administration, but the commissioner points to it as the foremost reason for major economic changes that have occurred under his leadership.

With all the changes baseball has undergone during Selig's time as commissioner, there is an obvious need to evaluate where the game currently stands. Did Bud Selig's plan really work, and will it continue to influence baseball policy as a new commissioner, Rob Manfred, has begun his tenure? To answer the first question the existence of subjectivity must be noted. If an individual prefers to see truly great teams and dynasties they are less likely to agree with Commissioner Selig's overall plan to begin with. A push for greater competitive balance works against individual teams having extended runs of success because the focus is on each team having an annual chance to succeed, so dynastic franchises are viewed as part of the problem. If an individual values consistent unpredictability they are more likely to agree with Commissioner Selig's overall plan because in a league of true competitive balance fans never know who is going to succeed in a given year. Tom Verducci (2014) of Sports Illustrated manages to capture the general attitude of baseball's current state as well as how it relates to Selig's influence:

> So people generally have come to like this new democratic baseball world that keeps teams in the race. A generation has grown up with expanded postseason play – now with 10 entrants instead of two – and has come to know postseason

> baseball as October Madness, not the coronation of kingly team. Selig was successful at executing his blueprint (p. 1).

Selig is justified in resting his legacy at the foot of competitive balance. He has made a significant contribution to attempt to increase competitive balance in baseball, and fans generally agree with the changes he has made and the resulting state of the game.

The second question is whether or not Major League Baseball will continue to focus so heavily on competitive balance now that Bud Selig has stepped away and Rob Manfred has assumed the role of Commissioner of Major League Baseball. Since assuming the role on January 25, 2015, Manfred's comments have indicated he will continue to treat competitive balance as a cornerstone issue for Major League Baseball. In an interview with Baseball America's John Manuel (2015) regarding the structure of the league's draft, Manfred said he feels the MLB Draft is an important structure for maintaining competitive balance in the league. Manfred has also cited his desire to institute an international draft in order to increase competitive balance and help the game of baseball in general (Weinstein, 2015). Further, in a trip to Tampa Bay to meet with Rays players and officials Manfred stated that a priority in anything MLB does, draft related or not, is maintaining competitive balance (Topkin, 2015). In his short time as MLB Commissioner, Rob Manfred has made clear that he intends to continue treating competitive balance as a foundational issue for Major League Baseball.

National Basketball Association. The National Basketball Association (NBA), much like Major League Baseball, recently saw its longtime commissioner step down from his position. David Stern stepped down after 30 years on the job and was succeeded by Adam Silver. During Stern's tenure as commissioner, his NBA has often been cited as the least competitively balanced of the four major professional sports in North America (Schmidt & Berri, 2003; Vrooman, 2009; Rockerbie, 2012). Because of this, Stern made it a priority to increase competitive balance in the league.

In his first year as NBA Commissioner, David Stern enacted his first policy to increase competitive balance by adopting a soft salary cap. The NBA's salary cap is considered 'soft' because it allows for exceptions in which teams can exceed the salary cap in order to keep star players (Késenne, 2000). By comparison, the NFL and NHL have hard salary caps that do not allow for exceptions. Even with the inception of the soft salary cap in 1984 the NBA has exhibited a three-decade-long run as the least competitively balanced of the four leagues, highlighted by only nine teams winning an NBA Championship in this time period (Rockerbie, 2012). The relative failure of the soft salary cap did not dissuade Stern from continuing to focus on competitive balance.

Stern always envisioned an NBA in which competitive balance was better, and the league was not dominated by only a few teams. In fact, Stern was known to refer to competitive balance quite often, whether he was talking about balance sheets or standings (Bryant, 2011). Bryant goes on to add Stern would like to model the NBA after the NFL. This is assumed to be because the NFL has the most parity of the four professional leagues by a substantial margin (Rockerbie, 2012). After years of watching the NBA dominated by a few teams, Stern made his biggest play for increasing competitive balance when it came time to negotiate a new Collective Bargaining Agreement (CBA) in 2011. As the main negotiator for the NBA owners, Stern took on the players with the aim of making significant changes to the league's economic structure. Stern was looking for three core changes (Grow, 2012): 1. Eliminate guaranteed contracts for players. 2. Significantly reduce the players' share of total league revenues. 3. Eliminate the exceptions of the soft salary cap and turn it into a hard cap. The 2011 NBA season went into a lockout because Stern felt so strongly about the changes that needed to be made. He justified the lockout in the name of competitive balance (Howard, 2011).

Prior to the lockout Stern made his feelings on the subject well known by stating, "Our goal for our teams, our players, but particularly our fans, is to come up with a model that says that every NBA team can compete" (Garcia, 2010, p. 1). He did not waver in this statement as the league did in fact suffer a lockout stemming from the inability for the player's association and ownership unable to come to terms on a new CBA that reflected Stern's goals. When the new CBA was agreed to in 2011 it was championed as achieving two significant long-term objectives:

> One goal of the CBA was to reduce losses and help teams make a profit. The other goal was increased parity via a system that disperses players to all markets and prevents high-revenue producing teams from hoarding top talent and making it more difficult to create super teams (Zillgitt, 2013, p. 1).

Although the league championed the new CBA as a success, the only concrete achievements were reducing the players' share of total league revenues from 57% to a flexible amount ranging from 49 – 51.2% and increasing the luxury tax penalties on teams for exceeding the salary cap (Grow, 2012). Otherwise significant changes were not made to the structures of guaranteed contracts or the status of the soft salary cap. It is difficult to determine if the changes will have lasting effects on competitive balance in the league, but David Stern was quick to react in a matter he considered at the heart of competitive balance.

Almost immediately after agreeing to the new CBA and salvaging a majority of the 2011-2012 season, Stern took an opportunity to prove his commitment to competitive balance. Only a week after the new CBA was signed in November of 2011 David Stern intervened in a three-team trade by the then-owned-by-the-NBA New Orleans Hornets, the Los Angeles Lakers, and the Houston Rockets. The trade was considered fair by most basketball people (Wood, 2011), but was rejected in the name of competitive balance because Chris Paul was going from the small-market Hornets to the big-market Lakers. According to Mark Conrad, professor of law and ethics at Fordham, Stern viewed his rejection of the trade as necessary for allowing small-market to compete. In the wake of agreeing to the new labor contract he did not want to give an impression of inconsistency on the topic of competitive balance (Wood, 2011). Stern had invested a lot of league time, energy, and resources in negotiating the 2011 CBA and was ready to defend the fruits of those labors.

Wood (2011) provided a summary of Stern’s thought process when he claimed that the overarching goal of the 2011 CBA was to create a healthier league where each team, regardless of market size, was given the leverage and resources to compete every season. It is clear that from the beginning of his tenure as NBA Commissioner in 1984 that David Stern worked towards improving competitive balance for the league. Despite all of his efforts, it is unclear whether or not the efforts have been in vain. It may be too early to tell if the changes in 2011 will ultimately increase competitive balance, but we do know that David Stern will not be the person overseeing the progress of the league going forward. Three years after negotiating the 2011 CBA Stern stepped down as commissioner and handed the reins to Adam Silver.

It was widely assumed as commissioner, Adam Silver would continue to push the same progressive policies in relation to competitive balance that his predecessor had. Silver had served under Stern for 22 years before his appointment to commissioner, and was even described as Stern's 'attack dog' during CBA negotiations, playing, "the bad cop to David Stern's good cop in their routines to the press" (Hellin, 2013, p. 1). Silver's statements since becoming commissioner indicate he will continue to push for increased competitive balance. In an interview with New York Post columnist Ted Bontemps (2013), Silver said that the goal of the league is to have a system where all teams have roughly the same resources for building their franchises and competing on an annual basis. Silver believes competitive balance is paramount to the long-term success of the NBA.

Silver sees the level of competitive balance in the few years since the 2011 CBA as a vast improvement, and seems fixated on continuing that trend. His thoughts on the matter are summed up in an interview with Frank Isola (2015):

> What we're seeing is that the size of the market and the resources of an individual owner are less relevant. This is what we had hoped would happen, that with revenue sharing every team would have a better opportunity to compete. By

> having a harder cap what we're seeing is that regardless of market size teams can compete (p. 1).

When referring to a 'harder' cap, Silver is alluding to the changes in the luxury tax system, which is more punitive for teams that exceed the salary cap. The cap itself did not actually change from a soft to a hard salary cap. Like his predecessor, Adam Silver communicates the importance of increasing competitive balance in the NBA, and the policies enacted under his and Stern's terms as commissioner show a resolute commitment to competitive balance. Adam Silver has only been the Commissioner of the NBA since February 1, 2014, and to date has had very few opportunities to enact any real change as it relates to competitive balance. When the NBA's new television deal begins in the 2016-2017 season there will be a large influx of money in the league which will significantly affect the league's salary cap. Any changes enacted by the league leading up to or shortly after this event may prove Silver's first chance to show his commitment to the promotion of competitive balance. Going forward it will become clearer whether or not his work as commissioner brings about actual increases in competitive balance.

National Hockey League. The National Hockey League (NHL) often finds itself as an afterthought in North American professional sports. In terms of its size, the NHL saw league revenues of $2.6 Billion for the 2012-2013 season, compared with $4.6 Billion for the NBA, $7.1 Billion for MLB, and $9.2 Billion for the NFL (Plunkett Research, 2014). In reality, the NHL does appear to be an afterthought as the NBA nearly doubles, MLB nearly triples, and the NFL nearly quadruples their revenues. Being smaller in terms of their revenues does not change how the NHL views the importance of competitive balance; they position it just as importantly as the other leagues.

In a press conference preceding the 2014 Stanley Cup Finals, NHL Commissioner Gary Bettman led off by stating, “This will be our 15th consecutive year without a repeat champion, which is a testimony to the NHL’s unparalleled competitive balance” (Heika, 2014). It is no coincidence that he led off his state of the union address by addressing the topic of competitive balance. As NHL broadcaster Daryl Reaugh says regarding competitive balance in the NHL, “It’s Gary Bettman’s dream, he wants everyone to get a chance, and here we are” (Dowbiggin, 2012). In terms of championship teams no other league can boast the same claim about a lack of repeats. Bettman’s dream of a competitively balance NHL appears to be approaching, if not already a reality.

Similar to David Stern’s insistence on measures to ensure competitive balance during CBA negotiations, Gary Bettman has engaged in extreme tactics in order to move the NHL in the direction he envisioned. Under Bettman’s leadership the NHL has engaged in three lockouts, in 1994-1995, 2004-2005, and 2011-2012, each of which has either caused games to have been missed or the entire season to be missed. The 1994-1995 lockout shortened the season from 84 to 48 games and ended with the NHL winning minor concessions from the players that included a rookie salary cap, which was a far cry from their intention of instituting a league-wide salary cap (McIndoe, 2014).

Ten years later the 2004-2005 lockout resulted in the cancelation of the entire NHL season, as Commissioner Bettman was resolved to achieve the results he had sought a decade earlier. His comments at the time reflect this resolve. In response to stalled negotiations Bettman said, “Until the union is willing to acknowledge and address the economic problems we’re having…the negotiations are not going to progress (Piercy, 2012). On the importance of the negotiations Bettman stated, “These negotiations are not just about next season. It is next season and all the ones that will follow. It's about the future of our game” (Piercy, 2012). Finally, when addressing the lack of an agreement as the season approached Bettman said, “I will give you the same response that I gave you the two times the union made the proposal — the proposal doesn't address our problems. The proposal is flawed in many respects including the assumptions that it makes” (Piercy, 2012). Commissioner Bettman remained steadfast in his demands to overhaul the NHL economic structure.

The NHL work stoppage became the worst in modern sports history, as it was the first time a professional sports league lost an entire season, it included the most games lost in a work stoppage, and it was the longest-lasting shutdown in the history of pro sports (Staudohar, 2005). In the end, the NHL benefitted from Bettman’s determination as the owner’s won out on almost every issue under contention, including the all-important league-wide salary cap (Staudohar, 2005). It took an extreme measure, but the NHL and player’s union agreed on a 6-year CBA that saw Commissioner Bettman’s objectives met.

As the CBA agreed to in 2005 was set to run out, the NHL under Commissioner Bettman again locked out the players in 2012. At issue under this lockout were matters of: revenue sharing, escrow payments, international player participation, the salary cap, roster sizes and guaranteed contracts, unrestricted free agency, player safety, division realignment, long-term front-loaded contracts, the NHLPA executive director's successor, and negotiating strategy (Kobritz & Levine, 2013). In all, the 2012-2013 NHL lockout caused a shortening of the season from 82 to 48 games.

Much like the resolution to the 2004-05 lockout, the NHL and its owners prevailed by a wide margin when all the dust settled as they reduced the players' share of hockey-related revenue, imposed limits on free agent contract length, and increased revenue sharing (Staudohar, 2013). Proteau (2012) described Bettman's negotiating style during CBA negotiations as one that demanded "total parity". Through three labor shutdowns, of which the last two were particularly aggressive, Gary Bettman has effectively implemented his plan to increase competitive balance in the NHL. Stubits (2013) confirms this notion by claiming the reason there hasn't been a repeat Stanley Cup Champion in 15 years is because parity does exist.

National Football League. In the world of competitive balance, the National Football League (NFL) is king. Not only is it the most competitively balanced league in all of sports, it is also the most powerful economically (Quirk & Fort, 1992; Vrooman, 2009; Rockerbie, 2012; Lenten, 2015). It is not by coincidence that the NFL holds this title; they employ a hard salary cap and free agency system that promote competitive balance (Larsen, et. al, 2006). The hard salary cap prevents any team, regardless of how much money they make, from spending more on players than the 31 other teams in the league; unlike the NBA. As a result of the hard salary cap teams are not able to offer a significantly higher salary to top players to entice them to play for their teams once they hit free agency; unlike MLB. Additionally, they have the most extensive revenue-sharing program among professional sports which keeps each franchise economically healthy even in years of poor performance; more so than the NHL, NBA, and MLB (Vrooman, 2009). Each of these structures is in place to build the economic model that the NFL desires. Vrooman (2015) states unequivocally, “The economic model of the NFL seeks absolute parity among the clubs so that any given team can defeat any other” (p. 110).

The NFL through the years has laid the groundwork for a model that consistently produces competitive balance. Their combination of a salary cap, free agency, and a revenue-sharing system made possible through behemoth television contracts provide the foundation for competitive balance that the NFL and its administration pride themselves on (Posnanski, 2014). The foundation for a cohesive NFL was laid by former NFL Commissioner Paul Tagliabue and has been carried on by the current commissioner, Roger Goodell. In 1992, while arguing in favor of new free agency rules to grow competitive balance, Commissioner Tagliabue said:

> Our league is a joint and common venture of the teams, it is unique because the teams join together to present one product, the National Football League, and they compete on the field but join together in a business sense to share the revenues, the profits, and the losses (George, 1992).

Tagliabue retired in 2006 following a successful CBA negotiation between the owners as players. At the time he had two years remaining on his deal as NFL Commissioner, but decided to step down, citing the continued proliferation of competitive balance the new CBA ensured (Matuszewski, 2006). His successor would be Roger Goodell, who remains the Commissioner of the NFL as of the writing of this paper.

Goodell succeeded to the Commissioner role after serving five years as the league's chief operating officer, and was widely considered Tagliabue's most trusted advisor. When faced with heading another round of CBA negotiations in 2011, Goodell stressed the importance of maintaining competitive balance much like his predecessor had. In speaking with Kansas City Chiefs' season ticket holders, Goodell stated:

> There are systems that we have to make sure that we maintain [competitive balance]. When you come into a season, every fan thinks that their football team has a chance to win the Super Bowl and that's what I believe the 32 clubs are working towards. There are several issues that need to be addressed (regarding competitive balance) in (negotiation) this Collective Bargaining Agreement (Looney, 2011).

As the 2011 CBA negotiations approached, the NFL locked out the players in an attempt to force the players to accept the NFL's terms on a new rookie pay scale, an increased share of league revenues for owners, and two additional regular season games (Feldman, 2011).

Locking out the players, and ultimately preventing them from earning a salary had worked for the NFL in the past, but in this instance the players instituted a new strategy; they decertified their union in order to bring an antitrust suit against the NFL. After several court decisions the NFL and the players agreed on a new 10-year CBA in which the NFL won on the rookie pay scale and increases in league revenues for owners, and the players won by keeping the regular season schedule at 16 games (Grow, 2013). Despite victories for each side, Goodell won the concessions he was most interested in and agreed to a new CBA that seems to favor the NFL and their quest to maintain competitive balance (Deubert, Wong, & Howe, 2012). Commissioner Goodell achieved what his predecessor had done in both 1992 and 2006 – he negotiated a CBA that paved the way for the most competitively balanced league in professional sports to maintain their position. The economic structures that are currently in place as results of the 1992 and 2006 CBA's are identified as the main factors which have led to the NFL exhibiting the highest levels of competitive balance among the four professional American sports in almost all academic research on the topic.

Economic Structures Affecting Competitive Balance

Based on the literature and the words of the most powerful leaders in each sport it is undeniable that competitive balance is important. As Sanderson (2002) says, there is nothing more important in professional sports than the production and maintenance of competitive balance. Because of this importance we see Roger Goodell and the other league commissioners' talk about the economic systems, or restraints that they have in place to promote competitive balance. One of the issues consistently at hand is determining how competitive balance can be increased. Sanderson and Sigfried (2003) describe this issue:

> Every sport and sports league has had to confront the fundamental issue of relative strengths among competitors. There has not been a uniform, one-size-fits-all approach or set of rules to resolve this problem. Inasmuch as uncertainty of outcome is a key component of fan demand, wide disparities in inputs and, thus, in likely outcomes are seen as inimical to the long-term health and financial

> viability of the individual enterprise. How to handle weak teams or inferior opponents—to prevent lower quality competitors from free-riding on higher quality rivals—can be as much or more of a problem as dealing with perennially strong ones, because there is at least some interest in seeing the very best individual performers and teams (p. 258).

As Sanderson and Sigfried (2003) say, there is no uniform, one-size-fits-all approach to resolving the issues of competitive balance. However, Peeters and Szymanski (2014) describe the economic systems and restraints typically exhibited by North American sport leagues as primarily in place to promote competitive balance. Based on comments from the top administrators in each sport we know the motivation for the economic restraints is a maximization of competitive balance. Peeters and Szymanski (2014) also list the economic restraints adopted by the leagues, stating that they are justified in the name of competitive balance. Key among the restraints are salary restrictions and free agency, television broadcast rights, and revenue sharing.

Salary restrictions and free agency. Each of the four major sports leagues in North America has a form of salary restriction, either in salary caps or luxury taxes, as well as a free agency system which allows for player movement from team to team. Payroll caps, whether in the form of salary caps or luxury taxes have been shown to improve competitive balance (Quirk & Fort, 1992; Fort & Quirk, 1995; Vrooman, 1996; Rascher, 1997; Késenne, 2000; Dietl et al., 2012). Salary caps are often accompanied by salary floors, meaning that teams have a minimum they *have* to spend along with the maximum they cannot exceed. In leagues where there is no salary cap, a luxury tax is used to restrict salaries. A salary cap may be a hard cap or soft cap. If a salary cap is hard, it means that teams are not allowed to exceed the dollar amount for player salaries. If a salary cap is soft, it means that teams are allowed to exceed the dollar amount but will pay a penalty for doing so (Keefer, 2015). Additionally, each league has its own rules in place to govern free agency and support a more even distribution of talent.

MLB. Unlike the other three major sport leagues, Major League Baseball does not use any form of a salary cap to limit team spending on players. What MLB does use, is a luxury tax that they refer to as a 'competitive balance tax'. Dixon (2013) explains this tax and its purpose:

> The [luxury tax] is a surcharge on the aggregate payroll of a sports team that exceeds a predetermined limit set by league officials. The "luxury" tax was conceived to prevent large-market teams (those with the most money) from buying all the best players in the league, and in effect buying championships. The main goal of the "luxury" tax is to create a more balanced league, because redistribution among clubs counteracts financial imbalances (p. 524).

While teams are not prevented from going over a certain amount of payroll, they are required to pay a tax on the amount of which they exceed the predetermined luxury threshold. Delgado (2014) explains the rules of MLB's luxury tax including its calculation, how it is enforced, and how collected luxury tax money is distributed.

The luxury tax threshold, or the amount at which teams have to pay a luxury tax for exceeding, is agreed to in the CBA. The current CBA for MLB runs through the 2016 season, and the threshold is set at $189 million. The rate of taxation for a club that exceeds $189 million in team payroll is a dynamic rate that changes based on the team's history related to the luxury tax threshold. Under the current CBA, teams could pay anywhere from 17.5% to 50% on the amount that they exceed the threshold. The rate begins at 17.5% and continues to rise for each consecutive year that a team remains above the luxury tax threshold, maxing out at 50%. If a team is able to get below the threshold for one year they are able to reset their tax rate back down to 17.5%.

Despite being called a 'competitive balance' tax by MLB, the money collected is not actually redistributed to other teams in the league. Rather, MLB's extensive revenue sharing program covers any redistribution from large market to small market teams. Taxes collected from teams that exceed the threshold are distributed in four ways. First, an initial $5 million is held in case there is a need for a future refund. Then, the remaining money is split between three sources. 50% of the money goes to fund player benefits, 25% goes to fund baseball development in countries without high school baseball, and 25% goes to MLB's Industry Growth Fund. If there are no refunds that need to be issued, the initial $5 million that was held is distributed to the Industry Growth Fund. According the MLB CBA (2012) the Industry Growth Fund is used to enhance fan interest in the game, increase baseball's popularity, and ensure industry growth.

MLB's competitive balance tax may not directly allocate funds to smaller market teams from the taxes collected, but it does appear to curb spending as only five teams have exceeded the luxury threshold since its induction in 2003 (Delgado, 2014). Of those five teams, the Angels and Tigers have each paid a tax once, equally just over $2 million total. The Dodgers have paid each of the last two seasons, including 2014 in which they set the record for highest payroll ever. The Red Sox have paid six times, but not since 2011. The Yankees are the clear outlier when it comes to luxury tax, having paid each year since 2003.

NBA. There are many nuances to the NBA's salary restrictions and how it affects free agency. In order to parse out what is relevant to the purpose of the study, the author defers to the judgment of Larry Coon on much of the information regarding the salary cap. Coon is a computer scientist and Information Technology Director at the University of California, Irvine. He is known as the NBA salary cap guru and operates the website *cbafaq.com*. Coon is so widely considered the NBA salary cap guru that he is cited more frequently than basketball inventor James Naismith (Beck, 2009).

The NBA has a soft salary cap that was first instituted in 1984 with the purpose of improving competitive balance, although the evidence shows this has not occurred relative to the other major professional sport leagues (Rockerbie, 2012). The NBA's soft salary cap is based off of the NBA CBA which stipulates that each year the cap be set at 44.74% of Basketball Related Income (BRI); therefore as the league makes more money the salary cap continues to rise. The current CBA runs through the 2020-2021 season and also stipulates a salary floor of 90% of the cap that teams must spend on payroll (Coon, 2015). The reason for having a soft cap rather than a hard cap is to allow teams to retain their own players. In reality few NBA teams are actually under the salary cap during the season because the soft cap allows so many exceptions for teams to sign and/or trade for players (Coon, 2015). Since teams do not have to stay under the salary cap, teams that make more money can routinely add players and their salaries as long as they fit into one of the numerous exception categories and simply pay the fines for being over the salary cap. Teams that do not make as much are not afforded this same opportunity. Therefore, larger-market teams are still able to spend more even with a form of a salary cap.

Since the NBA operates under a soft cap with many exceptions, the league has also instituted a luxury tax, or fine system in an effort to control team spending over the cap. Essentially teams are fined on a percentage basis for the amount they exceed the salary cap each year. Coon (2015) describes how the luxury tax is set up, enforced, and how luxury tax funds are later distributed. The luxury tax threshold must be crossed in order for a team to incur payment under the system. Like the salary cap, the luxury tax threshold, or tax level as the NBA refers to it as, is determined by the CBA and changes based on BRI. Under the current CBA the tax level is set at 53.51% of BRI – almost 10% higher than the salary cap to reflect the allowance of exceptions. Beginning with the 2014-2015 season, the rate of luxury tax increased for repeat offenders, similar to MLB's luxury tax. The new luxury tax rate for a first-time offender falls between 150% and 375% based on the amount they are over the threshold. A repeat offender adds 100% to those numbers, meaning their tax rate would fall between 250% and 475% based on the amount they are over the threshold.

Distribution of luxury tax money is set up to be done at the discretion of the league, unlike MLB who have a set percentage paid out to three areas. Up to 50% of tax money can be distributed to non-tax paying teams, although there is no requirement to give it to the non-paying teams. All money that is not distributed to teams must be used for 'league purposes' – so at least 50% of tax money will be used for league purposes each year. Since it is up to the discretion of the league on how to distribute tax money, they can basically do whatever they want with it, including giving it back to teams. In fact, this is exactly what they did in 2011 as 100% of tax money was used to fund the league's revenue sharing program. Since then the practice has been to distribute 50% of luxury tax money to non-paying teams in equal shares and to put the other 50% towards league revenue sharing. This ensures that all luxury tax money gets redistributed throughout the league.

The complicated combination of the soft salary cap and luxury tax system make for many peculiarities and exceptions for free agency in the NBA. While a team cannot exceed the salary cap to sign or trade for a player, they can do so if it falls into one of the many exceptions. According to Coon (2015) there are at least eleven exceptions that allow for this to happen. Another rule agreed to in the CBA that affects free agents is that the NBA employs a maximum contract value. This is another complicated formula in which a maximum amount of years and total value of a contract is determined. A main determinant of this is the type of contract that the player is signing. Due to the number of exceptions allowed, there are at least 16 different contract scenarios that the NBA uses to determine maximum length and monetary value of contracts.

The NBA's use of a soft salary cap combined with a luxury tax system is in place to improve competitive balance by achieving a sense of balance among team spending (Epstein & Kisska-Schulze, 2014). Unfortunately, this combination has not had the intended results as the NBA still exhibits the least competitive balance of the four major professional sport leagues (Schmidt & Berri, 2003; Vrooman, 2009). This seems counterintuitive, especially compared to MLB which has no salary cap structure at all, and only uses a luxury tax system which Rockerbie (2012) describes as relatively useless when it comes to encouraging balance among club payrolls. Because of the lack of any salary cap system in MLB, the relative lack of competitive balance in the NBA appears paradoxical when compared with MLB. The lack of competitive balance in the NBA makes more sense in comparison to the NHL and NFL because of their hard salary cap systems.

NHL. As earlier detailed the NHL has gone through a myriad of labor issues in the last 30 years leading to several work stoppages. These work stoppages have primarily been due to unresolved issues that led to competitive balance problems for the league. Two of the most contested issues in these labor agreements had been the salary cap and free agency. This did not change with the most recent round of CBA negotiations in 2012. The hard salary cap now in place was so important to the NHL that the owners forfeited the entire 2004-2005 season in order to obtain it (Kobritz & Levine, 2013). Despite obtaining a hard cap following the lockout season, many clubs still reported losing money due to the players receiving 57% of hockey-related revenue (HRR). Since the portion of revenue that goes to players is used to determine the hard salary cap, the ‘cost certainty’ that the NHL was seeking by instituting the hard salary cap was not being achieved (Kobritz & Levine, 2013).

The agreement reached in early 2013 on a new CBA brought some changes to the NHL’s hard salary cap. Key among the changes in the 10-year CBA was a new 50-50 split on HRR which lowered the player’s share by seven percent; much like the recent

NFL and NBA negotiations, this was the most important issue at hand (Staudohar, 2013). In order to get the players to agree to a decrease in their share of the pie, the NHL agreed to pay the players a $300 million over the first three years of the deal to initially offset the reduction in the salary cap. This payment effectively moved the player portion above 50% for the first few years of the deal (Staudohar, 2013).

Another key issue for both sides in the 2012 negotiations was how to reform free agency. Under the previous CBA players whose contracts ended could not become unrestricted free agents until they had reached age 27 or had been in the league for seven years (Kobritz & Levine, 2013). If they did not reach either of these statuses they were treated as restricted free agents, and if new teams wanted to sign them they could have draft pick compensation attached to their signing, similar to the NFL's restricted free agency rules. Under the new CBA (NHL, 2013b), restricted free agents and their accompanying draft pick compensation is tied to the salary at which their new team signs them. If the salary is under $1,110,249 annually, there is no draft pick compensation to their former team. However, this sliding scale of draft pick compensation goes all the way up to a restricted free agent's former team receiving four first-round picks if their new salary averages over $8,410,976 annually (NHL, 2013b). This effectively keeps most teams from attempting to sign star players who are restricted free agents.

An issue at hand that involved both the salary cap and free agency rules was the existence of contracts in which teams would take advantage of loopholes in order to try and circumvent the salary cap. Kobritz and Levine (2013) referred to these as 'long-term, front-loaded contracts', where teams would sign players to long contracts in which the average value would sharply decline after the first few years in order to keep the team out of salary cap trouble in the twilight of the deal. The new CBA quashed these type of contracts in two ways. First, they put a seven-year limit (eight if a team re-signs their own free agent) on free agent deals. Second, they put rules in place that govern allocation of salary throughout the deal to eliminate sharp changes in salaries from year-to-year. There can be no more than a 35% change in salary from year-to-year, and in the entirety of the contract there can be no greater than a 50% difference in the highest and lowest single-season salary (Boron, 2013). The changes made to the CBA reflect the NHL's push for more competitive balance by solidifying the hard salary cap by reducing the player's share of HHR and reeling in free agency rules to keep teams from leveraging loopholes in the salary cap.

NFL. The NFL's rules regarding the salary cap and free agency are meticulously detailed in the 2011 CBA agreement between the NFL and NFL Players Association (NFLPA). The NFL uses a salary cap and salary floor system as well as free agency. They use both of these restraints to level the playing field between large market teams that would otherwise have more funds to spend on player salaries in absence of a salary cap and their small market counterparts. Larsen et al. (2006) explains how the two restraints work in tandem:

> The NFL's salary cap in combination with free agency reduces the possibility that one team is able to sign all of the premier talent in the free agent pool, because each team has a fixed amount that it may spend on players' salaries in a given year (p. 382)

In their study, Larsen et al. (2006) find that free agency and salary cap restrictions in the NFL tend to promote competitive balance. For the purposes of this study it is not important to discuss all of the rules related to the NFL salary cap and free agency, but pertinent points are discussed.

In the NFL, the salary cap is a hard cap, so teams are not allowed to exceed the specified amount. According to the NFL's CBA the salary cap is calculated as a percentage of league revenues from the previous season and each player's salary cap share towards the team total is calculated as the pro-rated sum of their base salary and bonuses over the life of the contract (Keefer, 2015). The NFL's salary cap has seen a steady increase in recent years due to effects of larger television contracts on league revenues (Patra, 2015). The NFL announced that the salary cap for the 2015 season will be a new record-high of $143.28 million per team, up from $133 million in 2014 and $123 million in 2013 (Bien, 2015).

The free agency system in the NFL allows for players to offer their services to the highest bidder when their current contract expires. However, there are two types of free agent statuses in the NFL: restricted free agents and unrestricted free agents. The difference between the two and how it impacts their ability to sign with different teams is explained on NFL.com (2013). When a player's contract expires they become a restricted free agent if they have accrued three seasons, and they become an unrestricted free agent if they have accrued four or more seasons. An accrued season is defined as spending six or more games on a team's active/inactive, reserved/injured, or reserve/physically unable to perform lists.

The distinction between a restricted an unrestricted free agent is important. In the case of a restricted free agent, their old club retains a “right of first refusal”. What this means is that a restricted free agent is free to sign an offer sheet with any other team. But when the offer sheet is signed, their old club has the choice to match that offer sheet and retain them on the same terms. If their old team chooses not to match the offer sheet they may receive draft pick compensation from the player’s new team. For an unrestricted free agent there are no offer sheets or draft pick compensation attached to them, they are true free agents.

In addition to the rules regarding restricted and unrestricted free agents, NFL teams also use franchise and transition tags which can reduce a player’s ability to become true free agents. The rules for franchise and transition tags are explained on NFL.com (2013). Under the franchise tag, a team can keep a free-agent-to-be from signing with another team. If a team places an ‘exclusive’ free agent franchise tag on a player they are required to tender him a one-year contract equal to the average of the top five salaries of players at their position or the amount of what a ‘non-exclusive’ franchise player would earn, whichever is greater. If a team decides to place a ‘non-exclusive’ franchise tag on a player that player will receive a one-year contract equal to the average of the top five salaries of players at their position or 120% of their previous years’ salary, whichever is greater. The transition tag allows a team the right of first refusal on any contract an otherwise unrestricted free agent would sign with another team. Teams have the ability to place the franchise tag and transition tag on one player per year. Players can be franchise tagged more than once, but if they are tagged a third time they receive the salary of a tagged quarterback, regardless of their position. This effectively keeps any nonquarterback from being tagged more than twice since the amount would be so high (Florio, 2013).

Television contracts. Television contracts play an important part in determining competitive balance in leagues. As sports leagues have moved into an age with national and local television contracts that match or even surpass gate revenues, the money brought in from these contracts becomes increasingly important. In such a scenario, a television audience may be more valuable than a gate attendance. Mongeon and Winfree (2012) put it into perspective as they posit:

> If, for example, winning impacts the television audience more than game attendance, then the competitive balance of the league will depend more on the television market. Also, league policies such as revenue sharing will have a greater impact on talent investment and competitive balance if the revenue that is shared is the most sensitive to winning. This may help explain why some leagues have focused on sharing television revenue more than stadium revenue (p. 74).

Understanding the totality of television contracts for the leagues is imperative. The totality of the contracts is not just their total value, but also understanding how the money from television contracts is distributed throughout the league. Each of the four leagues has received antitrust exemption from Congress in order to share television broadcast revenue between league clubs (Mitten & Hernandez, 2014). Through this exemption each of the four leagues divides their national television revenue equally among all league clubs. However, there are differences between how national television deals negotiated by the league and local television deals negotiated by individual teams affect competitive balance.

MLB. Major league baseball has both national television deals for the entire league, as well as offering each team the ability to negotiate local television deals for their games that are not nationally televised. This makes sense from a practical standpoint, because each team plays 162 games over the course of a season. While popular teams like the Yankees may find themselves on national television several times during a season, teams with a much smaller national following are less likely to have many, if any, nationally televised games.

MLB has agreements with three networks (ESPN, FOX, and TBS) on its national television deals. All three of the deals were announced in 2012 and each is an 8-year deal running through 2021. MLB's deal with ESPN is worth approximately $700 million per year and gives them broadcasting rights to regular season games on Sunday nights, Mondays and Wednesdays, select postseason games, and regular season tie-breaking games (ESPN, 2012). The deal between MLB and FOX is worth $500 million per year and gives them exclusive broadcasting rights to the World Series and All-Star game, as well as a split of divisional round and championship round playoff games, and regular season Saturday games (Ourand & Fisher, 2012). Additionally, FOX negotiated the right to broadcast some of their regular season games on their all-sports cable channel, FOX Sports 1. The agreement between MLB and TBS is worth over $300 million per year and gives them broadcasting rights to Sunday regular season games, and a split with FOX on the divisional and championship round playoff games (Ourand & Fisher, 2012). All told, the national television deals bring in about $1.55 billion per year to MLB. This is important because like each of the other leagues, national television revenues are shared equally among all teams (Bloom, 2014).

In MLB, similar to what occurs in the NBA and NHL, there are also local or regional broadcast rights. MLB teams in larger markets have a decided advantage when it comes to negotiating local or regional broadcast rights because these deals are not subject to the same revenue sharing rules as the national television deals, which are split evenly. Haupert (2014) relays the importance of local television contracts in MLB:

> Local television contracts are a large and important variable in determining franchise values today. In the last two decades, they have become more important as they have grown in magnitude. What is more, the growth of local television revenue has not been constant across clubs, leading to a great divergence in franchise values (p. 78).

More than just franchise values, it is easy to see how this influences competitive balance. Without a salary cap and without having to share revenues from local television deals, teams with the largest television contracts have more money with which to sign players. This has never been more apparent than now, says Horner (2014) when we see the Dodgers signing a new television deal worth approximately $280 million annually compared to the Brewers deal which pays them approximately $21 million annually. She claims that if MLB does not do something to offset this disparity that only a few teams will strive financially and that competitive balance throughout the league will be negatively affected.

NBA. The NBA, much like MLB, has both national and local television deals to consider. The national broadcasting rights for the NBA are held jointly by ESPN/ABC and TNT. Under the current deal, the NBA receives approximately $575 million per year from these national television deals. Money from the national deal is split evenly among the 30 franchises, so each team receives approximately $19.17 million per year (Mongeon & Winfree, 2011). This deal will expire following the 2015-2016 season and will be replaced with a new deal that was agreed to in October of 2014. Under the new deal, ESPN/ABC and TNT will still retain the rights but the price is increasing by quite a bit. The new deal, starting in the 2016-2017 season is for 9 years and $24 billion. This equates to approximately 2.66 billion per year, or $88 million per team (Sandohir, 2014). As Commissioner Adam Silver said of this large influx of television money, “It will have a profound effect [on the salary cap] and I'm sure the union has already begun studying it just as we've begun studying it” (ESPN, 2014). While this is certainly good news for the league as a whole and will undoubtedly raise the salary cap, the local television contracts play their part in competitive balance as well.

New national television money is good for everyone in the league and it marks more money for the league and more salary for the players due to its even split. In the NBA’s revenue sharing plan (discussed a length below), each team contributes 50% of their total annual revenue, with exceptions, into a revenue sharing pool for the league (Lombardo, 2012). The inherent problem, similar to MLB’s, is that the local television contracts for NBA teams creates a huge disparity between large- and small-market teams. According to Forbes, the Lakers own the richest local television deal, worth approximately $200 million per year, which comes in at $155 million more than the next richest deal owned by the Rockets (Settimi, 2014). Less than 10 years ago, in the 2007-2008 season, the average local television deal for an NBA team was only $12 million per year (Mongeon & Winfree, 2011). With large-market teams beginning to reap the benefits of larger local television deals, they will have the ability to go over the NBA's soft cap and pay luxury taxes that other franchises will not be able to afford. Even with contributing 50% of team revenues to the league-wide revenue sharing pool, local television money is a threat to the already fragile competitive balance model in the NBA.

NHL. Television contracts for the NHL are similar to MLB and the NBA where there is a mix of national and local television contracts. The major difference seen for the NHL is the total value of these contracts. Following the lockout season of 2004-2005, ESPN decided to not renew their television deal with the league. This left the NHL scrambling and they eventually agreed to a 2-year national rights deal with Outdoor Life Network (OLN) that paid the league $65 and $70 million for the broadcast rights (Staudohar, 2005). Recently, the NHL has agreed to two new national television contracts. The first is a 10-year deal that was signed with NBC in 2011 and pays the league $1.9 billion total, or $190 million per year (Condor, 2011). This deal was considered a major upgrade as it tripled the previous deal with OLN and became the largest television deal for the league ever.

The NHL, because of its strong presence in Canada, has the advantage of being able to negotiate national television deals in two countries. In a more recent deal the league agreed to 12-year $4.9 billion pact with Rogers Communications in 2013 for Canadian broadcasting rights to the league. Together with the NBC deal, the NHL will be bringing in just under $600 million per year in national television contracts (Wyshynski, 2013). While this is a big step up for a league that struggled to get a $65 million national deal in the last decade, it still puts the league well behind the other three professional sports leagues in terms of national television money.

Teams in the NHL also have local television contracts to concern themselves with. While each team gets an even share of the new national television deals, the same is not true of local television deals. Since the NHL's revenue sharing contribution (explained below) is based on a percentage of its HRR, teams with higher local television contracts end up contributing more. But a large discrepancy in local television contracts, like the Lakers in the NBA, still allow teams with larger television contracts to spend more money. According to Forbes the Toronto Maple Leafs top the list for NHL local television deals at over $40 million per year, which is twice as much as all but six other teams, with the Dallas Stars and Los Angeles Kings enjoying the eighth largest local television deals at $20 million per year (Settimi, 2013). The Montreal Canadiens are expected to soon top this list as they have agreed to a new 12-year deal with the French-language sports network RDS. Although financial details of the deal were not disclosed it is expected to double their current deal of $31 million per year (Settimi, 2013).

Though reliable figures on other local television deals are sparse, it is known that the San Jose Sharks and the NHL are actively trying to renegotiate their local television deal. The Sharks currently have 14 years left on a deal that pays them just $7 annually (Purdy, 2014). The team and league both view this as unacceptable as the other California teams (Kings and Ducks) have local television deals that pay them north of $20 million per year. The Florida Panthers, based in hockey's sixth-largest city, recently signed a 10-year, $115 million local television deal (Settimi, 2013). Based on the numbers of the Sharks and Panthers, it is reasonable to assume that teams like the Predators, Blues, and Hurricanes in much smaller American markets also have local television contracts that leave them at a competitive disadvantage.

NFL. Compared to the other three leagues the NFL has a rather straightforward set of television deals because the NFL does not have any local contracts outside of preseason games and local radio (Tainsky & McEvoy, 2012). Due to a lack of significant local television contracts NFL teams are on a relatively even playing field because all national television revenue is split evenly between the 32 franchises (Vrooman, 2015). The NFL has television deals with FOX, NBC, CBS and ESPN that total approximately $5 billion per season (Ramachandran & Clark, 2014). Because of the enormity of the national television contracts, NFL teams receive more than double the money than franchises in the other three leagues. Tainsky and McEvoy (2012) describe how television revenue puts the NFL on such solid footing, "The league's television contract provides its member teams a stable financial base before ever selling a single ticket, hot dog, or licensed sweatshirt" (p. 251). This stable financial base created by the league's television contracts allows the NFL to engage in the largest revenue sharing program of the four leagues.

Revenue sharing. The existence of salary restrictions, free agency, and mammoth television contracts requires the need for a system to redistribute the wealth among all members of a league. In order for competitive balance to occur on the field, there must be a balance among all members of a league to work within the structures put in place to promote competitive balance. Zimbalist (2010) said that leagues must, "require a certain degree of revenue balance across teams to be viable" (p. 26) in order to properly utilize a combination of salary caps and floors. Revenue sharing in professional sports is the system that attempts to ensure a certain degree of revenue balance. Slone (2015) puts this in the perspective of competitive balance by stating, "The intuition is that if revenue sharing transfers resources from large clubs to small clubs the extra revenue will be spent on either more or better players, thus improving the playing performance of small clubs" (p. 3).

The literature on the effects of revenue sharing is not in agreement. Some studies have found that revenue sharing does not increase competitive balance (El Hodiri & Quirk, 1971; Fort & Quirk, 1995; Vrooman, 1995) while others have found that revenue sharing does increase competitive balance (Atkinson, Stanley, & Tschirhart, 1988; Késenne, 1996; Marburger, 1997). Despite mixed results in the literature revenue sharing is used by all four leagues as a cornerstone structure for promoting competitive balance.

MLB. The revenue sharing program for MLB underwent significant changes with the last CBA. Under the 2006 CBA, roughly 28% of national revenue and 31% of local revenue was shared equally between the 30 teams (Vrooman, 2009). Under the current CBA 50% of all revenue is shared equally among the teams. Of the revenue that is shared, 50% comes from media revenue and 31% comes from local team revenue (Vrooman, 2015). In addition to this, the clubs in the 15 largest markets will be excluded from receiving revenue sharing starting in 2016 (MLB, "Summary", 2012). MLB has taken clear steps under the new CBA to promote competitive balance by redistributing wealth from the richest teams to the poorest teams in the sport.

NBA. The NBA engages in a revenue sharing program in accordance with the CBA in order to promote competitive balance in the league. At the top of the list for NBA revenue sharing are the league's national television contracts, as described earlier, which are split evenly among each of the 30 franchises. Additionally, Lombardo (2012) lays out the basics of the revenue sharing program beyond the national television deal, which went into effect for the 2013-2014 season, "The core of the plan calls for all teams to contribute an annually fixed percentage, roughly 50 percent, of their total annual revenue, minus certain expenses such as arena operating costs, into a revenue sharing pool" (p. 1). After collecting the revenue from the teams, the NBA then determines what the average payroll salary for the league was and distributes that amount to each team. Following that, they distribute additional money to teams based off of their contribution to the pool relative to the average team payroll.

For example, if the average team payroll for the league was \$60 million and a team contributed \$50 million from their total annual revenue, that team would be given \$60 million and then an additional \$10 million. A team that contributed \$70 million would receive no additional funds as they would be redistributing their money to the teams falling below the average team payroll. Lombardo (2014) describes the system as one where some teams end up being net recipients and some are net payers. For comparison, he provides some highlights of revenue sharing from the 2012-2013 season where the Lakers, Knicks, and Bulls were the top three payers at \$45 million, \$23 million, and $17 million respectively. By comparison the Hornets and Bucks were the top recipients at $21 million and $17 million respectively.

As a provision of the CBA, the NBA's Planning Committee began meeting in October of 2014 to check on how the revenue sharing plan is working and suggest any changes. As one team executive said, "It's a checkpoint to see how the process is working. It's a very complex formula. They want to take a deep dive and make sure that it is working" (Lombardo, 2014). As of the writing of this paper no changes to the structure of league-wide revenue sharing have been amended.

NHL. The NHL's revenue sharing program is a more complicated system than the other three leagues because only teams that have revenues below the league median mark and teams classified as 'small-market' are eligible to receive shared revenues (Peeters, 2015). In their summary of the current CBA the NHL (2013a) describes how revenue sharing is handled. The NHL first commits 6.055% of HRR to revenue sharing. Of the total revenue sharing pool, 50% is raised by the top 10 revenue grossing clubs while the remaining 50% is raised from, "a combination of centrally generated League revenues and a flat tax (35%) on each Club's Playoff gate receipts" (p. 7). Clubs that are in either the top half of league revenues or in a large media market are disqualified from received revenue through the program. For the teams that do qualify, they receive a minimum compensation equal to 70% of the lower limit for payroll ranges in a given year. Additionally, the Revenue Sharing Oversight Committee has the discretion to distribute funds on top of the minimum compensation based on other formulas laid out in the CBA.

Similar to MLB and the NBA, the NHL revenue sharing plan is designed to redistribute wealth from the richest teams to the poorest teams.

NFL. The NFL's revenue sharing model is considered the best of the four leagues. As mentioned earlier, it is in part made possible by the large amount of national television money that the league disperses evenly between the 32 franchises. According to Vrooman (2015) the NFL shares 67% of total revenue. The 67% comes from two sources: national media revenue of which all is shared evenly and makes up 50% of all league revenue, and gate revenue which makes up 20% of all league revenue and is shared at a 34% rate. The additional 20% of total revenue is made up of venue revenue which is not shared. The NFL, at a 67% sharing rate is almost 20% higher than any of its three counterparts, and as Staudohar (2013) says, "The NFL has the most revenue sharing of the major team sports, which is perhaps the chief reason for its success" (p. 4).

Measures of Competitive Balance

Despite the rich literature on the topic there is no agreement on how to measure competitive balance (Uyar & Surdam, 2012). Due to a lack of agreement there exists literature on various dimensions of competitive balance and alternative approaches to measuring it. Zimbalist (2002) put in in perspective when he said, "there are almost as many ways to measure competitive balance as there are to quantify the money supply" (p. 112). In response to the assortment of methods applied to measure competitive balance, Evans (2014) provides a review of these measures.

Evans (2014) makes an important distinction between measures of concentration and measures of dominance. He defines measures of concentration as those that measure, "the extent of closeness between teams in a league in a season" and measures of dominance as those that measure, "the extent to which the same team persists in winning over a number of seasons" (p. 3). This distinction is essential because it defines the focus of a particular study. In a study of concentration, the identity of the teams being measured *does not matter*, but in a study of dominance the identity of the team(s) being measured *does matter*. When a researcher examines competitive balance from a concentration approach they are concerned with measuring the competitive balance of the league as a whole, it is more of a cumulative measure. When a researcher examines competitive balance from a dominance approach they are concerned with seeing how individual teams are performing within the league over the course of the study. In dominance studies the performance of individual teams relative to the rest of the league is the focus. Without such a distinction between concentration and domination it would be difficult to articulate the purpose of a study; as Evans (2014) states, "From a policy perspective it may be important to be able to differentiate between, and assess, the two aspects of competitive balance separately" (p. 3). The review of literature for measures of competitive balance will therefore follow the format and recommendations of Evans (2014) and define the measures, as he does, as either:

1) Measures of concentration
2) Measures of dominance
3) Measures attempting to combine concentration and dominance

Defining measures into one of these three categories is helpful for two reasons. First, it assists in defining measures by considering them in relation to other like measures.

Second, it provides a framework for designating the purposes of studies in order to not confuse one with the other.

Measures of concentration. The following measures do not take team identity into consideration; rather they look at the level of competitive balance for a league within a particular season. These measures are concerned with the closeness of competition within the league.

Range. Range is a very simple statistical measure that can be applied to sports leagues as a measure of concentration. Range is simply the difference in winning percentages between the best team(s) and worst team(s) in the league. Range can fall anywhere from 0-1. A range of 1 would indicate that the best team won every game and the worst team lost every game. A range of 0 would indicate that every team in the league ended with identical records, and would be considered perfect competitive balance. Noll (1991) used range to measure competitive balance in the NBA from 1946-1989 and found that the most competitively balanced period of time existed in the early 1950’s when the league contracted from 17 to 8 members by eliminating the weakest teams. He found there was not a noticeable variance in any of the periods except for the early 1950’s. Quirk and Fort (1992) provide an analysis for each of the four major professional sports leagues in 10-year periods, from 1901-1990. Measuring competitive balance by range is limited in its ability to only look at one variable, and therefore is not often used.

Standard deviation. Standard deviation is described by Evans (2014) as, “A statistical measure of dispersion related to the mean as a measure of central tendency. Applied to sports league, it provides a measure of the concentration of the teams in the league for a competition period (p. 7). Standard deviation is one of the most widely used statistical measures for competitive balance and due to its popularity there are several versions and applications of the measure. There are many variations of the standard deviation approach to measuring competitive balance, presented here are four common applications of the standard deviation measure under which most of the research falls.

Standard deviation of share of maximum possible wins or absolute points. As the title would suggest, this measure of standard deviation can be used to measure a league's competitive balance in either win percentage if the league standings are based only on wins and losses (maximum possible wins), or points if the league standings are based on a point system (absolute points). Each of these measures is based on a range in which the lowest value is 0 and indicates a perfectly balanced league. The upper bound for each measure is based on the number of teams in the league and the amount of games played and the number of teams in each league, so it varies from league to league.

The win percentage method has been used by Scully (1989) to compare MLB's National League and American League from their inceptions through 1987. It has also been used by Fort and Quirk (1992) to compare each of the four major professional sports leagues in 10-year periods, from 1901-1990. Totty and Owens (2011) used standard deviations of win percentage to analyze the effect of salary caps on competitive balance in professional sport, and found evidence that salary caps may be causing a decrease in competitive balance.

An absolute points measure of standard deviation is applicable to leagues that incorporate a points system to determine end of the year league standings. In North

American professional sport this would be applicable to the NHL. As is found in much of the literature, this has been applied to European soccer leagues. Szymanski and Kuypers (1999) used an absolute points standard deviation to measure competitive balance in European football leagues from 1946-1995. Their findings suggested a decrease in competitive balance over the time period. Koning (2000) used this same approach to analyze competitive balance in Netherlands' top football league.

In both of these measures there are two important limitations. First, leagues which play fewer games are more likely to produce higher standard deviations, which can make them appear less competitively balanced than they actually are. Second, since the upper bound in both cases is determined by number of teams in the league and the number of games played by each team, a change in either of these can affect the upper bound limit even if competitive balance is not really affected (Evans, 2014). Because of these two limitations, there are variations of the standard deviation measure in the literature.

Idealized standard deviation. The idealized standard deviation (ISD) attempts to remedy the problems with standard deviation measures by weighting the standard deviation by a factor called the idealized standard deviation (Evans, 2014). The ISD measurement was developed by Noll (1988) and Scully (1989), and is commonly referred to as the Noll-Scully metric. Quirk and Fort (1992) describe the Noll-Scully metric as a comparison of league performance to what the league would look like if it exhibited a maximum degree of competitive balance, or as they refer to it, 'equal playing strengths'. They also provide two assumptions for this metric: there are only two outcomes (win or loss), and a win is equally as probable as a loss. Cain and Haddock (2006) challenged the first assumption on the basis that some sports exhibit three outcomes due to draws or ties. Trandel and Maxcy (2011) challenge the second assumption on the existence of other influences that might provide an advantage, such as a home team effect. As Evans (2014) notes, each of these challenges are not centered at the ISD approach, rather they produce differed ISD's and need to be considered because the quantifications may vary.

The ISD approach has been widely used in competitive balance research. Quirk and Fort (1992) used it to compare each of the four major professional sports leagues in 10-year periods, from 1901-1990. Vrooman (1995) used it to analyze MLB, NHL, and NFL seasons from 1970-1992. Berri, Schmidt, and Brook (2007) used it to measure MLB competitive balance from 1901-2006. Due in large part to the challenges of the ISD assumptions by Cain and Haddock (2006) and Trandel and Maxcy (2011), a normalized standard deviation was proposed.

Normalized standard deviation. The normalized standard deviation proposed by Goosens (2006) has a scale ranging from zero to one, where zero represents a league in perfect competitive balance and one represents a league in perfect competitive imbalance. The normalized standard deviation has two distinct advantages. First, the range of zero to one does not change regardless of the number of teams in the league or the number of games played. Second, the standard deviation is applicable to both open and closed league structures. Because of these two advantages, it provides the ability to compare standard deviations between leagues in which standings are point based and leagues in which standings are win percentage based (Evans, 2014). Normalized standard deviations are very popular in the competitive balance research, with relative standard deviations being the most prevalent.

Relative standard deviation. The relative standard deviation (RSD) is the single statistical measure used most frequently for measuring competitive balance (Owen, 2010; Rockerbie, 2012; Owen & King, 2015) RSD is described by Uyar and Surdam (2012) as, "the most basic and the most often used statistic for evaluating on-field competitive balance". The RSD measurement is the next step in the ISD application, and like ISD was introduced by Noll (1988) and Scully (1989). It was then applied by Quirk and Fort (1992) and Fort and Quirk (1995) in their analyses of the four major professional sports in North America. The RSD is measured by taking the actual standard deviation (ASD) and dividing it by the aforementioned ISD. Its range has a lower bound of 1, which represents a perfectly balanced league. As the RSD moves further above 1, it indicates less competitive balance (Uyar & Surdam, 2012).

Coefficient of variation. The coefficient of variation measurement is a supplement to the standard deviation measurement. It is measured by taking the standard deviation measurement and dividing it by the mean. As Goosens (2006) notes, the typical standard deviation measurement is not valid for comparing leagues that are based on points systems but have different standards for awarding points. Some leagues operate on a 2, 1, 0 point system where others work on a 3, 1, 0 system. In each of these systems the first number is the amount of points awarded to the winning team, the second is the amount of points awarded to a losing team that forces overtime, and the third is the amount of points awarded to a team that loses in regulation. As Evans (2014) notes, this measurement method is valid only if the possibility of a draw is zero.

Excess tail frequencies for win percentages. This measurement is based off of the differences between the actual and idealized distribution of win percentages. The basis for the comparison is a traditional bell curve where two thirds of observations lie within the first standard deviation, 95% of observations lie within the second standard deviation, and 99% of observations lie within the third standard deviation. The measure itself, as described by Evans (2014) is the percentage difference between actual cases that reside within the tails and the amount of cases that would reside in the tails in a perfectly balanced league. This method has been used by Quirk and Fort (1992) to measure competitive balance in each of the four North American sport leagues. It has also been used by Lee and Fort (2005) to test competitive balance in MLB.

Relative entropy. Relative entropy (R) is a measure of competitive balance proposed by Horowitz (1997) in which all teams would have to play the same amount of games against each other. Horowitz describes it as, "the degree of uncertainty about which team might have won a randomly-selected game relative to the maximum uncertainty possible" (p. 376). The measurement has a maximum value of one, which would represent a perfectly balanced league. The minimum value is dependent on the number of teams in the league, as more teams in a league will reduce the range of the measure (Evans, 2014). Horowitz (1997) used relative entropy to test competitive balance in MLB from 1903-1995. As Evans (2014) notes, this measurement is applicable to open leagues but not for comparison between leagues since the minimum value fluctuates due to the number of teams in the league.

Herfindahl-Hirschman Index. Herfindahl-Hirschman Index (HHI) is commonly used to measure concentration in an industry by looking at the market share of a firm in relation to the number of firms in an industry. For competitive balance purposed, ‘firm’ is replaced by ‘team’ and ‘market share of firm’ is replaced with ‘league attainment for a team over a season’ (Evans, 2014). There are three adaptations of the HHI measure because there is a key difference in HHI’s application to sports leagues and industry. The structure of sport leagues, unlike most industries, restricts the range of the measure because it is dependent on the number of teams, which is artificially regulated by the league (Evans, 2014).

Share of all wins in a season. In this measure the limitation of league structure is embraced, and therefore is the most direct application of HHI (Evans, 2014). The range of the measure is determined by the number of teams, but regardless of the range the minimum value is representative of a perfectly balanced league and the maximum value is representative of a perfectly imbalanced league where the best team wins all of their games. Michie and Oughton (2004) use this measure to investigate competitive balance in European football. Depken (1999) uses HHI to test the effect of free agency on competitive balance in MLB from 1920-1996. He finds that free agency has negatively impacted competitive balance in the AL, but had no effect on competitive balance in the NL.

Deviated HHI. This measure is proposed by Depken (1999) with the purpose of recalibrating the metric to have a minimum value of zero, rather than a minimum value that fluctuates based on the number of teams. He did so to account for the variations in the number of teams in MLB due to expansion during the studies' timeframe. In this measure, the maximum value still varies based on the number of teams. In his study Depken finds that overall competitive balance has increased over time, and he uses this as a basis for then looking into issues that may have affected competitiveness such as free agency, the institution of the DH in the AL, and integration.

Normalized HHI. This measure attempts to fill the gap that is left by the previous two measures of HHI. By normalizing the measure, Owen, Ryan, & Weatherspoon (2007) allow for comparison between leagues as the normalized measure does not fluctuate based on the number of teams in the league. Under the normalized HHI the range of the measure is from zero to one, with zero being a perfectly balanced league and one being a perfectly imbalanced league. They tested the measure with Depken's (1999) MLB data and found that the normalized HHI is better equipped to handle comparisons over time and across leagues. Owen et al. (2007) also found that Depken's application of HHI provided an overestimation of the increase in overall parity.

Concentration ratios. The concentration ratio measurement of competitive balance looks at a subgroup of teams in a league and assesses their concentration in relation to a comparable statistic (Evans, 2014). The literature has two examples of the application of concentration ratios, with each focusing on a different comparable statistic.

Attainable concentration ratio. This concentration ratio is so aptly named because it investigates a subgroup of teams compared to the maximum amount of points they could achieve (i.e. if they went undefeated). This measure was proposed by Koning (2000) and has a maximum value of one which indicates a perfectly imbalanced league, and a minimum value that represents a perfectly balanced league, which varies based on the number of teams in the league. In Koning's investigation of the top football league in the Netherlands he determines the values in two subgroups. The first subgroup only considers the top team in the league, while the second subgroup considers the top four teams in the league. He found that competitive balance decreased in the 1960's, increased in the 1970's, and that there is no trend in competitive balance since.

Five club concentration ratio. The five club concentration ratio (C5 ratio) is a measure proposed by Michie and Oughton (2004), and is represented by dividing the total points won by the top five clubs in the league by the total points won by all the clubs. The minimum and maximum values for the measurement depend on the number of teams and the number of points awarded for various outcomes (win, loss, overtime loss, and draw). The minimum value represents a perfectly balanced league, while the maximum value represents a perfectly imbalanced league where each of the top five teams would attain as many points as they could. Michie and Oughton (2004) use the measure to investigate competitive balance in England's top football league from 1947-2004. They find that for the first 40 years of their investigation that the C5 ratio was unchanged, but then saw a rise in the last 15 years of 6.4%. They find this change from 1989-2004 to be significant.

Index of dissimilarity. This measure is one that is often used in other disciplines. Evans (2014) notes that, "it is used in human geography to assess the degree of segregation between ethnic populations" (p. 29). In its application to competitive balance in sport leagues, Mizak and Stair (2004) use it to determine the amount of wins that would need to be reallocated in order to attain perfect competitive balance. The measure has a lower limit of zero, corresponding to a perfectly balanced league, and an upper limit which varies based on the number of teams and corresponds to a perfectly imbalanced league. Mizak and Stair (2004) apply the index of dissimilarity to MLB from 1986-2004 and selected years dating back to 1929 and find that competitive balance has increased over time.

Lorenz curve and Gini coefficients. Lorenz curves and Gini coefficients can be used in tandem to measure competitive balance in sports. Evans (2014), explains it by stating, "The Lorenz curve provides a graphical representation of inequality, which can be measured by the Gini coefficient" (p. 30). The Gini coefficient would approach one in a case of perfect imbalance. The Gini coefficient is adopted by Schmidt (2001) as a new measure at the time because he said, "The use of the Gini measure is quite common in studies where the central question deals with the degree of inequality. In this way, the use would seem to complement the earlier competitive balance literature" (p. 22). In his investigation of MLB from 1901-1998 Schmidt (2001) found that competitive balance began rising with the expansion of the AL in 1962 and the NL in 1963.

The use of Gini coefficients by Schmidt (2001) and Schmidt and Berri (2001) were subsequently questioned by Utt and Fort (2002) for two reasons. "First, they note that the most unequal outcome cannot have one team winning all the games played in the entire league so the maximum value is less than one and consequently the calculated Gini coefficient understates the level of inequality (Evans, 2014, p. 31). Second, "the measure…ignores a host of other complexities, such as unbalanced schedules, league expansion, and interleague play and presents additional challenges (Utt & Fort, 2002, p. 368). The first reason is addressed by Berri et al. (2007) as they conclude it is not really a problem because the scenario that Utt and Fort (2002) point to, while technically possible, would never be observed. The second reason is addressed by Mizak, Stair, & Rossi (2005) as they propose eliminating the denominator from the equation to satisfy all parties.

Surprise index for leagues. The surprise index is a measure proposed and used by Groot and Groot (2003) in their examination of competitive balance for France's top football league from 1945-2002. In this measure they look at the actual results of games in relation to a teams' final standing in the league. Games in which a team either recorded a win or draw against a team and finished higher in the standings is deemed a 'surprise'. To calculate the index the 'surprise' games are weighted to denote the degree of the surprise relative to the final standings. Evans (2014) notes that the final surprise index is then formulated, "by expressing the total number of surprise points relative to the number with a perfectly balanced league" (p. 30). The surprise index has a maximum value of one, which would represent a perfectly balanced league. The index has a minimum value of zero, which would represent a perfectly imbalanced league because it would mean there are no surprising results. In their application, Groot and Groot (2003) find that over the course of their study the competitive balance in France's top football league is slowly decreasing. They also examined the same time period using a normalized standard deviation and normalized three-club concentration ratio and found that each of the measures produces almost the exact same results. The surprise index has the advantage of being able to be used to compare leagues of different sizes.

Measures of dominance. The following measures take team identity into consideration, and look at the level of competitive balance for particular teams or a group of teams over the course of time. These measures are concerned with the same team winning over the course of predetermined number of seasons. Vrooman (1996) claims measures of competitive balance dominance are most important for issues such as free-agency, because in a competitively balanced league the existence of perennial winners and perennial losers becomes the exception rather than the standard.

Descriptive statistics. Descriptive statistics can be used in a number of ways to measure competitive balance dominance. Evans (2014) provides the six most prevalent in the competitive balance literature.

Number of league titles per team. This measure is the simple procedure of counting how many times a team has won a league title. It was first introduced in Rottenberg’s (1956) seminal article as he referred to it as a simple measure to test whether the reserve clause had been successful. Scully (1989) used this measure to analyze MLB dominance from 1901-1987. In Noll’s (1991) study he finds that the Celtics and Lakers had dominated 40 years of league history. Syzmanski and Kuypers (1999) used it to measure football dominance in five European countries from 1946-1998. They find that two clubs have dominated league play in England, Scotland, Spain, and the

Netherlands; while three clubs have dominated league play in Italy. Michie and Oughton's (2004) analysis of the English Premier League find it has been dominated by one team, and that the history of English football has several periods of one-club dominance.

Consecutive title wins. This measure of competitive balance dominance has been used by Szymanski and Kuypers (1999) as they point out that Scotland's top football division has had two separate runs of teams winning nine consecutive league titles. Lenten (2009) takes this measure a step further by calculating a weighted sum for the history of a league. In his calculation a higher number corresponds to team dominance, and therefore less competitive balance. Lenten applies his measure to two leagues, the Australian Football League (AFL) and the National Rugby League in Australia (NRL). Based on this measure he finds the AFL to be the more competitively balanced of the two.

Lifetime achievement of teams. In this measure, a team's average winning percentage over its entire existence is used to compare it against the winning percentages of teams in the same league or teams in other leagues. It is viewed in terms of the number of standard deviations away from the league average winning percentage, which is 0.5. A higher value means that the particular team has achieved a higher level of dominance, relative to the rest of the league. Quirk and Fort (1992) used this measure to identify 'over achievers' and 'under achievers' in MLB, NBA, NFL, and NHL for each ten-year period from 1901-1990.

Number of different title winners. This measure is indicative over dominance over the course of league history. A higher number would indicate more competitive balance because more teams would be winning championships. If the number is lower, fans from teams that have never won a championship have little expectation of doing so from year to year. Szymanski and Kuypers (1999) use this measure and report that from 1946-1998 twice as many teams have won league titles in England than in Scotland, therefore exhibiting better competitive balance. Buzzachi, Szymanski, & Valletti (2003) use this measure to analyze the competitive balance in MLB, NFL, NHL, and soccer in England, Belgium and Italy from 1950-1999. Based on this measure they conclude that the European soccer leagues are significantly less balanced than the North American professional sport leagues.

Number of top teams. This measure of dominance is an extension of the number of title winner's measure, because rather than simply title winners, it aims to include teams that are 'close' to the top of the league (Evans, 2014). Since the definition of 'close' to the top of the league is subjective, there are several versions found in the literature. Borland (1987) defines close to the top as the teams to compete in the finals of the Australian Rules football league divided by the number available berths in the finals. He finds that policies to increase competitive balance should have a positive effect on attendance. Ross and Lucke (1997) look at a 32-year period of MLB from 1961-1992 to get 16 years of observations prior to free agency and 16 years of observations post free agency. They define close to the top as teams that finished within five games of first place and found that since free agency was instituted, more teams have been competitive each year.

In his examination of MLB from 1975-1999 Eckard (2001) looks at five-year periods and considers teams that finished in the top four as close to the top. His findings suggest a drop in competitive balance in the AL and an increase in competitive balance in the NL. Recognizing that the number of teams considered as ‘top’ is subjective, Goosens (2006) chooses to use the top three finishers in each 5-year period she investigates for European soccer leagues. In her version, the number ranges from three to fifteen, with three representing complete imbalance as the same three teams finish there each year. A completely balanced league would correspond with a fifteen as that would mean three different teams enter the top three each season. Based on Goosens’ measure she finds that the Netherlands, Greece, and Belgium exhibit the highest degree of dominance while Denmark shows the most competitive balance.

In a slight variation of the number of top team measure, Curran, Jennings, & Sedgwick (2009) create a top four index in which they are concerned with the identity of the dominant teams in an attempt to explain deviations in competitive balance. Their measure is represented as a proportion of the number of times a team finishes in the top four by the total amount of available places. They apply their measure to the top football league in England from 1948-2008 and find that competitive balance has decreased over time.

Frequency of failure to win a league title. In this measure, the failure of teams to win a league title is examined. Noll (1991) describes this as, “the extent to which any team remains a doormat for a long period of time” (p. 40). He compares the actual frequency of failing to win a championship against the theoretical probability of failing to win a championship for a team in an equally balanced league. Noll uses the measure on the NBA from 1951-1989 and found that from 1951-1969 there were fewer perpetual doormats than from 1970-1989.

Time series association. Time series association measures require that two different periods of data be compared. Evans (2014) notes that this is typically done with consecutive periods. In this measure of dominance, leagues of different sizes and leagues that have different teams during the observed period present problems. Evans (2014) presents four approaches of the time series association measure.

Correlation coefficient. As a basic correlation measure, this has a range of -1 to +1, where -1 represents a perfectly balanced league and +1 represents a perfectly imbalanced league. The measure can be used in two ways. It can be used to measure how a team fares in relation to other teams in the league, or it can be used to measure how a particular team fares in relation to themselves across different seasons. Evans (2014) states this a strong measure of dominance because, “It captures the extent to which all of the teams in the league replicate performance. The dominance of a subset of teams could be masked by the differing performances of the other teams” (p. 38). Daly and Moore (1981) used the measure to rank the eight original MLB franchises in order of dominance from 1955-1973. Butler (1995) used the measure to MLB from 1947-1991 and found that free agency, market sizes, and a compression of baseball talent have all worked to increase competitive balance.

Adjusted churn. The measure of adjusted churn is proposed by Mizak, Neral, & Stair (2007) and is a measure of the absolute difference in end of the year standings for consecutive seasons. The measure has a range of zero to one, with zero representing minimal competitive balance because there is no 'churn', or change in standings from year to year. A value of one represents perfect competitive balance because it would mean maximum possible change. In their examination they look at the MLB's AL East Division from 1995-2007 as well as for each decade of the AL and NL from 1910-2007. Mizak et al. found that competitive balance has declined in both leagues since 1999, and the decline is more severe in the AL.

Autoregressive win percentage. In this time series association measure, Vrooman (1996) constructs an autoregressive model based on team win percentages. His model includes a one year lag and binary variables for small market teams and large market teams. The eight large market teams are denoted for either dominance in media revenues (5) or having a new ballpark (3). Vrooman constructs this model in order to test how free agency affects competitive balance. He applies the method to the 1970-1993 seasons and finds that since the inception of free agency, competitive balance in baseball has significantly increased.

Top 4 recurrence. Top 4 recurrence, as proposed by Curran et al. (2009), measures the probability of a team finishing in the top four, and then repeating that finish the next season. This measure, as a probability, ranges from 0% to 100%, where 100% would indicate top team dominance because the top four finishers would finish in those positions again in back-to-back seasons. A measure of 0% would indicate perfect competitive balance because there would be no repeat top four finishers from one year to another. Curran et al. calculate the measure for each season of England's top football division from 1948-2008 and find that competitive balance has decreased over time.

Herfindahl-Hirschman Index. The HHI, as previously discussed, is often used to measure concentration within an industry, and therefore is widely used as a competitive balance measure of concentration. The measure can be adapted to measure dominance over a number of seasons, and it has been used three different ways in the literature.

Titles won. The dominance measure of titles won is a function of two items. In a given period of time it examines how many teams have won titles as well as the relative title shares of those teams. The minimum value for a given time period, indicating perfect competitive balance, would be 1/N, where N = the total number of teams in the given league. The maximum value, indicating perfect competitive imbalance, would be one, and would mean the same team won the title each year of observed period. This measure is used by Humphreys (2002) in his investigation of MLB decades from the 1900s to the 1990s. He calls the measure the Competitive Balance Ratio (CBR) and concludes that the measure does a better job than other, more popular methods, at distinguishing differences in competitive balance and is a useful competitive balance measure.

Number of top or bottom positions. This measure, proposed by Eckard (2001), is aimed at providing a more comprehensive measure of dominance by including more teams than simply title winners. In his study, Eckard (2001) calculates the measure for both the top four and bottom four teams in MLB from 1975-1999 because, as he states, "More concentration means less balance, as the same teams repeat their positions as contenders or tail-enders" (p. 215). His findings suggested recent drops in competitive balance in the AL and recent increases in competitive balance in the NL.

Virtual league appearances. In an interesting twist on much of the competitive balance literature, Eckard (1998) proposes this measure in order to measure the dominance of teams that, while competing in the same sport, are not in direct competition with each other. He applies the measure to college football because due to its system of highly regional competition and national ranking system. Eckard claims that the existence of a national rankings list creates a 'virtual league'. If a team finds itself on this list that is a measure of dominance, recurring inclusion on the list from year to year would indicate more dominance. He uses this to measure competitive balance in college football by calculating measures for both top 10 and top 20 lists from 1924-1984. Eckard does so to test whether competitive balance has changed since NCAA regulation changes in 1952. He finds that, based on his measure, competitive balance in NCAA football has declined since regulation changes in 1952.

Lorenz curve and Gini coefficients. Similar to HHI, Lorenz curves and Gini coefficients have been used to measure competitive balance in terms of both concentration and domination. For measuring competitive balance domination, the Lorenz curve is positioned above the line of perfect equality to show teams in order of decreasing success, while the Gini coefficient is calculated by ranking the teams in increasing order of success (Evans, 2014). By doing so a measure between -1 and 1 is calculated, with -1 corresponding with a perfectly balanced league and 1 corresponding with a perfectly imbalanced league. Quirk and Fort (1992) graph the measure for the big four American sports leagues from 1901-1990. Szymanski and Kuypers (1999) use this measure to rank the top football leagues in Europe based on dominance and they found that from least competitively balanced to most competitively balanced, the ranking was the Netherlands, Spain, Italy, Scotland, and England. Goosens (2006) applies the measure from 1963-2005 for eleven national football leagues in Europe.

Measures combining concentration and dominance. Measures that combine both concentration and dominance have a marked advantage of, “providing a single statistic as a measure for both aspects” (Evans, 2014, p. 3). But Evans (2014) also points out that they are not always the best statistic based on what is being measured because, “they do not fully replace the independent measures of concentration and dominance” (p. 3). The literature provides four approaches for combining measures of concentration and dominance.

Distribution of ‘lifetime’ win percentages in a league. As the name would indicate, this is a measure that looks at competitive balance for a league over its entirety. A team is measured by its Lifetime Achievement statistic (discussed earlier under descriptive statistics for measures of dominance), while the league is measured by the concentration of teams whose Lifetime Achievement statistics fall within an idealized standard deviation of the league population mean for winning percentage of 0.5. A larger concentration, or proportion, of teams close to the league mean would indicate a more competitively balanced league. A larger concentration of teams away from the league mean would indicate more dominant teams, and therefore a less competitively balanced league. Quirk and Fort (1992) use this measure with three standard deviation limits to examine each of the big four American sport leagues from 1901-1990.

ANOVA-type measures. As Evans (2014) notes, these measures are contingent on the same teams competing with each other over a specified amount of time. Therefore, these measures are applicable to closed leagues that did not experience expansion over the course of the study. Open leagues, where either the identity or total amount of teams is subject to change, would not be applicable. Two ANOVA-type measures are proposed in the literature.

Variance decomposition. In this application, Eckard (1998) develops an ANOVA-type measure to test his hypothesis that regulation enacted in 1952 by the NCAA has reduced competitive balance. As he states it, “Less balance within a conference means that schools with good, middling or poor records tend to repeat them year after year” (p. 214). Eckard’s measure combines both concentration and domination because he, “assesses competitive balance by separating the pooled variance of win percentages into the variation in a team’s performance between seasons and the variation between teams” (Evans, 2014, p. 46). Eckard effectively develops a measure to view competitive balance on both the league and team level. Eckard’s findings support his hypothesis that competitive balance has decreased in college football since regulation changes in 1952. In fact, Eckard finds that this occurred in all five conferences and that the mean difference between pre- and post-enforcement periods is statistically significant.

The variance decomposition measure has undergone scrutiny since Eckard (1998) proposed it. Eckard (2001) provides an update on the measure in order to point that the measure is restricted to leagues in which the same teams participate throughout the specified period and play the same amount of games. This measure is later criticized by Humphreys (2002) as he concludes that the condition in which this measure holds up would be, “a situation that rarely occurs in actual sports leagues” (p. 135). Eckard (2003) responds to this criticism and shows it to be wrong, and Humphreys (2003) accepts that his own original criticism was in fact incorrect.

Competitive balance ratio. As discussed under HHI measures of dominance, the measure of competitive balance ratio (CBR) was proposed by Humphreys (2002). He also proposed it as a combined measure of concentration and dominance. His argument for the validity of the measure is that CBR makes it easier to compare different time periods. In order to test this, Humphreys compares CBR for each decade of MLB from the 1900s to the 1990s to standard deviation measures (concentration) and HHI measures (dominance). As a result of this comparison Humphreys states, “The CBR uncovers important distinctions between several periods that have similar [standard deviations] and HHIs” (p. 139). He points to particular decades where the CBR shows stark differences between the periods and the other measures produce similar values.

Specifically, Humphreys (2002) points to the National League in the 1920’s where the same subset of teams consistently occupied the upper and lower portion of the standings and he concludes that, “The CBR captures this relative stratification in standings but the other two measures do not” (p. 142). Humphreys also references the 1990s and states that the CBR shows changes in competitive balance in the second half of the decade that measures like standard deviations do not. CBR, similar to variance decomposition, is not without its criticisms. Eckard (2003) shows that CBR is actually just an adapted measure of his variance decomposition.

Mobility gain function. Mobility gain function is a competitive balance measure proposed by Lenten (2009). Evans (2014) explains that in this measure potential changes in team performance are categorized. The categories of potential change are relative to the previous season performance of the team as well as to the league's average performance. The categories are then weighted in order to produce a model to determine the impact it would have on competitive balance for the league. The final measure of concentration is the averaged product of the total 'gain' of all of the teams in the league. If the gain function moves towards the league average performance from one season to another, it represents an increase in competitive balance. Evans details the three possible scenarios related to mobility gain function as a measure of dominance for individual teams. In the first scenario, a particular team has not 'gained' more competitive balance relative to the previous season. In the second scenario, an individual team got closer to competitive balance but still resided either above or below the league average. In the third scenario, a team moved from one side of the league average to either league average or the other side.

The mobility gain function is applied to the Australian Football League (AFL) from 1898-2006 and National Rugby League in Australia (NRL) from 1911-2006 (Lenten, 2009). As part of his application he compares it to six other popular measures of competitive balance. Lenten's results point to three specific findings. First, three of the six measures that he compares his mobility gain function to are strongly correlated.

Second, of the six measures that he compares he concludes that the standard deviation is the best. Finally, Lenten finds that his mobility gain function indicates competitive balance effects that are quite different from the other measures. It is pointed out that this is to be expected since the other measures he compared mobility gain function to do not include a combination of concentration and dominance. Because of the nuances picked up by mobility gain function in relation to the other, more common measures, Lenten suggests this measure be at least implemented in tandem with other measures of competitive balance for future studies.

Markov models. Markov models are common in other disciplines of study. As Evans (2014) notes, they are frequently used to model disease progression. The underlying tenet of a Markov model is that the state of what is being examined in one period affects the outcome in the next period. Applied to competitive balance in sports, a Markov model would hold that a team's performance in one season would be dependent on that same team's performance in the previous season. The literature contains two applications of Markov models to competitive balance in sports.

Transitional probability tests. In this measure of competitive balance, teams are designated as either winners, contenders, or losers based on their ranking to end each season. The probability that a team transitions from one designation to another is the transitional probability. In each season that is observed, teams have the opportunity to either remain in the same designation, or move to one of the other two designations. If the league were perfectly balanced, one would find that the designation as either winner, contender, or loser for one season was not dependent on the previous season's designation. This measure is a combination of concentration and dominance because it can be viewed over one season or a number of seasons for an entire league (concentration) or for each team in the league (dominance). As Evans (2014) sums this up by stating, “Competitive balance within a season can be assessed by comparing actual values with ‘balanced’ transitional probabilities and tested statistically for difference. The actual values can also be used to assess competitive balance between periods” (p. 52).

Transitional probability tests as a measure of competitive balance was implemented by Koop (2004). His application of the model examines MLB from 1901-2000 and he finds that with the exception of the Yankees, MLB enjoys a very high level of competitive balance. Koop points out that that Yankees profile as an exceptional team that is substantially different from other teams. Even with the Yankees skewing the numbers there is no evidence that competitive balance has worsened over time. The measure was also applied by Hadley, Ciecka, & Krautmann (2005) in their investigation of MLB from 1982-2003. They chose this time period in order to examine periods before and after the 1994 strike-shortened season. This study found that competitive balance has decreased in the period following the 1994 player’s strike.

Gini-type measure. In this Markov model, Buzzachi et al. (2003) put forward a single measure that is based on the number of teams that finish a season, “At, or near the top of a league compared to the theoretical number that would be expected if all teams had equal probability of winning each of their games” (Evans, 2014, p. 53). Buzzachi et al.’s measure has a lower limit of zero, which indicates a perfectly balanced league. The upper limit depends on the number of seasons that are being observed, and indicates a perfectly imbalanced league. This measure can be used to analyze both open and closed leagues. As such, Buzzachi et al. utilized it to compare the open soccer leagues of Europe to the closed professional leagues of North America. They find that when using the Gini-type measure, the closed leagues of North America exhibit significantly more competitive balance than their open counterparts in Europe.

Generalizability Theory

Generalizability Theory (G-Theory) is a method often used in education and psychology research and first introduced to these fields by Cronbach, Rajaratnam, and Gleser (1963). Generalizability Theory provides researchers with the advantages of being able to simultaneously examine multiple sources of variance and examine both the generalizability and consistency of measurement (Cronbach, Gleser, Nanada, & Rajaratnam, 1972). On this matter, Briesch et al. (2014) point out with Generalizability Theory, "a single construct of interest can be measured under many different conditions, including using different measures, employing different raters, or measuring the construct at different points in time" (p. 15). This theory applied to the context of competitive balance in sports would state that we are not necessarily interested in the result of a particular game on any particular day under a particular set of circumstances, but rather we are interested in obtaining an estimate of the teams (or league's) general level of competitiveness. In this way, Generalizability Theory can provide an estimate of competitive balance with the advantages described by Cronbach, Gleser, Nanada, and Rajaratnma.

When using Generalizability Theory two studies are carried out; a generalizability study (G-Study) and a decision study (D-study). As Briesch et al. (2014) say, “A G study focuses on estimation of the degree of measurement variance attributable to different sources of variance, or facets” (p. 15). This is based on the use of a repeated measures ANOVA and ultimately provide a percentage breakdown of the total variance associated represented by each facet in the model. The percentages provided are the most important information provided by the G-Study (Morrow, 1989). This percentages information is important because it “can help the researcher determine which sources of variance contribute to measurement error. Decisions are then made regarding the manipulation or control of error sources” (Morrow, 1989, p. 76). The following study will have three sources of variance: the individual teams (T), each game (G), and the interaction between the teams and games (T x G). When the G-study is completed the researcher will have an estimate of how much each variance source contributed to the measurement error, and could subsequently manipulate the test.

The next study carried out in utilizing Generalizability Theory is the D-Study. The results of a D-Study are two coefficients; the generalizability coefficient (G or ρ^2) and the dependability coefficient (Φ). Both of these coefficients are useful depending on the purpose of the research. Generalizability coefficients are used to make relative decisions and dependability coefficients are used to make absolute decisions; and both may be of interest in the same study (Briesch, et al., 2014). The following study will utilize both coefficients as the dependability coefficient will be used to make the absolute decision of whether or not a league exhibits competitive balance. The generalizability coefficient will be used more extensively to compare different years within the same league, as well as different years between different leagues in order to make relative decisions about which leagues and years exhibited higher levels of competitive balance.

Each of the coefficients produced is then used to determine a desirable level of generalizability or dependability; which is referred to as the minimum level of reliability. Both coefficients use the same minimum level of reliability in order to make their respective decisions. One of the complications in the Generalizability Theory literature, though, is that there is no universal standard for a minimum level of reliability – rather a satisfactory level is determined by the nature of the study (Nunnally & Bernstein, 1994). The satisfactory level is mostly determined by judging the relative importance of the decisions of the study. As Salvia, Ysselydke, & Bolt (2010) point out, there does seem to be agreement that as the stakes of the decision being made increase, so should the level of coefficient. The authors point out that a minimum level of reliability as low as 0.7 may be used if a researcher is in the early stages of trying to determine the merit of conducting a study (Salvia, Ysselydke, & Bolt, 2010). Conversely, a minimum level of reliability as high as 0.9 may be used if high-stakes decisions are to be made, such as whether or not to categorize children as special needs.

Due to the lack of consensus on a satisfactory level, studies using Generalizability Theory have considered minimum levels of acceptable reliability between .70 and .90 (Hopkins, 1997; Nunnally & Bernstein, 1994; Salvia et al., 2010). Within the fields of education and psychology .80 has often been used as an acceptable minimum level of dependability (Briesch et al., 2014). Using .80 as an acceptable minimum level of dependability has also been confirmed in fields such as physical activity (Ishikawa et al., 2013; Kang et al., 2014) and nursing (O'Brien, 2014; Sijtsma & van der Ark, 2015). Because a minimum level of dependability of .80 is widely used in educational, psychology, and physical activity research, the present study will use it as well. The decisions made by the present study on the topic of competitive balance do not represent decisions that require a more stringent level of dependability than those in the fields of education, psychology, or physical activity.

In determining whether Generalizability Theory is appropriate for a study, Briesch et al. (2014) provide a 7-step process for deciding whether to use it, and if so, how to use it.

Step 1: Confirm that GT is appropriate. Is the goal of the study to investigate the generalizability or dependability of measurement? The answer to this question should be yes. As Briesch et al. (2014) says, "Use of GT may be appropriate in those situations in which there is interest in determining the generalizability of rankings across particular instances of a facet or facets or the dependability of measurement in making criterion-related inferences" (p. 20).

Step 2: Identify relevant facets. Although simple, this step is important to the process, as a researcher must determine all facets that are to be included in a model. Facets in a Generalizability Theory study are any sources of variance in the study. A faced in a G-Theory study is the same as a factor in an ANOVA study (Briesch et al., 2014). Key in this process is clearly articulating the reasoning for facet choices and what each observation constitutes (Brennan, 1992). Observations will vary from one study to another because they are based on the purpose of each study. In a study looking at competitive balance in sports, the researcher would be interested in the game results of the teams in the league. Therefore, they would be looking at a study of subjects (the individual teams) and days (each game played), where teams and games are each facets of the study.

Step 3: Determine whether multiple instances of each identified facet can be sampled. In order to use Generalizability Theory, the researcher should be using multiple instances of the selected facets; if not, Generalizability Theory is not appropriate. In a study of competitive balance a researcher would need to analyze multiple games for each individual team in order for Generalizability Theory to be appropriate.

Step 4: Describe the measurement procedure and associated G-study design. Under this step, three pieces of information must be considered. The first is whether the researcher will use a crossed or nested design. According to Briesch et al. (2014), "A fully crossed design is one in which all objects of measurement are rated under all specified conditions of measurement" (p. 23) and is most desirable because it allows for interpretation of all facets and interactions. In a nested design, there are different levels of one facet for each level of another facet. Practically, a nested design occurs when all objects cannot be rated under the same specified conditions. This often occurs when a particular rater may only be able to assess a subset of items. A study that looks at competitive balance over the course of a season would exhibit a fully crossed design if every game for each individual team were measured.

The second piece of information to consider is a specification of facet classification. Here, a researcher will determine whether a facet is random or fixed. If the facet would be expected to vary when replicating a G-Study, that facet will be considered random. The facet would be considered fixed if the researcher will not be making generalizations beyond what is immediately sampled (Briesch et al., 2014). In most studies facets are treated as random so that results can be generalized. However, if a researcher is particularly interested in examining the dependability of an established measure, then treating a facet as fixed would be appropriate. A competitive balance study looking at subjects and days would be classified as a random facet due to the variability of win/loss results based on the particular day.

The third piece of information to consider is determination of an adequate sample size. While authors have discussed optimal sample sizes for G-studies (Shoukri, Asyali, & Donner, 2004), their recommendations are only applicable to a single-facet design. While there is no consensus, it is agreed on that having too few data points can result in uneven or negative components of variance (Shavelson, Webb, & Rowley, 1989; Smith, 1981). Although published studies have included fewer, the recommendation put forward by Conger, Conger, Wallander, Ward, and Dygdon (1983) is that G-studies include a minimum of 20 persons and 2 conditions per facet, or 40 data points for a one-facet design. Any study comprising an entire season of a North American professional sport league (MLB, NBA, NFL, NHL) would meet the recommendation of at least 40 data points, as an NFL season would represent the smallest amount of data points at 512 (32 teams x 16 games per team).

These first four steps prepare a researcher to use generalizability theory. If the study meets the requirements of these four steps the researcher can then move on to conducting the study.

Step 5: Select an appropriate software package. This is important to decide before data collection, because different packages have different capabilities, the packages are not interchangeable, and some packages can only handle specific types of data. The GENOVA suite of programs (Crick & Brennan, 1983) include GENOVA, urGENOVA, and mGENOVA. GENOVA can be used in generalizability studies for, "complete, balanced univariate designs and offer the option of entering raw data or mean squares in order to calculate variance components" (Briesch et al., 2014, p. 25). urGENOVA handles balanced and unbalanced designs, and some designs with missing data. mGENOVA was designed specifically for multivariate generalizability examinations. Important to note is that of all software packages available, a researcher can only get D-study results from GENOVA and mGENOVA.

In addition to the GENOVA suite of packages, Generalizability Theory analyses can be done using popular statistical packages like SPSS and SAS. The two largest drawbacks to both of these programs are that they do not support missing data and that they lack the ability to produce D-study results. The ability to produce D-study results is a vital consideration for any researcher interested in the results of a D-study, and therefore the GENOVA suite of packages is often used.

Step 6: Conduct the G-Study. As discussed above, the G-study will provide a breakdown of variance components for specified facets and interactions. This should be used to determine potential sources of error and make changes accordingly. Additionally, the researcher should take note of any problems with estimation such as negative or uneven variance components. In a competitive balance study which uses subjects and days as facets, a G-Study would provide information on three different variance sources: subjects (the individual teams), days (each game played) and the interaction of subjects and days.

Step 7: Conduct the D-Study. The results of the D-study, as earlier discussed, are used to make relative decisions (generalizability coefficient) or absolute decisions (dependability coefficients) about the data. Each coefficient is judged on its acceptable level of dependability, which will typically fall between .70 and .90 depending on the stakes of the decision being made. In a study where the researcher is utilizing both coefficients it is advisable to use the same acceptable level of dependability for each. In a competitive balance study, if the dependability coefficient exceeded the minimum level of reliability, that season would be determined to not be competitively balanced. If it fell below the minimum level, that season would be determined to be in competitive balance. The generalizability coefficient would be used to make relative comparisons between different seasons and different leagues. For example, if one season had a generalizability coefficient of .50, that season would be more competitively balanced than one that exhibited a generalizability coefficient of .60.

Interpreting results. After the seven steps have been completed the researcher is left with results from both the G-study and D-study, which require interpretation. Interpreting the results of both the G- and D-studies will be largely dependent on the purpose of the study. In an educational study the researcher may be concerned with interpreting whether or not students are engaged during a class discussion. In a psychology study the researcher may be concerned with obtaining reliable estimates of a patients' self-esteem. In a physical activity study the researcher may be concerned with how many days it takes for a step-counter to produce reliable estimates of a participants' physical activity. In a study of competitive balance in sports, a researcher is concerned with whether or not a league reaches a level of reliable estimates in regards to the outcome of its games. The researcher would also be concerned with the ability to compare these results across different seasons and for different leagues. In the event that a league is reaching an acceptable level of dependability, competitive balance in the league could be compromised.

CHAPTER III

METHODOLOGY

Generalizability Theory

In order to test competitive balance in professional sports in North America, this study will use Generalizability Theory as the method for testing. To do this, the study must be pass the 7-step process presented by Briesch et al. (2014).

Step 1. The goal of the study should be to investigate the generalizability of rankings across specific instances of a facet or the dependability of measurement in order to make criterion-related inferences. The goal of the present study is to both test the generalizability of rankings of professional sport leagues and the dependability of the measure of competitive balance, so Generalizability Theory is appropriate.

Step 2. The researcher must identify all relevant facets in order to clearly articulate the reasoning for facet choices and what each observation constitutes. This is a one-facet in study looking at subjects and days (S x D). Each team in a league represents the subject, and each game they play over the course of a season represents the day. Therefore, the one-facet design will be identified as teams and games (T x G). Teams and games are each facets of the study, similar to factors in an ANOVA study. They each represent sources of variance in the study. Additionally, there is the interaction between teams and games as a potential source of variance. So, in this one-facet study, there are three different sources of potential variance.

Step 3. The researcher must determine if multiple instances of each identified facet can be sampled. In order to use Generalizability Theory, the researcher should be using multiple instances of the selected facets; if not, Generalizability Theory is not appropriate. The present study will be using multiple instances of the facet since each game played by each individual team over the course of observed seasons will be sampled. Thus, Generalizability Theory is appropriate.

Step 4. The researcher must describe the measurement procedure and related design of the study. There are three important parts to this step: First, the design must be identified as nested or crossed. Second, facets must be identified as random or fixed. Third, it must be determined whether or not the sample size is adequate.

The present study will be a fully crossed design because all objects of measurement fall under the same conditions of measurement. This is because every game for each individual team will be sampled, so the conditions of measurement are the same. The facet will be random as it would be expected to vary if the study were replicated since each team's win/loss result is expected to vary by the game they are playing. According to the 40 data points recommendation (Conger et al., 1983) this study meets the requirements for adequate sample size. The study will sample each of the four professional sport leagues in North America and will look at game results for each of them. Therefore, for one season of observations the NFL would have 512 data points (32 teams x 16 games each), the NBA and NHL would each have 2460 data points (30 teams x 82 games each), and MLB would have 4860 data points (30 teams x 162 games each).

Step 5. The researcher must select an appropriate software package for analyzing the data. Taking into account the different capabilities of statistical software, the present study will use GENOVA for data analysis. The two main reasons for choosing the GENOVA software is its ability to have raw data entered and its ability to produce D-Study results.

Steps 6 and 7. The researcher must conduct both the G-Study and D-Study. These will be detailed in the following Data Analysis section and reported in the Results section.

Based on the assessment of the aforementioned steps, generalizability theory is appropriate for the study. It is appropriate because the study requires results that measure reliability, which is precisely what Generalizability Theory does. In fact, Morrow (1989) argues, “Generalizability theory is perhaps the most appropriate methodology for estimating reliability available” (p. 75). This presents us with an interesting fit for a competitive balance study. Essentially, reliability is the antithesis of uncertainty; and uncertainty of outcomes is the foundation of competitive balance. Owen and King (2015) support this position by stating, “An appropriate degree of competitive balance, how evenly teams are matched, is regarded as central to this endeavor, as this affects the degree of uncertainty over the outcomes of individual matches and overall championships” (p. 731).

With this in mind, the results of the D-study would be interpreted in an opposite manner as they would typically be for a study using Generalizability Theory. In a typical Generalizability Theory study, the researcher would want to reach the minimum level of reliability in order to obtain reliable estimates of the measure. To use the example from earlier, if a study testing a student's level of science ability reaches the minimum level of reliability it means that the test should produce similar results even if components such as day, subject, or rater change. However, the present study will be interpreted in an opposite manner because lower coefficients, both generalizability and dependability, would be associated with higher levels of competitive balance. Higher coefficients would be associated with lower levels of competitive balance. If the coefficients were to cross the threshold of the minimum acceptable level of dependability, then the data being analyzed would be providing reliable estimates. In the case of looking at an entire season's worth of games for a league, if this were to happen it would represent a situation where reliable estimates of the final rankings (standings) could be made. This situation would certainly be of concern to any league that actively promotes competitive balance.

Data Collection

Data collection for the study consists of inputting game-by-game win/loss results for each of the four North American professional sports leagues (MLB, NBA, NFL, NHL) for a 10-year period dating from 2005 to 2014. This ten-year period was chosen for three reasons. First, in order to make comparisons between different seasons and different leagues more than one season's worth of data is necessary. Second, the most recent data provides us with the best representation of where each league is currently positioned in regards to competitive balance. Finally, the 2005 season was chosen as a logical starting point because of the structural rule changes which the NHL adopted prior to the 2005-2006 season. As documented, the 2004-2005 NHL season was completely wiped out due to a lockout. Prior to 2005 NHL games could end in ties, so instead of a typical win/loss record such as 42-40, NHL standings counted wins/losses/ties/overtime losses, so standings looked like 34-30-12-6. This would make it impossible to compare with the other leagues using G-Theory. When the NHL resumed play in 2005 they instituted changes to team records; specifically they eliminated the ability for games to end in ties by adopting sudden death shootouts. This makes for much easier comparison between the other leagues, which also do not count ties (except for very rare occasions in the NFL).

Game-by-game win/loss data for each team for each of the ten seasons was recorded into a spreadsheet with a "1" representing a win, a "0" representing a loss. There were four instances of NFL regular-season ties during the observed period. These four ties were treated as missing data in the dataset. In line with the data collection procedure of Mills and Fort (2014), all data was collected from the Sports reference league-specific coverage websites (Baseball-Reference, 2005-2014; Basketball-Reference, 2005-2014; Hockey-Reference, 2005-2014; Pro-Football-Reference, 2005-2014).

Data Analysis

Analysis of the data was done using the GENOVA for PC application. The first analysis investigated all 30 teams in MLB, all 30 teams in the NBA, all 32 teams in the NFL, and all 30 teams in the NHL together for each of the ten years. In an attempt to meet the assumption of independency of measures that Generalizability Theory is subject to, as it is an extension of intraclass correlation (ICC) and works within an ANOVA framework (Briesch, et al., 2014), the same analysis was done again using a randomized subset of 80% of the data for each league (i.e., 24 teams for MLB, the NBA, and the NHL and 26 NFL teams). Because there is one winner and one loser in each contest, the measures are not independent of each other as they influence the outcome of others. By randomly selecting 80% of the data we increase the level of independence of measures by eliminating subjects that would otherwise impact the outcome of the remaining teams.

Cronbach (1989) explained a similar approach when he analyzed the Stanford-Binet Intelligence Test. The test consists of 15 subtests which produce an index of four dimensions of intelligence. Rather than use an aggregate of each of the subtests to produce their intended index, Cronbach evaluated each singular subtest as a random sample of their particular index in order to determine the degree to which it measured its intended trait. By treating each random 80% sample of data for competitive balance in the same manner we will test the degree to which it measures the overall index of competitive balance. This will be done by analyzing the 80% randomized set of data in a paired t-test with the full dataset for including all teams in the league. If the paired t-tests for each league return significant results it would indicate that the samples are different from each other. However, if the paired t-tests for each league return non-significant results, it would indicate that the 80% randomized sample is similar to the full dataset and provides an estimate of the overall index of competitive balance. Both the full dataset results and the 80% randomized sample dataset will be reported, but only the 80% randomized dataset will be used to draw conclusions as it is intended to improve the independence of measures.

Another step taken by the researcher is to analyze each of the leagues based on their conference structures. Each of the leagues does not have the same divisional structure, but each does have two conferences, so the next step in the analysis looked at each conference in order to compare competitive balance between them. This practice has been very common for studies looking at competitive balance in MLB (see Berri, Schmidt, & Brook, 2007; Mills & Fort, 2014; Quirk & Fort, 1992) but not nearly as common in the other three sports despite the similarities.

MLB is divided into the National League (NL) and American League (AL). Each conference consists of 15 teams with teams in one conference playing teams from the same conference 88% of the time. The NBA is divided into the Eastern Conference and the Western Conference. Each conference consists of 15 teams and teams in one conference play teams from the same conference 63% of the time. The NFL is divided into the American Football Conference (AFC) and the National Football Conference (NFC). Each conference consists of 16 teams and teams in one conference play teams from the same conference 75% of the time during the regular season. The NHL is divided into the Eastern Conference and the Western Conference. Despite changes in the last year, for the duration of this study each conference consisted of 15 teams and teams in one conference played teams from the same conference 80% of the time.

Additionally, the analysis was set to check coefficients at four different intervals throughout the season. This was done in order to present a more accurate picture of reliability estimates during the season. For MLB each season can be seen at games 40, 80, 120, and 162. For the NBA and NHL each season can be seen at games 20, 40, 60, and 82. For the NFL each season can be seen at games 4, 8, 12, and 16. Moreover, in the event that the reliability coefficient reached the 0.8 threshold at a point during the season, the intervals would provide more information as to where it occurred.

For the purposes of the present study, the results of the D-study are most relevant. The reliability coefficients (generalizability and dependability) produced by the D-study disclose the information by which competitive balance can be examined. The dependability coefficient provides the information to make the absolute decision of whether or not a league is competitively balanced. When the reliability coefficients reach or exceed 0.8, the league is not competitively balanced. In terms of the present study, if the coefficient level of 0.8 is reached, then the number of games at which this occurs would represent the point in the season where one could make reliable estimates about the final standings.

The generalizability coefficient provides the information to make relative decisions and inferences. When comparing the generalizability coefficients between different seasons in the same league the researcher can make inferences about which seasons exhibited higher or lower levels of competitive balance. When comparing the generalizability coefficients between different leagues the researcher can make inferences about which league exhibits higher or lower levels of competitive balance. In order to maximize competitive balance, whether one is looking at the measure in absolute or relative terms, leagues would want to see the coefficients as low as possible. In order to provide more descriptive information regarding the reliability coefficients, the researcher will report ranges, means, standard deviations, 95% Confidence Intervals (CI), medians, and interquartile ranges (IQR) for the full-season coefficients for each league and conference. This information will help to interpret the results as they will provide more context to the coefficients.

CHAPTER IV

RESULTS

Results are displayed in five sections; one section for each of the four professional sport leagues (MLB, NBA, NFL, and NHL), and a fifth section offering a comparison for each of the four leagues. In each of the first four sections the results for the G-studies and D-studies will be shown as well as for each conference in the league. The fifth section will display a comparison of only the D-study results for each of the four professional sport leagues and will focus on the league as a whole rather than include conference results for each league. Although the results of the D-study are most relevant to the present study, the results of the G-study can aid in determining relative contributions of variance components.

Each year in the study represents a one-facet, crossed design, which produces both G-Study and D-Study results. Because of this design it is possible to compare the results between different years within the same league as well as between different leagues. The information in the tables will be used to make absolute decisions on whether or not a particular season was competitively balanced. The information will also be used to make relative comparisons and decisions on the levels of competitive balance exhibited in each season of each league. In each of the observations, the dependability coefficient (for absolute decisions) and the generalizability coefficient (for relative decisions) were the same. Each of the tables displays the generalizability coefficients (G-Coefficients), but the coefficients can also be used for absolute decisions since the two were the same in each observation.

MLB

The G-Study results for MLB from 2005-2014 show the majority of variance comes from the interaction component of teams and games (T x G). Each season displays a 0% variance from the Game (G) component. The percent variance of the Team (T) component ranges from 0.68% - 1.65% with a mean of 1.19%. The percent variance of the interaction component (T x G) ranges from 98.35% - 99.32% with a mean of 98.81%.

Table 1 displays the results of the D-Study for the 2005-2014 seasons and Table 2 displays the results of the 80% randomized data D-Study for the 2005-2014 seasons.

Table 1

D-Study, G-Coefficients, MLB 2005-2014

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
40	0.320	0.271	0.214	0.333	0.354	0.330	0.354	0.383	0.401	0.239
80	0.484	0.426	0.352	0.500	0.523	0.496	0.523	0.554	0.573	0.385
120	0.585	0.527	0.449	0.600	0.622	0.596	0.621	0.651	0.668	0.484
162	0.655	0.601	0.524	0.669	0.690	0.666	0.689	0.715	0.731	0.559

Table 2

D-Study, G-Coefficients, MLB 2005-2014, 80% Randomized Data

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
40	0.306	0.272	0.208	0.326	0.333	0.362	0.361	0.410	0.326	0.245
80	0.469	0.428	0.344	0.492	0.499	0.531	0.530	0.582	0.491	0.394
120	0.569	0.529	0.441	0.592	0.599	0.630	0.628	0.676	0.592	0.494
162	0.641	0.602	0.515	0.662	0.669	0.697	0.695	0.738	0.662	0.568

Using an alpha level of .05, a dependent samples *t* test was conducted to evaluate whether competitive balance differed significantly when analyzed as a full set of league data or as a random 80% sample of teams in the league. The results indicate that using the full set of data (*M* = .649, *SD* = .068) was similar to using a randomized 80% sample of the data (*M* = .645, *SD* = .066), with *t*(9) = .572, *p* = .582, *d* = .18. The results of the D-Study in Table 2 provide answers to the purpose of the study regarding absolute and relative decisions on competitive balance for MLB. First, it is clear that at no point during the observed years did any season reach the minimum level of reliability of 0.8. From an absolute decision, MLB is competitively balanced during the observed period. Playing a 162-game schedule in Major League Baseball is not enough to provide reliable estimates of the final rankings of the regular season. Second, we can see the relative levels of competitive balance by observing the G-coefficients at any of four points during each season. The data showed a range of full-season G-coefficients from .515 - .738, in 2007 and 2012 respectively. We are 95% confident that the interval (.604, .686) covers the mean difference in full-season G-coefficients for MLB from 2005-2014 (*M* = .645, *SD* = .066, *N* = 10. Additionally, the data displayed a median of .662 (*IQR* = .612 - .689). Based on the 95% CI and IQR we would expect individual seasons of MLB to be competitively balanced as we would not expect a season of MLB to cross the minimum level of reliability of 0.8. The data indicates that MLB has moved towards a less competitively balanced league over the 10 observed years as the G-coefficients have generally risen since 2005. While the three years with the least competitive balance at season’s end occurred from 2010 – 2012 (.697, .695, .738 respectively) they were followed by the second-most competitive year in 2014 (0.568).

Comparing different seasons as well as different points within those seasons provides us with information to make some inferences about competitive balance in MLB. The 2007 season provided the most competitively balanced season in the observed period with a full-season G-coefficient of 0.515. By comparison, at the 80-game mark of the 2010 (0.531), 2011 (0.530), and 2012 (0.582) seasons the G-coefficients matched or exceeded this number. Additionally, in each season except 2014, the full-season G-coefficient for 2007 was exceeded by the 120-game mark in every observed season. By any measure, the 2007 MLB season was the most competitively balanced of the observed period.

American League only (AL). The G-Study results for the American League from 2005-2014 show the majority of variance comes from the interaction component of teams and games (T x G). Each season displays a 0% variance from the Game (G) component. The percent variance of the Team (T) component ranges from 0.86% - 2.19% with a mean of 1.45%. The percent variance of the interaction component (T x G) ranges from 97.81% - 99.14% with a mean of 98.55%.

Table 3 displays the D-study results for the AL from 2005-2014 and Table 4 displays the results of the 80% randomized data D-study for the AL from 2005-2014.

Table 3

D-Study, G-Coefficients, AL 2005-2014

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
40	0.442	0.389	0.321	0.344	0.391	0.363	0.351	0.322	0.473	0.257
80	0.613	0.560	0.486	0.512	0.562	0.533	0.520	0.487	0.642	0.409
120	0.704	0.656	0.586	0.611	0.658	0.631	0.619	0.588	0.729	0.509
162	0.762	0.720	0.657	0.680	0.722	0.698	0.687	0.658	0.784	0.583

Table 4

D-Study, G-Coefficients, AL 2005-2014, 80% Randomized Data

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
40	0.433	0.378	0.332	0.354	0.385	0.357	0.358	0.309	0.465	0.249
80	0.604	0.548	0.499	0.522	0.556	0.526	0.527	0.472	0.635	0.398
120	0.696	0.645	0.599	0.621	0.653	0.625	0.626	0.573	0.723	0.498
162	0.756	0.711	0.668	0.689	0.717	0.692	0.693	0.644	0.779	0.573

Using an alpha level of .05, a dependent samples *t* test was conducted to evaluate whether competitive balance differed significantly when analyzed as a full set of league data or as a random 80% sample of teams in the league. The results indicate that using the full set of data (M = .695, SD = .057) was similar to using a randomized 80% sample of the data (M = .692, SD = .057), with $t(9) = 1.077$, $p = .310$, $d = .34$. As with the overall MLB D-study results, at no point did a season reach the minimum level of reliability of 0.8. In terms of an absolute decision, the AL is competitively balanced during the observed 10-year period. Playing a 162-game schedule in the AL is not enough to provide reliable estimates of final season rankings. The data show a range of full-season G-coefficients from .573 - .779, in 2014 and 2013 respectively. We are 95% confident that the interval (.657, .728) covers the mean difference in full-season G-coefficients for the AL from 2005-2014 (*M* = .692, *SD* = .057, *N* = 10). Additionally, the data displayed a median of .693 (*IQR* = .673 - .716). Based on the 95% CI and IQR we would expect individual seasons in the AL to be competitively balanced as we would not expect an individual season in the AL to cross the minimum level of reliability of 0.8. The G-coefficients for the AL trend differently than in MLB as a whole. There is not a clear pattern to determine if competitive balance in the AL is increasing or decreasing. The G-coefficients do show that in 7 of the 10 observed years, the G-coefficients for the AL were higher than MLB as a whole. The AL only had lower G-coefficients in the 2010-2012 seasons. This indicates that the AL, overall, is less competitively balanced than MLB as a whole.

National League only (NL). The G-Study results for NL from 2005-2014 reveal the majority of variance comes from the interaction component of teams and games (T x G). Each season displays a 0% variance from the Game (G) component. The percent variance of the Team (T) component ranges from 0.31% - 1.91% with a mean of 1.01%. The percent variance of the interaction component (T x G) ranges from 98.09% - 99.69% with a mean of 98.99%.

Table 5 displays the D-study results for the NL only from 2005-2014 and Table 6 displays the 80% randomized data D-study for the NL from 2005-2014.

Table 5

D-Study, G-Coefficients, NL 2005-2014

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
40	0.183	0.123	0.110	0.326	0.333	0.312	0.369	0.438	0.333	0.232
80	0.309	0.219	0.199	0.491	0.499	0.476	0.539	0.609	0.499	0.376
120	0.401	0.296	0.271	0.592	0.599	0.577	0.637	0.701	0.599	0.475
162	0.475	0.362	0.334	0.662	0.669	0.648	0.703	0.759	0.669	0.550

Table 6

D-Study, G-Coefficients, NL 2005-2014, 80% Randomized Data

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
40	0.190	0.128	0.111	0.334	0.332	0.135	0.395	0.446	0.339	0.261
80	0.319	0.228	0.201	0.501	0.499	0.237	0.566	0.617	0.506	0.414
120	0.412	0.307	0.273	0.601	0.599	0.318	0.662	0.707	0.606	0.515
162	0.486	0.374	0.337	0.670	0.668	0.386	0.726	0.765	0.675	0.589

Using an alpha level of .05, a dependent samples *t* test was conducted to evaluate whether competitive balance differed significantly when analyzed as a full set of league data or as a random 80% sample of teams in the league. The results indicate that using the full set of data (M = .583, SD = .147) was similar to using a randomized 80% sample of the data (M = .568, SD = .159), with $t(9)$ = .561, p = .588, d = .18. Similar to the overall MLB D-study results and AL D-study results, at no point did a season reach the minimum level of reliability of 0.8. From an absolute decision perspective the NL was competitively balanced during the observed period. Playing a 162-game schedule in the NL is not enough to provide reliable estimates of final season rankings. The data showed a range of full-season G-coefficients from .337 - .765, in 2007 and 2012 respectively. We are 95% confident that the interval (.469, .666) covers the mean difference in full-season G-coefficients for the NL from 2005-2014 (*M* = .568, *SD* = .159, *N* = 10). Additionally, the data displayed a median of .628 (*IQR* = .411 - .674). Based on the 95% CI and IQR we would expect individual seasons in the NL to be competitively balanced as we would not expect an individual season in the NL to cross the minimum level of reliability of 0.8. The G-coefficients for the NL trend similarly to MLB as a whole, with increases in G-coefficients and then a drop off in 2014; indicating competitive balance in the NL is decreasing. The G-coefficients show that in half of the observed years, the G-coefficients for the NL were lower than MLB as a whole. This indicates that the NL is more competitively balanced than the AL, but there is no clear trend as to the NL's competitive balance versus MLB as a whole.

NBA

The G-Study results for the NBA from 2005-2014 show the majority of variance comes from the interaction component of teams and games (T x G). Each season displays a 0% variance from the Game (G) component. The percent variance of the Team (T) component ranges from 5.74% - 10.65% with a mean of 8.60%. The percent variance of the interaction component (T x G) ranges from 89.35% - 94.26% with a mean of 91.40%.

Table 7 displays the D-study results for the NBA from 2005-2014 and Table 8 displays the results for the 80% randomized data D-Study for the NBA from 2005-2014.

Table 7

D-Study, G-Coefficients, NBA 2005-2014

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
20	0.570	0.549	0.693	0.705	0.674	0.666	0.644	0.646	0.654	0.677
40	0.726	0.709	0.818	0.827	0.805	0.800	0.784	0.785	0.791	0.807
60	0.799	0.785	0.871	0.877	0.861	0.857	0.845	0.846	0.850	0.863
82	0.844	0.833	0.902	0.907	0.894	0.891	0.857[a]	0.882	0.886	0.896

[a]The 2011 season was 66-game season, this denotes the G-coefficient at 66 games

Table 8

D-Study, G-Coefficients, NBA 2005-2014, 80% Randomized Data

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
20	0.588	0.576	0.726	0.677	0.643	0.604	0.644	0.665	0.669	0.705
40	0.740	0.731	0.841	0.807	0.783	0.753	0.784	0.798	0.802	0.827
60	0.810	0.803	0.888	0.863	0.844	0.820	0.845	0.856	0.858	0.878
82	0.854	0.848	0.916	0.896	0.881	0.862	0.857[a]	0.890	0.892	0.908

[a]The 2011 season was 66-game season, this denotes the G-coefficient at 66 games

Using an alpha level of .05, a dependent samples *t* test was conducted to evaluate whether competitive balance differed significantly when analyzed as a full set of league data or as a random 80% sample of teams in the league. The results indicate that using the full set of data (M = .879, SD = .025) was similar to using a randomized 80% sample of the data (M = .880, SD = .024), with $t(9) = -.262$, $p = .799$, $d = -.08$. The results of the D-Study in Table 8 provide answers to the purpose of the study regarding absolute and relative competitive balance decisions for the NBA. First, it is clear that every observed year in the NBA reach the minimum level of reliability of 0.8. Consequently, the NBA is not competitively balanced from an absolute decision perspective. The data showed a range of full-season G-coefficients from .848 - .916, in 2006 and 2007 respectively. We are 95% confident that the interval (.865, .895) covers the mean difference in full-season G-coefficients for the NBA from 2005-2014 (*M* = .880, *SD* = .024, *N* = 10). Additionally, the data displayed a median of .886 (*IQR* = .858 - .895). Based on the 95% CI and IQR we would expect individual seasons of the NBA to not be competitively balanced as we would expect an individual season of the NBA to cross the minimum level of reliability of 0.8. Four of the observed ten seasons reached a G-coefficient of 0.8 by the midpoint of the season, with the 2012 season also checking in at .798 at the midpoint. Playing an 82-game schedule in the National Basketball Association is enough to provide reliable estimates of the final rankings, and in four of the ten observed years it would take no more than 40 games to reach reliable estimates of the final rankings.

Second, we can see the relative levels of competitive balance by observing the G-coefficients at any of four points during each season. The data indicates that the NBA has a major competitive balanced problem over the ten observed years as the G-coefficients all exceed 0.8 by seasons end. Two seasons (2007 and 2014) even eclipse 0.9 with three other seasons (2008, 2012, and 2013) exceeding .890. There is also little fluctuation in the full-season G-coefficients as the highest G-coefficient (0.916 in 2007) is only 0.068 higher than the lowest observed full-season G-coefficient (0.848 in 2006). This is the smallest range in full-season G-coefficients for any of the leagues. This is also displayed in the small range of the 95% CI (.30) and IQR (.37) relative to other leagues. This information strengthens the argument that the NBA will continue to exhibit competitive imbalance because a season falling below 0.8 would represent an outlier in the data.

Comparing different points of different seasons in the data provide intriguing conclusions about the NBA. As noted, four of the ten observed seasons reached or eclipsed the minimum level of reliability of 0.8 by the 40-game mark. The 40-game G-coefficient for these four seasons are higher than any full-season observed mark for any of the other three sports. These results align with those of previous studies identifying the NBA as the least competitively balanced league. Of particular interest when comparing NBA seasons to each other are the results of the 2011 season because due to a lockout, the season was shortened to 66-games. Despite only playing 66 games, the full-season G-coefficient for the 2011 season (0.857) was still higher than the 2005 and 2006 seasons in which a full 82-game schedule was played. The data indicates that the NBA continues to have a serious problem with competitive balance.

NBA eastern conference only. The G-Study results for the NBA's Eastern Conference from 2005-2014 show that the majority of variance comes from the interaction component of teams and games (T x G). Each season except one (2008) displays a 0% variance from the Game (G) component. The percent of variance from the Game (G) component ranges from 0 – 0.83% with a mean of 0.08%. The percent variance of the Team (T) component ranges from 3.23% - 10.91% with a mean of 8.32%. The percent variance of the interaction component (T x G) ranges from 89.09% - 96.67% with a mean of 91.60%.

Table 9 displays the results of the D-study for the NBA's Eastern Conference from 2005-2014 and Table 10 displays the 80% randomized results for the NBA's Eastern Conference from 2005-2014.

Table 9

D-Study, G-Coefficients, NBA EAST 2005-2014

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
20	0.579	0.400	0.668	0.645	0.690	0.693	0.710	0.659	0.649	0.627
40	0.734	0.571	0.801	0.784	0.817	0.819	0.830	0.794	0.787	0.770
60	0.805	0.667	0.858	0.845	0.870	0.872	0.880	0.853	0.847	0.834
82	0.849	0.732	0.892	0.882	0.901	0.903	0.890[a]	0.888	0.883	0.873

[a]The 2011 season was 66-game season, this denotes the G-coefficient at 66 games

Table 10

D-Study, G-Coefficients, NBA EAST 2005-2014, 80% Randomized Data

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
20	0.589	0.467	0.726	0.693	0.689	0.618	0.696	0.666	0.655	0.657
40	0.741	0.637	0.841	0.818	0.816	0.764	0.821	0.800	0.792	0.793
60	0.811	0.725	0.888	0.871	0.869	0.829	0.873	0.857	0.851	0.852
82	0.855	0.783	0.916	0.902	0.901	0.869	0.883[a]	0.891	0.886	0.887

[a]The 2011 season was 66-game season, this denotes the G-coefficient at 66 games

Using an alpha level of .05, a dependent samples *t* test was conducted to evaluate whether competitive balance differed significantly when analyzed as a full set of league data or as a random 80% sample of teams in the league. The results indicate that using the full set of

data (M = .869, SD = .051) was similar to using a randomized 80% sample of the data (*M* = .877, *SD* = .037), with *t*(9) = -1.145, *p* = .282, *d* = -.36. With the exception of the 2006 season, the NBA's Eastern Conference met or exceeded the minimum level of reliability of 0.8 in every observed season. The data showed a range of full-season G-coefficients from .783 - .916, in 2006 and 2007 respectively. We are 95% confident that the interval (.854, .900) covers the mean difference in full-season G-coefficients for the NBA Eastern Conference from 2005-2014 (*M* = .877, *SD* = .037, *N* =10). Additionally, the data displayed a median of .887 (*IQR* = .872 - .898). Based on the 95% CI and IQR we would expect individual seasons in the NBA's Eastern Conference to not be competitively balanced as we would expect an individual season in the NBA's Eastern Conference to cross the minimum level of reliability of 0.8. Five of the ten observed seasons had G-coefficients exceeding 0.8 by the midpoint of the season. There is not a clear trend of an increase or decrease in competitive balance in the NBA's Eastern Conference during the observed period. Additionally, there is not a trend in comparison with the NBA as whole, as six of the full-season G-coefficients for the Eastern Conference fell above and four fell below the observed full-season G-coefficients for the NBA.

NBA Western Conference only. The G-Study results for the NBA's Western Conference from 2005-2014 reveal the majority of variance comes from the interaction component of teams and games (T x G). Each season displays a 0% variance from the Game (G) component. The percent variance of the Team (T) component ranges from 5.80% - 13.77% with a mean of 8.86%. The percent variance of the interaction component (T x G) ranges from 86.23% - 94.20% with a mean of 91.14%.

Table 11 displays the D-study results for the NBA's Western Conference from 2005-2014 and Table 12 displays the 80% randomized data D-study results for the NBA's Western Conference from 2005-2014.

Table 11

D-Study, G-Coefficients, NBA WEST 2005-2014

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
20	0.560	0.642	0.714	0.762	0.667	0.634	0.552	0.624	0.627	0.717
40	0.718	0.782	0.833	0.865	0.801	0.776	0.711	0.768	0.770	0.835
60	0.793	0.843	0.882	0.906	0.858	0.839	0.787	0.833	0.834	0.883
82	0.839	0.880	0.911	0.929	0.892	0.877	0.803[a]	0.872	0.873	0.912

[a]The 2011 season was 66-game season, this denotes the G-coefficient at 66 games

Table 12

D-Study, G-Coefficients, NBA WEST 2005-2014, 80% Randomized Data

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
20	0.619	0.647	0.667	0.780	0.597	0.667	0.564	0.633	0.598	0.763
40	0.765	0.786	0.801	0.876	0.748	0.801	0.721	0.775	0.749	0.865
60	0.830	0.846	0.858	0.914	0.816	0.858	0.795	0.838	0.817	0.906
82	0.869	0.883	0.892	0.936	0.859	0.892	0.810[a]	0.876	0.859	0.929

[a]The 2011 season was 66-game season, this denotes the G-coefficient at 66 games

Using an alpha level of .05, a dependent samples *t* test was conducted to evaluate whether competitive balance differed significantly when analyzed as a full set of league data or as a random 80% sample of teams in the league. The results indicate that using the full set of

data (*M* = .879, *SD* = .037) was similar to using a randomized 80% sample of the data (*M* = .880, *SD* = .036), with $t(9) = -.287$, $p = .780$, $d = -.09$. Just like the NBA as a whole, each observed season in the NBA's Western Conference met or exceeded the minimum level of reliability of 0.8.The data showed a range of full-season G-coefficients from .810 - .936, in 2011 and 2008 respectively. We are 95% confident that the interval (.858, .903) covers the mean difference in full-season G-coefficients for the NBA's Western Conference from 2005-2014 (*M* = .879, *SD* = .036, *N* = 10). Additionally, the data displayed a median of .879 (*IQR* = .862 - .892). Based on the 95% CI and IQR we would expect individual seasons in the NBA's Western Conference to not be competitively balanced as we would expect an individual season in the NBA's Western Conference to cross the minimum level of reliability of 0.8. Four of the ten seasons met or exceeded a 0.8 G-coefficient by the midpoint of the season. There is not a clear trend of an increase or a decrease in competitive balance in the NBA's Western Conference during the observed period. Additionally, there is not a clear trend in comparison to the NBA as a whole, as half of the seasons fell below and half of the seasons fell above the observed full-season G-coefficients for the NBA.

NHL

The G-Study results for the NHL from 2005-2014 show the majority of variance comes from the interaction component of teams and games (T x G). Each season displays a 0% variance from the Game (G) component. The percent variance of the Team (T) component ranges from 0.51% - 3.52% with a mean of 2.16%. The percent variance of the interaction component (T x G) ranges from 96.48% - 99.49% with a mean of 98.81%.

Table 13 displays the results of the D-study for the NHL from 2005-2014 and Table 14 displays the results for the 80% randomized data D-study for the NHL from 2005-2004.

Table 13

D-Study, G-Coefficients, NHL 2005-2014

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
20	0.422	0.378	0.093	0.273	0.254	0.249	0.213	0.351	0.343	0.381
40	0.594	0.549	0.170	0.429	0.406	0.398	0.351	0.520	0.511	0.551
60	0.687	0.646	0.236	0.530	0.506	0.498	0.448	0.565[a]	0.611	0.648
82	0.750	0.714	0.296	0.606	0.583	0.576	0.526		0.682	0.716

[a]The 2012 season was 48-game season, this denotes the G-coefficient at 48 games

Table 14

D-Study, G-Coefficients, NHL 2005-2014, 80% Randomized data

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
20	0.420	0.333	0.098	0.328	0.257	0.214	0.186	0.328	0.357	0.388
40	0.591	0.499	0.179	0.494	0.409	0.352	0.314	0.494	0.526	0.559
60	0.684	0.600	0.246	0.594	0.509	0.449	0.407	0.539[a]	0.624	0.655
82	0.748	0.672	0.308	0.667	0.587	0.527	0.484		0.694	0.722

[a]The 2012 season was 48-game season, this denotes the G-coefficient at 48 games

Using an alpha level of .05, a dependent samples *t* test was conducted to evaluate whether competitive balance differed significantly when analyzed as a full set of league data or as a random 80% sample of teams in the league. The results indicate that using the full set of

data (*M* = .601, *SD* = .132) was similar to using a randomized 80% sample of the data (*M* = .595, *SD* = .135), with $t(9) = .619$, $p = .551$, $d = .20$. The results of the D-Study in Table 14 provide answers to the purpose of the study regarding absolute and relative competitive balance decisions for the NHL. First, it is clear that every observed year in the NHL fell below the minimum level of reliability of 0.8. The data showed a range of full-season G-coefficients from .308 - .748, in 2007 and 2005 respectively. We are 95% confident that the interval (.511, .678) covers the mean difference in full-season G-coefficients for the NHL from 2005-2014 (*M* = .595, *SD* = .134, *N* = 10). Additionally, the data displayed a median of .627 (*IQR* = .530 - .689). Based on the 95% CI and IQR we would expect individual seasons of the NHL to be competitively balanced as we would not expect an individual season of the NHL to cross the minimum level of reliability of 0.8. Half of the observed 10 seasons failed to reach a G-coefficient of 0.6 at any point in the season. Playing an 82-game schedule in the National Hockey League is not enough to provide reliable estimates of the final rankings. Second, we can see the relative levels of competitive balance by observing the G-coefficients at any of four points during each season. The range of full-season G-coefficients for the NHL is the largest of the four observed sports. The highest full-season G-coefficient (0.748 in 2005) is 0.440 larger than the lowest full-season G-coefficient (0.308 in 2007). The data shows that compared to the NBA, which has the same season length, the NHL enjoys a much better level of competitive balance.

Conclusions on the relative competitive balance in the NHL can be drawn from comparing different seasons as well as different points within those seasons. The 60-game G-coefficient levels of 2005 (0.684), 2006 (0.600), 2013 (0.624), and 2014 (0.655) are higher than the full-season G-coefficient levels of five other observed seasons. Similar to the NBA, the NHL had a season shortened due to a lockout in 2012. As a result, the league only played a 48-game schedule. The 48-game, full-season G-coefficient for the 2012 season (0.539) was higher than three observed full-season G-coefficient marks. It was also higher than the 60-game G-coefficients of four observed seasons.

The 2007 season is also of interest, as it is the lowest observed full-season G-coefficient (0.308) for any of the entire league observations in the study. The 2007 full-season G-coefficient falls well outside of the 95% CI and IQR for the data. Its position well outside of the 95% CI and IQR indicates that the 2007 NHL season is an outlier. Table 16 shows the NHL’s Eastern Conference G-coefficients and shows that the low 2007 G-coefficient was mainly due to an extraordinarily low G-coefficient (.163) in the NHL’s Eastern Conference. As the lowest G-coefficient in the study, it means that the 2007 NHL season (and especially the Eastern Conference) exhibited the closest clustering of teams around a .500 winning percentage. The 2007 NHL season is the single most competitively balanced season of any league in the observed period. Additionally, the 2011 NHL season exhibits the second-lowest, full-season G-coefficient (0.484) of the entire league observations; and also falls outside of the 95% CI and IQR for the data.

NHL Eastern Conference only. The G-Study results for the NHL’s Eastern Conference from 2005-2014 show that the majority of variance comes from the interaction component of teams and games (T x G). Each season displays a 0% variance from the Game (G) component. The percent variance of the Team (T) component ranges from 0.17% - 3.38% with a mean of 2.17%. The percent variance of the interaction component (T x G) ranges from 96.62% - 99.83% with a mean of 97.83%.

Table 15 displays the D-study results for the NHL's Eastern Conference from 2005-2014 and Table 16 displays the 80% randomized data D-study results for the NHL's Eastern Conference from 2005-2014

Table 15

D-Study, G-Coefficients, NHL EAST 2005-2014

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
20	0.412	0.345	0.034	0.380	0.232	0.219	0.227	0.336	0.358	0.387
40	0.584	0.513	0.065	0.551	0.376	0.359	0.369	0.503	0.527	0.559
60	0.678	0.612	0.094	0.648	0.475	0.457	0.468	0.549[a]	0.626	0.655
82	0.742	0.683	0.124	0.716	0.553	0.535	0.546		0.696	0.722

[a]The 2012 season was 48-game season, this denotes the G-coefficient at 48 games

Table 16

D-Study, G-Coefficients, NHL EAST 2005-2014, 80% Randomized Data

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
20	0.441	0.332	0.045	0.408	0.269	0.201	0.244	0.342	0.345	0.410
40	0.612	0.498	0.087	0.580	0.424	0.334	0.393	0.510	0.513	0.582
60	0.703	0.598	0.125	0.674	0.525	0.430	0.492	0.555[a]	0.613	0.676
82	0.764	0.670	0.163	0.739	0.602	0.507	0.570		0.684	0.740

[a]The 2012 season was 48-game season, this denotes the G-coefficient at 48 games

Using an alpha level of .05, a dependent samples *t* test was conducted to evaluate whether competitive balance differed significantly when analyzed as a full set of league data or as a random 80% sample of teams in the league. The results indicate that using the full set of data (*M* = .587, *SD* = .183) was similar to using a randomized 80% sample of the data (*M* = .599, *SD* = .176), with *t*(9) = -1.663, *p* = .131, *d* =-.53. Similar to the NHL as a whole, there is not an observed season in the NHL's Eastern Conference that meets or exceeds the minimum level of reliability of 0.8. The data showed a range of full-season G-coefficients from .163 - .764, in 2007 and 2005 respectively. We are 95% confident that the interval (.490, .709) covers the mean difference in full-season G-coefficients for the NHL's Eastern Conference from 2005-2014 (*M* = .599, *SD* = .176, *N* = 10). Additionally, the data displayed a median of .636 (*IQR* = .559 - .725). Based on the 95% CI and IQR we would expect individual seasons in the NHL's Eastern Conference to be competitively balanced as we would not expect an individual season in the NHL's Eastern Conference to cross the minimum level of reliability of 0.8. There is not a clear trend of an increase or decrease in competitive balance in the NHL's Eastern Conference during the observed period. Additionally, there is not a clear trend in comparison to the NHL as a whole as half of the seasons fell below and half of the seasons fell above the observed, full-season G-coefficients for the NHL. As discussed earlier, the full-season G-coefficient for the NHL's Eastern Conference in 2007 is the lowest observed G-coefficient (0.163) for a conference and/or league in the entire study. Based on the 95% CI and IQR the 2007 season is a substantial outlier in the data and indicates that the 2007 NHL Eastern Conference represented the closest example of any league or conference to perfect competitive balance.

NHL Western Conference only. The G-Study results for the NHL's Western Conference from 2005-2014 show that the majority of variance comes from the interaction component of teams and games (T x G). Each season except one (2008) displays a 0% variance from the Game (G) component. The percent variance from the Game (G) component ranges from 0% - 0.30% with a mean of 0.03%. The percent variance of the Team (T) component ranges from 0.92% - 3.90% with a mean of 2.28%. The percent variance of the interaction component (T x G) ranges from 96.10% - 99.06% with a mean of 97.69%.

Table 17 displays the D-study results for the NHL's Western Conference from 2005-2014 and Table 18 displays the 80% randomized results for the NHL's Western Conference from 2005-2014.

Table 17

D-Study, G-Coefficients, NHL WEST 2005-2014

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
20	0.448	0.423	0.160	0.156	0.267	0.288	0.210	0.380	0.326	0.386
40	0.619	0.595	0.275	0.270	0.422	0.447	0.348	0.550	0.492	0.557
60	0.709	0.688	0.363	0.357	0.523	0.548	0.444	0.595[a]	0.592	0.653
82	0.769	0.751	0.438	0.432	0.599	0.624	0.522		0.665	0.720

[a]The 2012 season was 48-game season, this denotes the G-coefficient at 48 games

Table 18

D-Study, G-Coefficients, NHL WEST 2005-2014, 80% Randomized Data

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
20	0.431	0.443	0.175	0.194	0.261	0.319	0.231	0.383	0.298	0.426
40	0.602	0.614	0.298	0.325	0.415	0.483	0.375	0.554	0.459	0.598
60	0.694	0.705	0.389	0.419	0.515	0.584	0.474	0.599 [a]	0.560	0.690
82	0.756	0.766	0.465	0.496	0.592	0.657	0.552		0.635	0.753

[a]The 2012 season was 48-game season, this denotes the G-coefficient at 48 games

Using an alpha level of .05, a dependent samples *t* test was conducted to evaluate whether competitive balance differed significantly when analyzed as a full set of league data or as a random 80% sample of teams in the league. The results indicate that using the full set of data (M = .612, SD = .120) was similar to using a randomized 80% sample of the data (M = .627, SD = .107), with $t(9) = -1.791$, $p = .107$, $d = -.57$. In line with the results of the NHL as a whole and the NHL's Eastern Conference, there is not an observed season in the NHL's Western Conference that meets or exceeds the minimum level of reliability of 0.8. The data showed a range of full-season G-coefficients from .465 - .766, in 2007 and 2006 respectively. We are 95% confident that the interval (.561, .694) covers the mean difference in full-season G-coefficients for the NHL Western Conference from 2005-2014 (M = .627, SD = .107, N = 10). Additionally, the data displayed a median of .617 (IQR = .562 - .729). Based on the 95% CI and IQR we would expect individual seasons of the NHL's Western Conference to be competitively balanced as we would not expect an individual season of the NHL's Western Conference to cross the minimum level of reliability of 0.8. There is not a clear trend of an increase or decrease in competitive balance in the NHL Western Conference during the observed period. In comparison to the NHL as a whole, eight Western Conference seasons had higher full-season G-coefficients. This indicates that the Western Conference is less competitively balanced than the NHL.

NFL

The G-Study results for the NFL from 2005-2014 reveal the majority of variance comes from the interaction component of teams and games (T x G). Each season displays a 0% variance from the Game (G) component. The percent variance of the Team (T) component ranges from 6.91% - 12.06% with a mean of 9.75%. The percent variance of the interaction component (T x G) ranges from 87.94% - 93.09% with a mean of 90.25%.

Table 19 displays the results of the D-study for the NFL from 2005-2014 and Table 20 displays the results for the 80% randomized data D-study for the NFL from 2005-2014.

Table 19

D-Study, G-Coefficients, NFL 2005-2014

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
4	0.354	0.229	0.335	0.336	0.311	0.254	0.323	0.278	0.285	0.297
8	0.523	0.373	0.502	0.503	0.474	0.406	0.489	0.436	0.444	0.458
12	0.622	0.471	0.602	0.603	0.575	0.506	0.589	0.537	0.545	0.559
16	0.687	0.543	0.669	0.669	0.643	0.577	0.656	0.607	0.615	0.629

Table 20

D-Study, G-Coefficients, NFL 2005-2014, 80% Randomized Data

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
4	0.382	0.252	0.304	0.353	0.298	0.264	0.309	0.227	0.253	0.255
8	0.552	0.402	0.466	0.521	0.459	0.418	0.472	0.370	0.404	0.406
12	0.649	0.502	0.567	0.620	0.560	0.519	0.573	0.468	0.504	0.506
16	0.712	0.574	0.636	0.685	0.630	0.589	0.641	0.540	0.576	0.577

Using an alpha level of .05, a dependent samples *t* test was conducted to evaluate whether competitive balance differed significantly when analyzed as a full set of league data or as a random 80% sample of teams in the league. The results indicate that using the full set of data (M = .630, SD = .045) was similar to using a randomized 80% sample of the data (M = .616, SD = .054), with $t(9)$ = 1.258, p = .240, d = .40. The results of the D-Study in Table 20 provide answers to the purpose of the study regarding absolute and relative competitive balance decisions for the NFL. First, it is clear that every observed year in the NFL fell below the minimum level of reliability of 0.8. In absolute terms, the NFL is competitively balanced. Playing a 16-game schedule in the National Football League is not enough to provide reliable estimates of the final rankings. Second, we can see the relative levels of competitive balance by observing the G-coefficients at any of four points during each season. The data showed a range of full-season G-coefficients from .540 - .712, in 2012 and 2005 respectively. We are 95% confident that the interval (.582, .650) covers the mean difference in full-season G-coefficients for the NFL from 2005-2014 (M = .616, SD = .054, N = 10). Additionally, the data displayed a median of .610

(*IQR* = .576 - .640). Based on the 95% CI and IQR we would expect individual seasons of the NFL to be competitively balanced as we would not expect an individual season of the NFL to cross the minimum level of reliability of 0.8.

The data indicates that the NFL is trending towards more competitive balance, as four of the last five seasons represent four of the five most competitively balanced seasons in the observed time period. Relative to the other three leagues there has not been much change to the structures that we would expect to affect competitive balance in the NFL. Based on this relative lack of change, it is somewhat surprising to see a trend as there have not been forces enacted to move the level of competitive balance one way or the other. In addition to a trend towards more competitive balance there are other comparisons that can be made between seasons in order to provide more context. There appears to be little fluctuation in G-coefficients from year to year, as evident by the range of G-coefficients, the 95% CI, and the IQR, which are only smaller than those observed in the NBA. The relative lack of full-season G-coefficient range compared to the MLB and NHL observations may be in part due to the fact that he NFL has, by a large margin, the fewest number of observations per season in the study. The 12-game G-coefficient levels of the 2005 (0.649) and 2008 (0.620) seasons are higher than the full-season G-coefficient levels of five other observed full-seasons. .

American Football Conference only (AFC). The G-Study results for the AFC from 2005-2014 show that the majority of variance comes from the interaction component of teams and games (T x G). Each season displays a 0% variance from the Game (G) component. The percent variance of the Team (T) component ranges from

5.23% - 17.56% with a mean of 11.05%. The percent variance of the interaction component (T x G) ranges from 82.44% - 94.77% with a mean of 88.95%.

Table 21 displays the D-study results for the AFC from 2005-2014 and Table 22 displays the 80% randomized D-study results for the AFC from 2005-2014.

Table 21

D-Study, G-Coefficients, AFC 2005-2014

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
4	0.414	0.306	0.460	0.348	0.181	0.294	0.242	0.376	0.313	0.328
8	0.586	0.469	0.630	0.516	0.306	0.454	0.390	0.546	0.477	0.494
12	0.680	0.570	0.719	0.615	0.398	0.555	0.489	0.644	0.578	0.594
16	0.739	0.638	0.773	0.681	0.469	0.625	0.561	0.706	0.646	0.661

Table 22

D-Study, G-Coefficients, AFC 2005-2014, 80% Randomized Data

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
4	0.436	0.330	0.453	0.319	0.169	0.330	0.218	0.371	0.332	0.315
8	0.607	0.497	0.624	0.484	0.290	0.496	0.358	0.541	0.499	0.479
12	0.699	0.597	0.713	0.584	0.379	0.597	0.455	0.639	0.599	0.579
16	0.755	0.664	0.768	0.652	0.449	0.663	0.527	0.702	0.665	0.647

Using an alpha level of .05, a dependent samples *t* test was conducted to evaluate whether competitive balance differed significantly when analyzed as a full set of league data or as a random 80% sample of teams in the league. The results indicate that using the full set of

data (*M* = .650, *SD* = .087) was similar to using a randomized 80% sample of the data (*M* = .649, *SD* = .096), with $t(9) = .091$, $p = .930$, $d = .03$. The results of the D-Study show that every observed year in the AFC fell below the minimum level of reliability of 0.8. The data showed a range of full-season G-coefficients from .449 - .768, in 2009 and 2007 respectively. We are 95% confident that the interval (.590, .709) covers the mean difference in full-season G-coefficients for the AFC from 2005-2014 (*M* = .649, *SD* = .097, *N* = 10). Additionally, the data displayed a median of .664 (*IQR* = .649 - .693). Based on the 95% CI and IQR we would expect individual seasons in the AFC to be competitively balanced as we would not expect an individual season in the AFC to cross the minimum level of reliability of 0.8. There is not a clear trend of an increase or decrease in competitive balance in the AFC during the observed period. However, the data indicates some differences between the AFC and the NFL as a whole. First, there are three seasons in which the full-season G-coefficient for the AFC exceeds 0.7, which only occurs in one season for the entire NFL. Second, seven of the ten observed seasons in the AFC have higher full-season G-coefficients than the NFL overall. The data point toward the AFC is less competitively balanced than the NFL overall.

National Football Conference only (NFC). The G-Study results for the NFC from 2005-2014 show that the majority of variance comes from the interaction component of teams and games (T x G). Each season displays a 0% variance from the Game (G) component. The percent variance of the Team (T) component ranges from 3.70% - 15.53% with a mean of 8.88%. The percent variance of the interaction component (T x G) ranges from 84.47% - 96.30% with a mean of 91.12%.

Table 23 displays the D-study results for the NFC from 2005-2014 and Table 24 displays the 80% randomized D-study results for the NFC from 2005-2014.

Table 23

D-Study, G-Coefficients, NFC 2005-2014

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
4	0.294	0.133	0.183	0.340	0.424	0.224	0.403	0.173	0.275	0.274
8	0.455	0.235	0.309	0.508	0.595	0.366	0.575	0.295	0.431	0.431
12	0.556	0.315	0.402	0.607	0.688	0.464	0.670	0.385	0.532	0.531
16	0.625	0.381	0.473	0.674	0.746	0.535	0.730	0.455	0.603	0.602

Table 24

D-Study, G-Coefficients, NFC 2005-2014, 80% Randomized Data

	Year									
No. of games	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
4	0.327	0.131	0.193	0.327	0.372	0.247	0.416	0.159	0.306	0.257
8	0.493	0.232	0.324	0.493	0.542	0.396	0.587	0.274	0.468	0.409
12	0.593	0.312	0.418	0.593	0.639	0.496	0.681	0.361	0.569	0.509
16	0.660	0.377	0.489	0.660	0.703	0.567	0.740	0.430	0.638	0.580

Using an alpha level of .05, a dependent samples *t* test was conducted to evaluate whether competitive balance differed significantly when analyzed as a full set of league data or as a random 80% sample of teams in the league. The results indicate that using the full set of data (M = .582, SD = .120) was similar to using a randomized 80% sample of the data (M = .584, SD = .120), with $t(9) = -.227$, $p = .825$, $d = -.07$. The results of the D-study show that every observed year in the NFC fell below the minimum level of reliability of 0.8. The data showed a range of full-season G-coefficients from .377 - .740, in 2006 and 2011 respectively. We are 95% confident that the interval (.510, .659) covers the mean difference in full-season G-coefficients for the NFC from 2005-2014 (*M* = .584, *SD* = .120, *N* = 10). Additionally, the data displayed a median of .609 (*IQR* = .509 - .660). Based on the 95% CI and IQR we would expect individual seasons in the NFC to be competitively balanced as we would not expect an individual season in the NFC to cross the minimum level of reliability of 0.8. There is not a clear trend of an increase or decrease in competitive balance in the NFC during the observed period. However, the data indicates differences between the NFC and the NFL as whole. The data indicates the opposite of what was observed for the AFC, as six of ten observed seasons in the NFC have lower full-season G-coefficients than the NFL overall. The NFC has three seasons in which the full-season G-coefficient falls below the lowest observed full-season G-coefficient for the NFL overall (0.540 in 2012). The data point toward the NFC being more competitively balanced than the NFL overall and even more competitively balanced than the AFC.

Four League Comparison

Table 25 displays the full-season G-coefficients for each of the four leagues in the study and Table 26 displays the full-season G-coefficients for 80% of the randomized data for each of the four leagues.

Table 25

D-Study, G-Coefficients, Big Four Leagues 2005-2014

	Year									
League	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
MLB	0.655	0.601	0.524	0.669	0.690	0.666	0.689	0.715	0.731	0.559
NBA	0.844	0.833	0.902	0.907	0.894	0.891	0.857[a]	0.882	0.886	0.896
NHL	0.750	0.714	0.296	0.606	0.583	0.576	0.526	0.565[b]	0.682	0.716
NFL	0.687	0.543	0.669	0.669	0.643	0.577	0.656	0.607	0.615	0.629

[a]The 2011 NBA season was shortened to 66 games, this denotes the G-coefficient at 66 games

[b]The 2012 NHL season was shortened to 48 games, this denotes the G-coefficient at 48 games

Table 26

D-Study, G-Coefficients, Big Four Leagues 2005-2014, 80% randomized data

	Year									
League	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
MLB	0.641	0.602	0.515	0.662	0.669	0.697	0.695	0.738	0.662	0.568
NBA	0.854	0.848	0.916	0.896	0.881	0.862	0.857[a]	0.890	0.892	0.908
NHL	0.748	0.672	0.308	0.667	0.587	0.527	0.484	0.539[b]	0.694	0.722
NFL	0.712	0.574	0.636	0.685	0.630	0.589	0.641	0.540	0.576	0.577

[a]The 2011 NBA season was shortened to 66 games, this denotes the G-coefficient at 66 games

[b]The 2012 NHL season was shortened to 48 games, this denotes the G-coefficient at 48 games

Figure 1 charts the full-season G-coefficients per league, while Figure 2 charts the full-season G-coefficients per league for 80% of the randomized data.

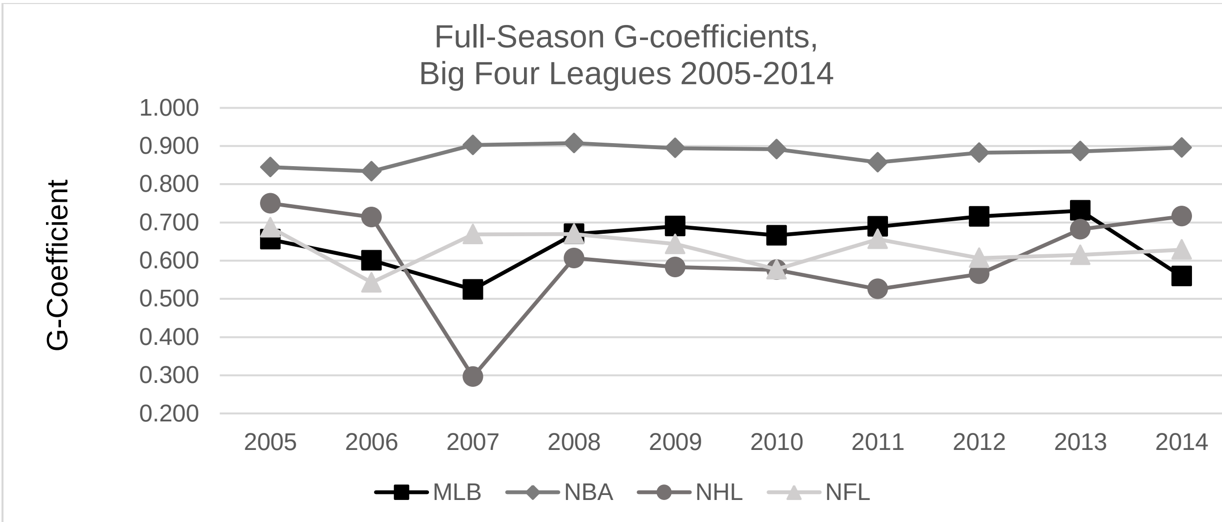

Figure 1. Four League G-Coefficient Comparison

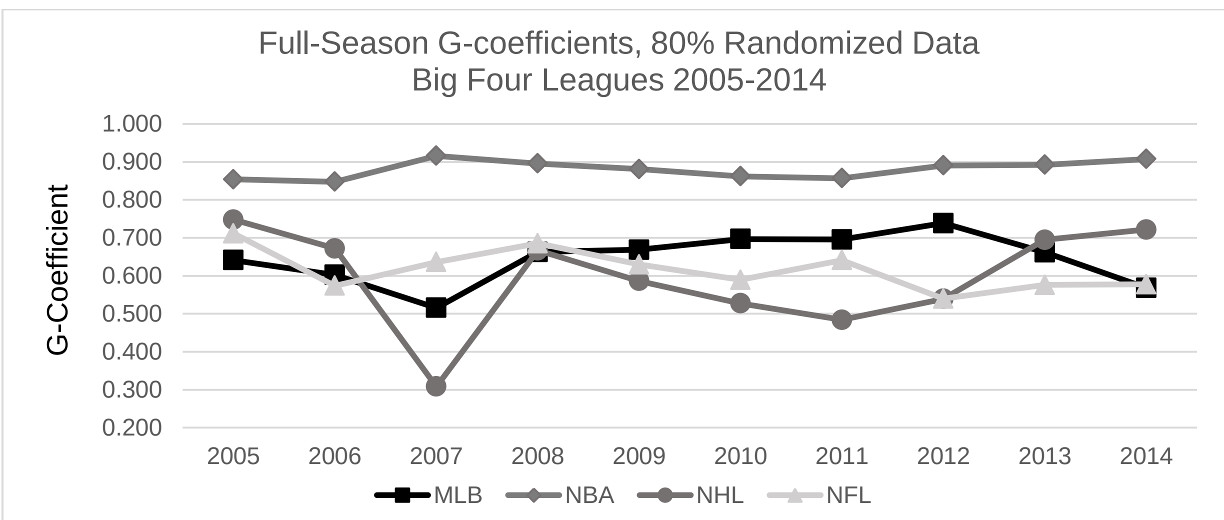

Figure 2. Four League G-Coefficient Comparison, 80% Randomized Data

In agreement with the existing literature, information provided in Table 26 and Figure 2 exhibit that it is apparent the NBA is the least competitively balanced as it exhibits the highest full-season G-coefficients. The data indicates the MLB, overall, is the second least competitive of the four leagues, although it has individual seasons that fall below the G-coefficients of the NHL and NFL. Further, at no point does any MLB season cross the minimum level of reliability of 0.8. Interestingly, determining the most competitively balanced league between the NHL and the NFL is a more difficult task. In the 10 observed seasons the NFL had a lower G-coefficient four times, and the NHL had a lower G-coefficient six times. The NFL has been more consistent in its observed G-coefficients while the NHL has tended to fluctuate more often, as evidenced by the wider ranges in both 95% CI's and IQR's for the NHL. The NHL exhibits a lower mean score (.595 to .616), but the NFL exhibits a lower median score (.610 to .627).

The results of the study facilitate several fascinating findings related to the use of Generalizability Theory as a competitive balance measure and the levels of competitive balance exhibited by the Big Four North American professional sports leagues. The first major finding, which speaks to both Generalizability Theory as a potential measure of competitive balance and the levels of competitive balance, is the relative rank of the leagues based on their levels of competitive balance. The NFL and NHL appear to be the most competitively balanced of the four leagues, with the ability to make a case for each as the single most competitively balanced. MLB is the next most competitively balanced, with the NBA exhibiting the least amount of competitive balance of the four leagues. The rank of most-to-least competitively balanced found in this study align with the ranks provided by Rockerbie (2012) in his study of the four leagues using RSD, the most often used measure of competitive balance. Related to this, the NBA was the only league of the four to exceed the minimum level of reliability of 0.8, and therefore be deemed not competitively balanced on absolute terms. The results point to the NBA not only being the least competitively balanced, but in a league of their own when it comes to competitive imbalance. These findings are in agreement with existing literature which positions the NBA as the least competitively balanced of the four leagues (Schmidt & Berri, 2003; Vrooman, 2009; Rockerbie, 2012). The major findings of this study are promising for the use of Generalizability Theory as a measure for competitive balance as it exhibits levels of competitive balance akin to those of previous studies employing accepted measures of competitive balance.

CHAPTER V

DISCUSSION

The purpose of this study is two-fold: to test the appropriateness of Generalizability Theory as a method for measuring competitive balance in professional sports, and to use the results to make absolute and relative decisions regarding levels of competitive balance in the Big Four North American professional sports leagues. The results of the study show promise for Generalizability Theory as an alternative method for measuring competitive balance because they correspond to the levels of competitive balance found by studies incorporating generally accepted measures of competitive balance. By comparing the results of this study to those of other known methods, we have established some validity for the use of Generalizability Theory as a measurement tool for competitive balance in sports. The results also present insight into the relative levels of competitive balance in the four North American professional sport leagues that require further discussion.

First, the results point to using Generalizability Theory as a method for measuring competitive balance to have merit worth further exploration. Generalizability Theory could serve as a useful tool in measuring competitive balance due to its ability to check relative levels of competitive advantage at any exact point of the season. For the purposes of this study, the researcher investigate four points in each season; but Generalizability Theory allows a researcher to check these levels at any point they desire with ease. The ability to do so would be very useful in pinpointing the exact point during a season where a league crossed a minimum level of reliability. This becomes of the utmost importance for a league like the NBA which regularly crosses the minimum level of reliability.

This study found the NBA to be the least competitively balanced of the four North American professional sport leagues. In fact, it was the only of the four leagues to have G-coefficients reach the minimum level of reliability of 0.8. Not only did the NBA exhibit the least amount of competitive balance relative to the other three leagues, it also surpassed the minimum level of reliability to make an absolute decision on the league's competitive balance. The NBA is the only league that is not competitively balanced. All ten seasons of the NBA reached a minimum level of 0.848, and therefore, all observed seasons in the NBA provided reliable estimates of the final rankings. At some point before each team played their 82^{nd} and final game, we are confident that we could have stopped any of the ten seasons observed in this study and ended up with a similar ranking of teams. The level of competitive imbalance is so pronounced in the NBA that in the 2007, 2008, 2013, and 2014 seasons we reach the minimum level of reliability by the midpoint of the season. By comparison, the highest observed full-season G-coefficient for each of the other three leagues was .748 (NHL in 2005), .738 (MLB in 2012), and .712 (NFL in 2005).

The result of the NBA being the least competitively balanced of the four leagues falls in line with the existing literature on competitive balance (Schmidt & Berri, 2003; Vrooman, 2009; Rockerbie, 2012). To reiterate this point, Rockerbie (2012) goes as far as to say the level of competitive balance in the NBA is significantly worse than each of the other three sports when measured using a relative standard deviation (RSD) method.

Our results fall in line with his interpretation since it was the only of the four leagues to exceed the minimum level of reliability. The point is exacerbated by the fact that four of the ten observed seasons exceeded the minimum level of reliability by the midpoint of the season. In those seasons the NBA would not have had to play more than 40 games in order to gain reliable estimates of the final rankings of teams for their 82-game schedule.

The results for MLB, the NHL, and the NFL tend to fall in line with the existing literature on competitive balance. As Rockerbie (2012) points out the NBA is the least competitively balanced, followed by MLB as the next least competitively balanced, with the NHL and NFL exhibiting the most competitive balance. Our results again agree with the existing literature as the same hierarchy of competitive balance was exhibited when using Generalizability Theory as the method of measurement. The mean full-season G-coefficients for our study were, ranked from least competitively balanced to most competitively balanced: NBA .880, MLB .645, NFL .616, NHL .595. We do see one striking difference in our results compared to the existing literature. While the consensus in the existing literature is that the NHL and NFL exhibit the greatest measures of competitive balance, it is generally agreed on that the NFL demonstrates by far the greatest level of competitive balance (Vrooman, 2009; Rockerbie, 2012; Lenten, 2015). Our results display the NHL actually exhibiting a greater level of competitive balance if you look at the mean scores of full-season G-coefficients (.595 for the NHL and .616 for the NFL). However, the NFL has a lower median score (.610) than the NHL (.627), and there is more fluctuation in the full-season G-coefficients for the NHL. This is exhibited by a range of (.308, .748), a 95% CI of (.511, .678), and an IQR of (.530 - .689) for the NHL. By comparison, the NFL exhibits a range of (.553, .712), a 95% CI of (.585, .647), and an IQR of (.576 - .640). Based on this data and the fact that the outlier G-coefficients of .308 from the 2007 season and, to a lesser extent, .484 from the 2011 NHL season substantially affects the NHL's overall mean G-coefficient, it is reasonable to advocate for the NFL being more competitively balanced. Despite possible dissention regarding the status as the NFL as the overwhelming leader in competitive balance; the overall results of the study support the existing literature on competitive balance which ranks the NBA as the least competitively balanced league, followed by MLB, and then the NHL and NFL.

Supported by the existing literature and the findings of the present study, Generalizability Theory as a method for measuring competitive balance has merit worth further investigation. With that in mind it is appropriate to move to the second portion of the purpose of the study: comparing the relative levels of competitive balance among the Big Four North American professional sports leagues. Our results demonstrate analogous distributions of competitive balance as found by accepted measures of competitive balance in the existing literature. Using Generalizability Theory provides a researcher with the unique ability to easily check the level of competitive balance at any point during a season, and compare that data to any other point of any other season. For the purposes of this study each season was partitioned in quarters, with the main purpose focusing on full-season levels of competitive balance. But with the ability to investigate each season with specific competitive balance levels, rather than just rankings, we would be remiss not to discuss the findings.

Perhaps the most striking result of our study is the relative level of competitive balance between the NHL and NFL. Existing literature holds that the NFL is far and away the most competitively balanced of the four leagues, yet our results show that making a distinction between the NFL and NHL as more competitive would be difficult. Rather, the NFL and NHL are both the most competitively balanced leagues. In terms of the structures in place to promote competitive balance and their intended effects it is not surprising that the NHL would rank at least closely to the NFL. The NHL and NFL are the only two leagues to enforce a hard salary cap. Salary caps have been shown to increase competitive balance (Fort & Quirk, 1995; Vrooman, 1996; Rascher, 1997); thus, our results should not be overly surprising. The NHL's salary cap was introduced in 2005, the first year of our study, so the ten years observed show a side by side comparison of competitive balance for two leagues which have enforced hard salary caps. Alternatively, the soft salary cap imposed by the NBA continues to look completely ineffective as a league with absolutely no salary cap (MLB) far outperforms them in terms of competitive balance. Both the NBA and MLB also enforce luxury taxes, and the NBA's system appears to be more favorable towards promoting competitive balance because a larger percentage of the collected tax money is redistributed to teams that did not incur luxury tax fines. Considering the systems in place, the results of this study together with previous studies indicate a hard salary cap as being an effective tool for promoting competitive balance. The results also show that a soft salary cap appears to be ineffective to the point where a league would be better off having no salary cap than a soft salary cap.

The existing literature on revenue sharing is not definitive regarding its effects on competitive balance. Yet, all four leagues have instituted revenue sharing programs with the intent of increasing competitive balance through a redistribution of wealth (Sloane, 2015). The tremendous growth in the value of television contracts has been pivotal in the development of revenue sharing programs for sports leagues (Mongeon & Winfree, 2012). The NFL has the most extensive revenue sharing program as 67% of its total revenues are shared evenly amongst the 32 teams. This level of revenue sharing is posited as a reason for the strong competitive balance of the league since the rate of sharing is almost 20% higher than the other three leagues (Staudohar, 2013). With changes to the revenue sharing systems in MLB, the NBA, and the NHL each of those leagues shares right around 50% of its revenues based on how the formulas work out each year for the NBA and NHL (MLB is set at 50%). Our results indicate that the salary cap structure, more than revenue sharing, is working to promote competitive balance. If revenue sharing were a significant factor we would expect the NFL to exhibit higher levels of competitive balance than the NHL since it has the same salary cap structure but shares revenues at a much higher rate.

If the information regarding salary cap structures, television contracts, and revenue sharing is looked at collectively, the evidence points to the salary cap structure having the most profound effect on competitive balance. We would expect MLB and the NBA to at least exhibit comparable levels of competitive balance to the NHL if the percentage of, and total amount of revenue being shared were a significant factor in determining competitive balance. Reasonably, we would expect the levels of competitive balance in MLB and the NBA to exceed those of the NHL given the lack of total NHL revenues compared to MLB and the NBA. Perhaps most damning to the argument of the importance of revenue sharing vis a vis competitive balance is what we would expect to find between the NFL and NHL. They both have a hard cap, so the most significant difference is the percentage of, and total amount of revenue shared. Even with the NFL sharing a higher percentage of revenues and having a television deal at least eight times larger than the NHL, our results show that the NHL is as, if not more, competitively balanced than the NFL. This runs somewhat counter to the findings of Vrooman (2009). He concluded that American sports leagues have been dominated by sportsmen owners since 1990, and that in a league of sportsmen owners, revenue sharing can increase competitive balance with or without salary caps. Again, if this were the case we would expect to see higher levels of competitive balance for MLB and the NBA. Truly, if revenue sharing were the most substantial variable for promoting competitive balance we would expect to find the NFL as by far the most competitively balanced league, in line with a bulk of the existing literature.

The imposition of salary caps by professional sports leagues has been done in order to promote competitive balance. Current literature holds that salary caps do, indeed, increase competitive balance (Quirk & Fort, 1992; Vrooman, 1996; Rascher, 1997; Késenne, 2000; Dietl et al., 2012). Fort and Quirk (1995) considered salary caps as the only truly effective structure leagues could impose to improve both competitive balance and financial stability. The understanding that salary caps in professional sports leagues increases competitive balance is so widely accepted that Zimbalist (2010) refers to cap systems as tools for promoting competitive balance as a grounded economic theory. Our findings agree with the extensive literature on the degree to which salary caps can promote and increase competitive balance; it appears to be the most important variable in determining competitive balance.

It should be noted that in an ever-changing business like professional sports there are more factors at play than those discussed in this study. As with any business, adapting to the shifting wants and needs of a customer base is essential for survival. However, it is important to never lose sight of what your core product is and how it is delivered. In sport the core product is the suspense of unscripted drama. Unscripted drama creates suspense because the result is an uncertainty of outcomes. Fans crave an uncertainty of outcomes so much that they purchase the product of sport in advance with an expectation that it will deliver.

This phenomenon of purchasing sport in advance certainly describes the actions of fans as they attend sporting events live, but it does not stop there. According to a study by the Consumer Electronics Association (CEA, 2013), 93% of Americans receive television programming from sources other than an antenna. They also found that 83% of Americans receive their television programming through traditional pay-for services such as cable and satellite. These numbers are important for putting into perspective just how much money Americans spend on sport-specific programming.

Cable and satellite providers pay each network that is part of their package a per-subscriber fee. In 2014, ESPN, which positions itself as the “worldwide leader in sports”, commanded the highest fee of any network, at $6.04 per subscriber. The next highest fee was for TNT, at $1.48 per subscriber (Molla, 2014). In fact, sport-specific networks accounted for four of the top ten networks in per-subscriber fees, and 57% of the total fees for those top ten networks. At $6.04 per subscriber, ESPN alone accounts for over 13% of the typical $45 cost for all channels, with a median channel cost of $0.14 (Molla, 2014). We do not have to attend live sporting events to purchase the product in advance; the vast majority of us engage in this process every month when we pay our television bill. According to a 2011 ESPN poll, fans are abandoning the experience of attending a sporting event live for the experience of watching at home (Rovell, 2012). The poll discovered that when given the choice between being at the game and watching at home, only 29% would prefer to be at the game, down from 54% in 1998. This information coupled with the dramatic rise in the value of television contracts for professional sports leagues point to a potential change in the formula for promoting competitive balance in the future.

There are many factors to consider for future studies related to Generalizability Theory and competitive balance. Future studies should look to weigh the efficacy of Generalizability Theory against specific, accepted measures of competitive balance. Doing so would provide more evidence towards, or possibly against, the further use of Generalizability Theory in the sport economic research. The present study served to provide an initial evaluation of Generalizability Theory as a measurement tool for competitive balance, and found that there is merit in its use. As noted, there are certainly more factors at play than the ones considered in this study, as such future studies should analyze the effects of other factors on competitive balance. As we continue to move towards the scenario where fans prefer to consume sporting events at home rather than attending them live, it is reasonable to assume the factors affecting the basic economics of sports leagues will continue to adapt. This will almost certainly affect the factors used to promote competitive balance. There is less risk involved with watching a game at home as opposed to attending it live. If the game becomes uninteresting, the channel can be changed without a monetary loss on the consumers' part. As television contracts and sport-specific programming continue to grow, perhaps competitive balance will become less of a concern for sport leagues. However, the continuing push by each of the four leagues in this study suggest that increasing competitive balance is as important as ever, despite the growing importance of television. Future studies should be cognizant of these changing factors and attempt to include them in studies using Generalizability Theory, or measuring competitive balance in general.

Limitations

While the findings of the present study indeed add to the existing literature on competitive balance in the four North American professional sport leagues, we must recognize its limitations. While the researcher used a randomized 80% sample in order to address the violation of independency of measures assumption which Generalizability Theory is subject to since it works within the ANOVA framework, there are still potential independency issues that should be noted. There is inherent dependency in win/loss data for a sports league as teams within the league influence the outcomes of other teams. Further investigations should be done to analyze this impact.

Additionally, this study analyzed competitive balance by using a one-facet design consisting of teams and games. This omits other potential facets that could represent sources of variance. As suggested, future studies should include more factors in order to ascertain their effect on competitive balance. The study analyzed a 10-year period for each of the four leagues; extending this period would provide a more comprehensive account of competitive balance.

Conclusion

By utilizing Generalizability Theory to measure competitive balance in the four North American professional sports leagues we can draw two distinct conclusions. The first is Generalizability Theory appears to be a viable method for measuring competitive balance in sports. The second is that our results align with existing literature on competitive balance as to the relative rankings of the four North American professional sports leagues regarding their levels of competitive balance. The NBA is the least competitively balanced of the four leagues and has a serious, ongoing issue as they are the only league to exhibit an absolute measure of competitive imbalance. The other three leagues exhibit competitive balance on absolute terms, with MLB ranking as the third most balanced league, and the NFL and NHL ranking as the most competitively balanced leagues.

REFERENCES

Atkinson, S.E., Stanley, L.R., & Tschirhart, J. (1988). Revenue sharing as an incentive in an agency problem: An example from the National Football League. *RAND Journal of Economics*, *19*, 27-43.

Barrabi, T. (2014, Oct 29). Bud Selig retirement: MLB's competitive balance is key to commissioner's legacy, but financial system remains imperfect. *International Business Times.* Retrieved from http://www.ibtimes.com/bud-selig-retirement-mlbs-competitive-balance-key-commissioners-legacy-financial-system-1715730

Baseball-Reference. (2015). Schedule and Results, 2005-2014. Retrieved from http://www.baseball-reference.com/

Basketball-Reference. (2015). Schedule and Results, 2005-2014. Retrieved from http://www.basketball-reference.com/

Beck, H. (2009, Jan 12). Got a question about the NBA salary cap? He's your man. *The New York Times*. Retrieved from http://www.nytimes.com/2009/ 01/13/sports/ basketball/ 13cap.html?_r=0

Berri, D.J., Schmidt, M.B. & Brook, S.L. (2007). The wages of wins: Taking measure of the many myths in modern sport. Stanford, CA: Stanford University Press.

Bien, L. (2015, Mar 2). The 2015 NFL salary cap explained. *SBNation*. Retrieved from http://www.sbnation.com/nfl/2015/3/2/8134891/nfl-salary-cap-2015-franchise-tag-explained

Bloom, H. (2014, Sep 5). NFL revenue-sharing model good for business. *Sporting News*. Retrieved from http://www.sportingnews.com/nfl/story/2014-09-05/nfl-revenue-sharing-television-contracts-2014-season-business-model-nba-nhl-mlb-comparison-salary-cap

Bontemps, T. (2013, Aug 8). Slip through the ‘Nets’: Payroll legal, but NBA wanted harder cap. *New York Post*. Retrieved from http://nypost.com/2013/08/08/slip-through-the-nets-payroll-legal-but-nba-wanted-harder-cap/

Borland, J. (1987). The Demand for Australian Rules Football. *Economic Record, 63*(3), 220-230.

Boron, A. (2013, July 3). NHL free agency primer: New rules, new dates for 2013 thanks to lockout, new CBA. *SBNation*. Retrieved from http://www.diebytheblade.com/2013/7/3/4489046/nhl-free-agency-primer-new-rules-2013

Brennan, R.L. (1992). *An NCME instructional model on Generalizability Theory. Instructional Topics in Educational Measurement, Module 14*. Madison, WI: National Council on Measurement in Education.

Briesch, A.M., Swaminathan, H., Welsh, M., & Chafouleas, S.M. (2014). Generalizability theory: A practical guide to study design, implementation, and interpretation. *Journal of School Psychology, 52*(1), 13-35.

Bryant, H. (2011, Sept 22). David Stern’s flawed NBA lockout logic. *ESPN*. Retrieved from http://espn.go.com/espn/commentary/story/_/id/7001461/david-stern-nba-lockout-logic-flawed

Butler, M. R. (1995). Competitive balance in Major League Baseball. *The American Economist, 39*(2), 46-52.

Buzzacchi, L., Szymanski, S., & Valletti, T. (2003). Equality of opportunity and equality of outcome: Open leagues, closed leagues and competitive balance. *Journal of Industry, Competition and Trade, 3*(3), 167-86.

Cain, L.P. & Haddock, D.D. (2006). Measuring parity: Tying into the idealized standard deviation. *Journal of Sports Economics, 7*(3), 330–338.

Cardinet, J., Tourneur, Y., & Allal, L. (1976). The symmetry of generalizability theory: Applications to educational measurement. *Journal of Educational Measurement, 13*, 119–135.

Condor, B. (2011, Apr 19). NHL, NBC sign record-setting 10-year TV deal. *NHL.com*. Retrieved from http://www.nhl.com/ice/news.htm?id=560238

Conger, A.J., Conger, J.C., Wallander, J., Ward, D., & Dygdon, J. (1983). A generalizability study of the Conners' teacher rating scale-revised. *Educational and Psychological Measurement, 43*, 1019–1031.

Consumer Electronics Assocation. (2013, July 30). Only seven percent of TV households rely on over-the-air signals, according to CEA study. Retrieved from https://www.ce.org/News/News-Releases/Press-Releases/2013-Press-Releases/Only-Seven-Percent-of-TV-Households-Rely-on-Over-t.aspx

Coon, L. (2015). Larry Coon's NBA salary cap FAQ: 2011 collective bargaining agreement. Retrieved from http://www.cbafaq.com/salarycap.htm#Q1

Crick, J.E., & Brennan, R.L. (1983). *GENOVA: A generalized analysis of variance system [Computer software and manual]*. Iowa City, IA: University of Iowa.

Cronbach, L. J. (1989). Review of the Stanford-Binet Intelligence Scale. *The tenth mental measurements yearbook*, 773-775. Lincoln, NE: Buros Institute of Mental Measurements.

Cronbach, L.J., Gleser, G.C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements*. New York, NY: Wiley.

Cronbach, L.J., Rajaratnam, N., & Gelser, G.C. (1963). Theory of generalizability: A liberalization of reliability theory. *The British Journal of Statistical Psychology, 16*, 137–163.

Crooker, J. R., & Fenn, A. J. (2007). Sports leagues and parity when league parity generates fan enthusiasm. *Journal of Sports Economics, 8*(2), 139-164.

Curran, J., Jennings, I. & Sedgwick, J. (2009). Competitive balance in the top level of English football, 1948-2008: An absent principle and a forgotten idea. *The International Journal of the History of Sport, 26*(11), 1735-1747.

Daly, G. & Moore, W.J. (1981). Externalities, property rights and the allocation of resources in major league baseball. *Economic Inquiry, 19*(1), 77-95.

Delgado, L. (2014, Dec 15). Explaining the luxury tax in Major League Baseball. *SportingCharts*. Retrieved from http://www.sportingcharts.com/articles/mlb/explaining-the-luxury-tax-in-major-league-baseball.aspx

Depken, C. A. (1999). Free-Agency and the competitiveness of Major League Baseball. *Review of Industrial Organization, 14*(3), 205-217.

Deubert, C., Wong, G.M., & Howe, J. (2012). All four quarters: A retrospective and analysis of the 2011 collective bargaining process and agreement in the National Football League. *UCLA Entertainment Law Review, 19*(1), 1-78.

Dietl, H.M., Franck, E., Lang, M., & Rathke, A. (2012). Salary cap regulation in professional team sports. *Contemporary economic policy*, *30*(3), 307-319.

Dixon, J. (2013). Scales of justice tilted against indigent persons. *Whittier Law Review, 35*, 513-529.

Dowbiggin, B. (2012, Sep 6). Welcome to the National Hockey League's parity party. *The Globe and Mail*. Retrieved from http://www.theglobeandmail.com/sports/hockey/welcome-to-the-national-hockey-leagues-parity-party/article1366952/

Eckard, E.W. (1998). The NCAA cartel and competitive balance in college football. *Review of Industrial Organization, 13*(3), 347–369.

Eckard, E.W. (2001). Baseball's blue ribbon economic report: Solutions in search of a problem. *Journal of Sports Economics*, *2*(3), 213-227.

Eckard, E. W. (2003). The ANOVA-based competitive balance measure. *Journal of Sports Economics, 4*(1), 74–80.

El-Hodiri, M., & Quirk, J. (1971). The economic theory of a professional sports league. *Journal of Political Economy, 79*, 1302-1319.

Epstein, A., & Kisska-Schulze, K. (2014). The NBA's 2011 collectively bargained amnesty clause - Exploring the fundamentals. *Atlantic Law Journal, 16*, 1-17.

ESPN. (2012, Oct 2). MLB completes new TV deals. *ESPN.com news services*. Retrieved from http://espn.go.com/mlb/story/_/id/8453054/major-league-baseball-completes-eight-year-deal-fox-turner-sports

Evans, R. (2014). A review of measures of competitive balance in the ‘analysis of competitive balance’ literature. *Bribeck Sports Business Centre Research Paper Series, 7*(2), 1-59.

Felder, A. (2014, Sep 5). Baseball’s new parity: A myth. *The Atlantic*. Retrieved from http://www.theatlantic.com/entertainment/archive/2014/09/bud-seligs-legacy-its-not-parity/379648/

Feldman, G.A. (2011, Feb 9). NFL lockout: The legal issues behind the NFL-CBA negotiations. *Huffington Post*. Retrieved from http://www.huffingtonpost.com /gabriel-a-feldman/the-legal-issues-behind-t_b_820579.html

Florio, M. (2013, Feb 28). 10 things to know about the franchise tag. *Pro Football Talk.* Retrieved from http://profootballtalk.nbcsports.com/2013/02/28/10-things-to-know-about-the-franchise-tag/

Forgues, B.L. (2012). Collective bargaining agreements and what has changed in the NBA and NFL. *The Verdict*. Retrieved from http://mslaw.edu/verdict-3/

Fort, R.D. (2003). Thinking (some more) about competitive balance. *Journal of Sports Economics, 4*, 280–283.

Fort, R., & Quirk, J. (1995). Cross-subsidization, incentives, and outcomes in professional team sports leagues. *Journal of Economic Literature*, *33*(3), 1265-1299.

Garica, A. (2010, Feb 14). Stern deflects union criticism, insists changes are needed. *NBA.com*. Retrieved from http://www.nba.com/2010/news/02/13/stern.presser/

George, T. (1992, July 31). Football; Tagliabue plugs line on N.F.L. restrictions. *The New York Times*. Retrieved from http://www.nytimes.com/1992/07/31/sports/football-tagliabue-plugs-line-on-nfl-restrictions.html

Goosens, K. (2006). Competitive balance in European football: Comparison by adapting measures: National measure of seasonal imbalance and top 3. *Rivista di Dirittoed Economia Dello Sport, II,* 77–122.

Groot, J., & Groot, L. (2003). The competitive balance of French football 1945-2002. *Économie Appliquée 56*(4), 91-113.

Grow, N. (2013). Decertifying players unions: Lessons from the NFL and NBA lockouts of 2011. *Vanderbilt Journal of Entertainment and Technology Law, 15*(3), 473-505.

Hadley, L., Ciecka, J. & Krautmann, A.C. (2005). Competitive balance in the aftermath of the 1994 players' strike. *Journal of Sports Economics, 6*(4), 379-389.

Haupert, M. (2014). New York, New York: A financial history of two baseball teams. *NINE: A Journal of Baseball History and Culture, 22*(2), 68-84.

Heika, M. (2014, June 6). NHL Commissioner Gary Bettman addresses several topics in his state of the union speech. *The Dallas Morning News*. Retrieved from http://starsblog.dallas news.com/2014/06/nhl-commissioner-gary-bettman-addresses-several-topics-in-his-state-of-the-union-speech.html/

Hellin, K. (2013, Aug 8). 2017 preview? Adam Silver says he wanted NBA hard salary cap. *NBC Sports*. Retrieved from http://probasketballtalk.nbcsports.com/2013/08/08/2017-preview-adam-silver-says-he-wanted-nba-hard-salary-cap/

Hockey-Reference. (2015). Schedule and Results, 2005-2014. Retrieved from http://www.hockey-reference.com/

Holbrook, E., Barreira, T., & Kang, M. (2009). Validity and reliability of Omron pedometers for prescribed and self-paced walking. *Medicine & Science in Sports & Exercise, 41*(3), 669-673.

Holbrook, E.A., Kang, M., & Morgan, D.W. (2013). Acquiring a stable estimate of physical activity in adults with visual impairment. *Adapted Physical Activity Quarterly, 30*(1), 59-69.

Horner, S.N. (2014) DMCA: Professional sports leagues' answer to protecting their broadcasting rights against illegal streaming. *Marquette Sports Law Review, 24*(2), 435-462.

Horowitz, I. (1997). The increasing competitive balance in Major League Baseball. *Review of Industrial Organization, 12*, 373-387.

Humphreys, B.R. (2002). Alternative measures of competitive balance. *Journal of Sports Economics, 3*(2), 133-148.

Humphreys, B.R. (2003). The ANOVA-based competitive balance measure: A reply. *Journal of Sports Economics, 4*(1), 81-82.

Ishikawa, S., Kang, M., Bjornson, K.F., & Song, K. (2013). Reliably measuring ambulatory activity levels of children and adolescents with cerebral palsy. *Archives of Physical Medicine and Rehabilitation, 94*(1), 132-137.

Isola, F. (2015). Adam Silver Q&A: NBA Commissioner discusses All-Star Game, potential changes to playoff system and future of Knicks and Nets. *New York Daily News*. Retrieved from http://www.nydailynews.com/sports/basketball/adam-silver-q-nba-commish-discusses-league-future-article-1.2110729

Justice, R. (2014, Nov 20). Selig hit a big homer on competitive balance. *MLB.com*. Retrieved from http://m.mlb.com/news/article/102020874/richard-justice-competitive-balance-seligs-biggest-hit

Kang, M., Bassett, D.R., Barreira, T.V., Tudor-Locke, C., Ainsworth, B., Reis, J.P., Strath, S. & Swartz, A. (2009). How many days are enough? A study of 365 days of pedometer monitoring. *Research quarterly for exercise and sport*, *80*(3), 445-453.

Kang, M., Bjornson, K., Barreira, T.V., Ragan, B.G., & Song, K. (2014). The minimum number of days required to establish reliable physical activity estimates in children aged 2–15 years. *Physiological measurement, 35*(11), 2229.

Keefer, Q.A.W. (2015). The sunk-cost fallacy in the National Football League: Salary cap value and playing time. *Journal of Sports Economics,* doi: 10.1177/1527 002515574515

Késenne, S. (1996). League management in professional team sports with win maximizing clubs. *European Journal for Sports Management, 2*, 14-22.

Késenne, S. (2000). Revenue sharing and competitive balance in team sports. *Journal of Sports Economics*, *1*(1), 56-65.

Késenne, S., & Pauwels, W. (2006). Club objectives and ticket pricing in professional team sports. *Eastern Economic Journal, 32*(3), 549-560.

Knowles, G., Sherony, K., & Haupert, M. (1992). The demand for major league baseball: A test of the uncertainty of outcome hypothesis. *American Economist, 36*, 72-89.

Kobritz, J. I., & Levine, J. F. (2013). Armageddon II: An analysis of the issues surrounding the NHL lockout. *American Journal of Management, 13*(1), 53-65.

Koning, R.H. (2000). Balance in competition in Dutch soccer. *Journal of the Royal Statistical Society: Series D (The Statistician), 49*(3), 419-431.

Koop, G. (2004). Modelling the evolution of distributions: An application to Major League baseball. *Journal of the Royal Statistical Society. Series A (Statistics in Society), 167*(4), 639-655.

Larsen, A., Fenn, A.J., & Spenner, E.L. (2006). The impact of free agency and the salary cap on competitive balance in the National Football League. *Journal of Sports Economics, 7*(4), 374-390.

Lee, Y.H. & Fort, R. (2005). Structural change in MLB competitive balance: The depression, team location and integration. *Economic Inquiry, 43*(1), 158-169.

Lenten, L.J. (2009). Towards a new dynamic measure of competitive balance: A study applied to Australia’s two major professional ‘football’ leagues. *Economic Analysis and Policy, 39*(3), 407-428.

Lenten, L.J. (2011). The extent to which unbalanced schedules cause distortions in sports league tables. *Economic Modelling, 28*(1), 451-458.

Lenten, L.J. (2015). Measurement of competitive balance in conference and divisional tournament design. *Journal of Sports Economics, 16*(1), 3-25.

Levin, R.C., Mitchell, G.J., Volcker, P.A.,& Will, G.F. (2000). *The report of the independent members of the Commissioner's Blue Ribbon Panel on Baseball Economics*. New York: Major League Baseball.

Lombardo, J. (2012, Jan 23). Inside NBA's revenue sharing. *Street & Smith's SportsBusiness Daily*. Retrieved from http://www.sportsbusinessdaily.com/ Journal/Issues/2012/01/23/ Leagues-and-Governing-Bodies/NBA-revenue.aspx

Looney, J. (2011, May 5). Goodell, Hunt stress importance of maintaining league parity. *KCChiefs.com*. Retrieved from http://www.kcchiefs.com/news/article-2/Goodell-Hunt-Stress-Importance-of-Maintaining-League-Parity/eebe6aca-6e49-4f96-a903-bcf8022fc480

Major League Baseball. (2012). 2012-2016 basic agreement. Retrieved from http://mlb.mlb.com/pa/info/cba.jsp

Major League Baseball. (2012). Summary of Major League Baseball Player Association – Major League Baseball labor agreement. *MLB.com*. Retrieved from mlb.mlb.com/mlb/ downloads/2011_CBA.pdf

Manuel, J. (2015, Jan 30). BA Q&A: MLB Commissioner Rob Manfred. *Baseball America*. Retrieved from http://www.baseballamerica.com/college/ba-qa-mlb-commissioner-rob-manfred/

Marburger, D.R. (1997). Gate revenue sharing and luxury taxes in professional sports. *Contemporary Economic Policy*, *15*, 114-123.

Matuszewski, E. (2006, Mar 20). NFL commissioner Paul Tagliabue will retire in July. *Bloomberg.com*. Retrieved from http://www.bloomberg.com/apps/news?pid=newsarchive&sid=anQWxnM4epew

McIndoe, S. (2014, Sep 3). The often forgotten 1994 NHL lockout. *Grantland*. Retrieved from http://grantland.com/the-triangle/the-often-forgotten-1994-nhl-lockout/

Michie, J., & Oughton, C. (2004). *Competitive balance in football: Trends and effects*. Football Governance Research Centre, Birkbeck College, University of London.

Milano, M., & Chelladurai, P. (2011). Gross domestic sport product: The size of the sport industry in the United States. *Journal of Sport Management, 25*(1), 24-35.

Mills, B., & Fort, R. (2014). League-level attendance and outcome uncertainty in US pro sports leagues. *Economic Inquiry*, *52*(1), 205-218.

Mitten, M., & Hernandez, A. (2014). The Sports Broadcasting Act of 1961: A comparative analysis of its effects on competitive balance in the NFL and NCAA Division I FBS football. *Ohio North University Law Review, 34*(3), 745-772.

Mizak, D., Neral, J. & Stair, A. (2007). The adjusted churn: An index of competitive balance for sports leagues based on changes in team standings over time. *Economic Bulletin*, *26*(3), 1-7.

Mizak, D. & Stair, A. (2004). The relationship between payroll and performance disparity in major league baseball: An alternative measure. *Economics Bulletin, 12*(9), 1-14.

Mizak, D., Stair, A. & Rossi, A. (2005). Assessing alternative competitive balance measures for sports leagues: A theoretical examination of standard deviations, gini coefficients, the index of dissimilarity. *Economic Bulletin, 12*(5), 1-11.

Molla, R. (Aug 5, 2014). How much cable subscribers pay per channel. *The Wall Street Journal*. Retrieved from http://blogs.wsj.com/numbers/how-much-cable-subscribers-pay-per-channel-1626/

Mongeon, K., & Winfree, J. (2012). Comparison of television and gate demand in the National Basketball Association. *Sport Management Review, 15*(1), 72-79.

Morrow, J. R. (1989). Generalizability theory. In M.J. Safrit (Eds.), *Measurement concepts in physical education and exercise science*, (pp. 73-96). Champaign, IL: Human Kinetics Books.

National Football League. (2011, Aug 4). Collective bargaining agreement. Retrieved from http://nflcommunications.com/current-cba/

National Football League. (2013, Mar 7). NFL free agency: Frequently asked questions. *NFL.com*. Retrieved from http://www.nfl.com/news/story/0ap1000000148425/article/nfl-free-agency-frequently-asked-questions

National Hockey League. (2013a, Jan 12). NHL/NHLPA proposed CBA – Summary of terms. Retrieved from http://cdn.agilitycms.com/nhlpacom/PDF/Summary-of-Terms-1-10-13.pdf

National Hockey League. (2013b, Feb 15). Collective bargaining agreement between National Hockey League and National Hockey League Players' Association. Retrieved from http://www.nhlpa.com/inside-nhlpa/collective-bargaining-agreement

Neale, W. (1964). The peculiar economics of professional sports. *Quarterly Journal of Economics, 78,* 1-14.

Nightengale, B. (2014, June 13). Is baseball's parity party a problem? *USA Today Sports*. Retrieved from http://www.usatoday.com/story/sports/mlb/2014/06/13/mlb-baseball-parity-bud-selig-competitive-balance/10395597/

Noll, R. (1974). Attendance and price setting. In R.G. Noll (Ed.), *Government and the sport business* (pp. 115-158). Washington, DC: Brookings Institution.

Noll, R. (1989). The economics of sports leagues. In Uberstine, G.A., Stratos, K.R., & Grad, R.J. (Eds). *Law of professional and amateur sports*. Deerfield, IL: The West Group.

Noll, R. (1991). Professional basketball: Economic and business perspectives. In Staudohar, P.D. and Mangan, J. A. (Eds). *The business of professional sports* (18-47). Champaign, IL: University of Illinois Press.

Nunnally, J.C., & Bernstein, I.H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.

O'Brien, J.E. (2014). *Competency assessment in nursing using simulation: A generalizability study and scenario validation process*. Retrieved from http:// repository.asu.edu/ attachments/137236/content/OBrien_asu_0010E_14115.pdf

Ourand, J., & Fisher, E. (2012, Sep 19). Fox, Turner to renew MLB packages; MLB net could get LDS games. *Street & Smith's SportsBusiness Daily*. Retrieved from http ://www.sportsbusinessdaily.com/Daily/Closing-Bell/2012/09/19/MLB-TV.aspx

Owen, P.D. (2010). Limitations of the relative standard deviation of win percentages for measuring competitive balance in sports leagues. *Economics Letters*, *109*(1), 38-41.

Owen, P.D., & King, N. (2015). Competitive balance measures in sports leagues: The effects of variation in season length. *Economic Inquiry*, *53*(1), 731-744.

Owen, P.D., Ryan, M, & Weatherspoon, C.R. (2007). Measuring competitive balance in professional team sports using the Herfindahl-Hirschman index. *Review of Industrial Organisation*, *31*(4), 289-302.

Pantuosco, L.J., & Stone, G.L. (2007). Capitalism for the cooperative: The NCAA and NFL model of parity and profit. *Journal of Economics and Economic Education Research*, *8*(2), 65-92.

Patra, K. (2015, Mar 2). NFL salary cap will be at $143.28 million in 2015. *NFL.com*. Retrieved from http://www.nfl.com/news/story/0ap3000000475775/article/nfl-salary-cap-will-be-14328-million-in-2015

Peeters, T. (2015). Profit-maximizing gate revenue sharing in sports leagues. *Economic Inquiry*, *53*(2), 1275-1291.

Piercy, J. (2012, Aug 17). Gary Bettman's language of NHL labour talks, then & now. *CBC Sports*. Retrieved from http://www.cbc.ca/sports/hockey/nhl/gary-bettman-s-language-of-nhl-labour-talks-then-now-1.1264736

Plunkett Research, Ltd. (2014, Nov 11). *Industry statistics sports & recreation business statistics analysis*. Retrieved from http://www.plunkettresearch.com/statistics/ sports-industry/

Posnanski, J. (2014, Feb 3). Fair or foul? Parity in American sports. *NBCSports.com*. Retrieved from http://www.nbcsports.com/football/nfl/fair-or-foul-parity-and-american-sports

Pro-Football-Reference. (2015). Games & Schedule, 2005-2014. Retrieved from http://www.pro-football-reference.com/

Proteau, A. (2012, Aug 31). Total parity not the best solution for the NHL. *The Hockey News*. Retrieved from http://sports.yahoo.com/news/total-parity-not-best-solution-183000999--nhl.html

Purdy, M. (2014, June 17). Bad TV deal could threaten Sharks' ability to stay in San Jose. *San Jose Mercury News*. Retrieved from http://www.mercurynews.com /sharks/ci_25983493/ purdy-bad-tv-deal-could-threaten-sharks-ability

Quirk, J., & Fort, R. (1992). *Pay dirt: The business of professional team sports*. Princeton, New Jersey (NJ): Princeton University Press.

Ramachandran, S., & Clark, K. (2014, Sep 3). In television deals, NFL plays by its own rules. *The Wall Street Journal*. Retrieved from http://www.wsj.com/articles/in-television-deals-nfl-plays-by-its-own-rules-1409797802

Rascher, D.A. (1997). A model of a professional sports league. *International Advances in Economic Research, 3*(3), 327-328.

Rockerbie, D. (2012). Exploring inter-league parity in North America: the NBA anomaly. MPRA Paper, University Library of Munich, Germany. Retrieved from http://econpapers.repec.org/paper/pramprapa/43088.htm

Ross, S.F., & Lucke, R.B. (1997). Why highly paid athletes deserve more antitrust protection than ordinary unionized workers. *The Antitrust Bulletin, 42*(3), 641-679.

Rottenberg, S. (1956). The baseball player's labor market. *Journal of Political Economy*, *64*, 242-258.

Rovell, D. (2012, Nov 16). Best seat in the house. *ESPN.com*. Retrieved from http://espn.go.com/espn/otl/story/_/id/8636927/nfl-taking-note-many-fans-watching-games-tv-beats-going-stadiums

Salvia, J., Ysselydke, J.E., & Bolt, S. (2010). *Assessment in special and inclusive education* (11th ed.)Boston, MA: Houghton Mifflin.

Sanderson, A.R. (2002). The many dimensions of competitive balance. *Journal of Sports Economics*, *3*(2), 204-228.

Sandomir, R. (2014, Oct 5). NBA is said to continue network deals. *The New York Times*. Retrieved from http://www.nytimes.com/2014/10/06/sports/basketball/nba-said-to-be-near-new-tv-deal-for-24-billion.html?_r=0

Sanderson, A.R., & Siegfried, J.J. (2003). Thinking about competitive balance. *Journal of Sports Economics*, *4*(4), 255-279.

Schmidt, M.B. (2001). Competition in Major League Baseball: The impact of expansion. *Applied Economics Letters*, *8*(1), 21-26.

Schmidt, M.B., & Berri, D.J. (2001). Competitive balance and attendance: The case of Major League Baseball. *Journal of Sports Economics, 2*(2), 145-167.

Schmidt, M.B., & Berri, D.J. (2003). On the evolution of competitive balance: The impact of an increasing global search. *Economic Inquiry*, *41*(4), 692-704.

Scully, G.W. (1989). *The business of Major League Baseball*. Chicago: University of Chicago Press.

Settimi, C. (2013, Nov 25). The NHL's richest local television deals. *The New York Times*. Retrieved from http://www.forbes.com/sites/christinasettimi/2013/11/25/nhls-richest-local-television-deals/2/

Settimi, C. (2014, Jan 22). The NBA's richest local television deals. *Forbes*. Retrieved from http://www.forbes.com/sites/christinasettimi/2014/01/22/the-nbas-richest-local-television-deals/

Shavelson, R.J., Webb, N.M., & Rowley, G.L. (1989). Generalizability theory. *American Psychologist*, *44*, 922–932.

Shoukri, M.M., Asyali, M.H., & Donner, A. (2004). Sample size requirements for the design of reliability study: Review and new results. *Statistical Methods in Medical Research, 13*, 251–271.

Sijtsma, K., & van der Ark, L.A. (2015). Conceptions of reliability revisited and practical recommendations. *Nursing research*, *64*(2), 128-136.

Sloane, P.J. (2015). The economics of professional football revisited. *Scottish Journal of Political Economy*, *62*(1), 1-7.

Smith, P.L. (1981). Gaining accuracy in generalizability theory: Using multiple designs. *Journal of Educational Measurement, 18,* 147–154.

Staudohar, P.D. (2005). The hockey lockout of 2004-05. *The Monthly Labor Review, 128*, 23-29.

Staudohar, P.D. (2013). The hockey lockout of 2012-2013. *The Monthly Labor Review, 136*, 1-10.

Strycker, L.A., Duncan, S.C., Chaumeton, N.R., Duncan, T.E., & Toobert, D.J. (2007). Reliability of pedometer data in samples of youth and older women. *International Journal of Behavioral Nutrition and Physical Activity*, *4*(1), 4-11.

Stubits, B. (2013, May 31). With last four cup champions still standing, is NHL parity a myth? *CBSSports.com*. Retrieved from http://www.cbssports.com/nhl/eye-on-hockey/22333246/with-last-four-cup-champs-still-standing-is-nhl-parity-a-myth

Szymanski, S., & Kuypers, T. (1999). *Winners and losers*. London: Viking.

Szymanski, S., & Késenne, S. (2004). Competitive balance and gate revenue sharing in team sports. *The Journal of Industrial Economics, LII*(1), 165-177.

Tainsky, S., & McEvoy, C.D. (2012). Television broadcast demand in markets without local teams. *Journal of Sports Economics, 13*(3), 250-265.

Topkin, M. (2015, Mar 26). Commissioner Rob Manfred calm on Rays stadium issue. *Tampa Bay Times*. Retrieved from http://www.tampabay.com/sports/baseball/rays /rays-pepper-new-commissioner-rob-manfred-with-questions/2223019

Totty, E.S., & Owens, M.F. (2011). Salary caps and competitive balance in professional sports leagues. *Journal for Economic Educators, 11*(2), 46-56.

Trandel, G.A., & Maxcy, J.G. (2011). Adjusting winning-percentage standard deviations and a measure of competitive balance for home advantage. *Journal of Quantitative Analysis in Sports*, *7(*1), 1-15.

Utt, J., & Fort, R. (2002). Pitfalls to measuring competitive balance with gini coefficients. *Journal of Sports Economics, 3*(4), 367-373.

Uyar, B., & Surdam, D. (2012). Searching for on-field parity: Evidence from national football league scheduling during 1991-2006. *Journal of Sports Economics*, *14*(5), 479-497.

Verducci, T. (2014, June 17). Mediocre baseball league: Extreme parity a result of Selig's plan. *Sports Illustrated*. Retrieved from http://www.si.com/mlb/2014/06/17/mediocre-teams

Vrooman, J. (1995). A general theory of professional sports leagues. *Southern Economic Journal*, *61(*4), 971-990.

Vrooman, J. (1996). The baseball player's labour market reconsidered. *Southern Economic Journal*, *63*(2), 339-360.

Vrooman, J. (2009). Theory of the perfect game: Competitive balance in monopoly sports leagues. *Review of Industrial Organization*, *34*(1), 5-44.

Vrooman, J. (2015). Sportsman leagues. *Scottish Journal of Political Economy*, *62*(1), 90-115.

Weinstein, A. (2015, Mar 18). MLB Commissioner Manfred: International draft would help ensure competitive balance. *Sporting News*. Retrieved from http://www. sportingnews. com /mlb/story/2015-03-18/mlb-commissioner-rob-manfred-baseball-international-draft-cuban-defectors-puig-cespedes-abreu

Wood, D.B. (2011, Dec 9). Chris Paul trade rejected: Was it really to save NBA parity? *The Christian Science Monitor*. Retrieved from http://www.csmonitor.com/USA/ Sports/2011/ 1209/ Chris-Paul-trade-rejected-Was-it-really-to-save-NBA-parity

Wyshynski, G. (2013, Nov 26). NHL Canadian TV deal: Winners and losers of the 12 year pact with Rogers. *Yahoo! Sports.* Retrieved from http://sports.yahoo.com/ blogs/nhl-puck-daddy/nhl-canadian-tv-deal-winners-losers-12-pact-212452256--nhl.html

Zillgitt, J. (2013, Feb 17). NBA Commissioner David Stern on the state of the league. *USA Today*. Retrieved from http://www.usatoday.com/story/sports/nba/2013/02/ 16/nba-david-stern-state-of-game/1925229/

Zimbalist, A. (2002). Competitive balance in sports leagues: An introduction. *Journal of Sports Economics, 3*(2), 111-121.

Zimbalist, A. (2010). Reflections on salary shares and salary caps. *Journal of Sports Economics, 11*(1), 17-28.