THE MANY DIMENSIONS OF COMPETITIVE BALANCE AND THE ATTENDANCE OF MAJOR LEAGUE BASEBALL

BY

XINRONG LEI

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Recreation, Sport and Tourism
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Doctoral Committee:

Professor Brad Humphreys, Director of Research
Associate Professor Bruce Wicks, Chair
Professor Xuming He, Contingent Chair
Assistant Professor Megan Janke, University of South Florida

ABSTRACT

This research proposed a set of measures of Competitive Balance which aims to address three dimensions of Competitive Balance: Closeness, Dominance and Consistency. Longitudinal MLB data is used for empirical evaluation purpose. The matched pair of teams is used as the basic research object in this study, and the growth model is applied to analyze the relationship between game attendances and the proposed measures of Competitive Balance. Research confirmed that Competitive Balance is multidimensional, and not every dimension of Competitive Balance is correlated with game attendance. Fans prefer changes, and they are not attracted by consecutive wins or losses. Rather fans are more like to go to games that can potentially affect teams' standings in their divisions or league. Fans show no specific preferences to upset games.

To Father and Mother

ACKNOWLEDGEMENT

I would never have been able to go this far without the support of my committee members and help from my friends.

I would like to express my deepest gratitude to my advisor, Professor Brad Humphreys, without his encouragement, guidance and persistent help this dissertation would not have been possible. In my hardest time, when doubting my direction and even the PHD program, it was Professor Brad Humphreys who guided me through the maze, flew across the county, and did everything he could to help me in my research. I am heartily thankful to my committee Chair Professor, Bruce Wicks, who patiently offered me a great deal of help time and again, and has always been the one I turned to for help when I found myself neck deep. I owe a deep gratitude to Professor Xuming He who participated on my committee as contingent chair, and was always supportive during the whole procedure. His wisdom and availability in campus gave me great confidence and energy and was vitally important to my success. I am grateful to Dr. Megan Janke, who helped me while she was at UIUC, and after leaving campus was kind enough to continue her participation on my committee and support me in every aspect.

My special thanks go to Jason Shahan, who always cheered me up and offered me his excellent suggestions in regards to writing and revising. Without him, it would be a lonely journey. My thanks and appreciations also go to Lauren Teffeau, Professor Leigh Estabrook and Professor Carole Palmer, my supervisors in GSLIS, who have provided me steady great support and made my research possible. It is a pleasure to thank Tiffany Richardson and Nicholas Watanabe, especially Nicholas for taking their time to discuss research and proof read my draft.

Lastly, I offer my best regards to all of those who supported me in any respect during the completion of the project, such as my friends in the Department of Statistic and friends in Graduate School of Library and Information Science.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

1.1 Background

Competition has been recognized as a positive factor in sports for over two-hundred years. In 1898, the first study in social psychology by Norman Triplett, showed that competition leads to better performance. He found that the cyclists' best records were always set when they were competing against others. In addition, he used other forms of sport to confirm his observation, such as wheel races and boat races. He concluded that the"…bodily presence of another contestant participating simultaneously in the race serves to liberate latent energy not ordinarily available" (Triplett, 1898). Following Triplett, researchers continued their studies on group dynamics and found that competition plays an important role in participant's performance as well. Competitive Balance is a main factor for maintaining diversity and innovation in the sports industry.

At the same time, competition is one factor that attracts sport spectators to sporting events. From a psychological perspective, for any performance-related activities, there is an optimal amount of stress when competition is introduced (Triplett, 1898), and this point of view is consistent with the concept of "eustress", which is a frequently referred to motivation for sports spectator's behavior (Branscombe & Wann, 1991; Branscombe & Wann, 1991; Branscombe & Wann, 1994; Sloan, 1989).

Eustress was originally developed by Richard Lazarus and is defined as a pleasant or curative stress in our life. Opposite of distress, eustress is healthy and is related to a feeling of fulfillment. Sport is enjoyable because it provides spectators with the stress they seek (Zuckerman, 1979), a stress which comes from the competition between teams. Therefore, competition between teams can not only boost better sport performance, but also attract more sports spectators leading to an increase in attendance.

In addition to providing stress, a well balanced game may arouse a spectator's curiosity, which is another intrinsic motivation intensively examined by researchers. Curiosity is defined as the desire to know, to see, or to experience that motivates exploratory behavior directed towards the acquisition of new information (Litman & Jimerson, 2004; Loewenstein, 1994) and is identified as an important component in the decision making process of sport spectators. Loewenstein called his theory as the "information-gap theory", that is, when people feel a gap between what they know and what they want to know, they are motivated to fill the gap and fulfill their curiosity. Researchers noticed that the approach to discovering information should also be pleasant; it should dispel undesirable states of ignorance or uncertainty rather than stimulate one's interest information (Litman & Jimerson, 2004; Loewenstein, 1994). Litman (2005) further promoted the optimal arousal model and curiosity-drive theory (Szymanski, 2003a; Szymanski, 2003b) which identified three components of demand for sports contents: demand of quality, the success of specific contestants, and the uncertainty of contest.

1.2 Theory of Uncertainty of Outcome

Consistent with the research on eustress and curiosity in psychology, the theory of uncertainty of outcome in sport economics states that unpredictable sporting events are desired by spectators (Rottenberg, 1956; Sloane, 1971; Neale, 1964; Canes,1974). If a league lacks Competitive Balance, spectators will lose interest in the games; if the outcome of the game is too obvious, spectators will not bother to attend the game. Spectators soon lose interest in a perennial loser or even in a team that always wins (Leeds & von Allmen, 2002). In return, the team performances will be further diminished due to the lack of social facilitation from the spectators (Amabile, 1996).

Consistent again with the theory of curiosity, the conventional wisdom supports that Competitive Balance must exist in healthy team sport leagues (Zimbalist, 2003). It is well accepted that uncertainty of outcome will increase attendance at sport events, and when the competitive pressure is absent, arrogance, laxity and inefficiency are fostered (Zimbalist, 2003). Uncertainty outcome helps to fulfill spectators' needs for suspense, thus bringing a thrill to spectators and making the game more enjoyable (Knobloch-Westerwick, David, Eastin, Tamborini & Greenwood, 2009). Uncertainty of outcome also has incentive effects on an athlete's performance, and therefore improves the game quality. Under the assumption that game quality and uncertainty are desired by the market, the more uncertain the result, the more spectators will attend the game (Fort & Quirk, 2004; Fort & Maxcy, 2003; Fort, 2003).

Because increased attendance is desirable by all sport participants and it is commonly asserted that promoting Competitive Balance in sports leagues will increase attendance (Forrest, Beaumont, Goddard, & Simmons, 2005a), the effort for promoting Competitive Balance is not only limited to academia, but also continues to be a concern within the industry. Many researchers may not be in agreement with the operational rules promoted by league Commissioners and owners with a 'league thinking', which is to make decisions for the sustainable benefits of all participants. In addition, many researchers do not agree with each other about how much Competitive Balance is needed in sport leagues, if any. Therefore, more research is necessary to examine these positions, especially the problems that have not yet been addressed, which are the challenges to be discussed in the following section.

1.3 Challenges

First, it is hard to determine the optimal amount of Competitive Balance because evenly distributed Competitive Balance is not a desirable thing to sports leagues. Based on Triplett's findings, when the "optimal amount of stress" is introduced into sport, athletes may perform better, and the audiences may enjoy the game more. However, it is hard to define the magnitude of the "optimal stress". Therefore there continues to be a tension between the needs of the league being more competitive versus the spectator's level of enthusiasm for the truly memorable teams, which fans talk about and sports writers write about for years to come, for example the 1927 New York Yankees, the 1921 Philadelphia Athletics, and the 1962 Green Bay Packers (Quirk & Fort, 1992).

Leeds (2008) argued that leagues may generate higher attendance and increased profits if the large market teams win more often; The same argument is made by Quirk and Fort's analysis of league market equlibum (Fort & Quirk, 2004) demonstrating that a league's income decreases when the large market team is defeated in its games. Therefore, it is of greater incentive to the league's teams to invest financially in more talent; especially those teams with the greatest market potential. Berri, Schmidt & Brook (2007) concluded that the relationship between team revenues and wins suggests that the perfect Competitive Balance would actually lower league revenues. At the same time, if the outcome is too random, the result of a game will be more like a gambling, therefore the spectators' population structure will change because more the games may attract more audience that are interested in gambling.

Second, given that optimal Competitive Balance exists, it is hard to find appropriate approaches to achieve the optimal levels of Competitive Balance. Operational rules that are promoted by the league commissioner have caused conflicts between league owners and players. Team owners have considerable autonomy in determining a player's salary and the location of the team (Winfield & Levin, 2007). Commissioners, who are supposed to represent all participants' interests, on the other hand, often behave under the influences of the league owners' wishes (Winfield & Levin, 2007; A. Zimbalist, 2003). These conflicts have led to work stoppages in MLB: the player's went on strike in 1972, 1980, 1981, 1990 and 1994-95; and were locked out by the owners in 1973, 1976 and 1990 (Zimbalist, 2003). These work stoppages led to shortened seasons and lower attendances, and hurt the interests of both the team owners and players.

As early as 1890, the owners instituted a reserve clause to prevent players from moving between teams, but players always want free labor markets. Policies such as salary caps/luxury taxes, college drafts, and monopoly exemptions are all applied in MLB in the name of promoting Competitive Balance. These policies are not so enjoyable for players because they actually decrease a player's salary and limit their freedom to provide their service to any team they wish.

Researchers have variety views about current operation rules, and sometimes these are conflicting. Berri et. al (2007) found that Competitive Balance appears to be dictated primarily by the underlying population of talent instead of league policies. By examining the economic structure of professional sports, El-Hodiri and Quirk (1971) found that operation rules do not make professional sport leagues exhibit any tendency toward Competitive Balances.

According to Coase theorem, in a world where everyone has perfect information and zero transaction costs, the allocation of resources in the economy will be efficient and will be unaffected by league rules regarding the initial impact of costs resulting from externalities

(Regan, 1972, in Medema & Zerbe, 2000; Maxcy,2002; Fishman, 2003; Schmidt & Berri, 2003; Lai, L., Ng, F. & Yung, P., 2008). That said, the allocation of property rights in a sports league would not impact the level of competition within a league, thus abandoning the reserve clause should not affect the distribution of players and not have a negative impact on Competitive Balance. Quirk's research shows that large market teams will dominate small market teams, and competitive imbalance will be invariant under a variety of institutional constraints designed to alter it (Vrooman, 1995).

Critics of free agency argue that Competitive Balance was more the result of dragging a good team down than bringing a bad team into contention (Vrooman, 1995); Palomino and Rigotti (2000) found that revenue share increases Competitive Balance but decreases incentives to win. According to Zimbalist (Zimbalist, 2003), revenue sharing, introduced to baseball in the name of reducing imbalance, actually contributes to baseball's imbalance. Beside revenue sharing, Zimbalist listed four other factors that increased imbalance: increased revenue inequality, more synergies from cross-ownership, the inversion of the draft's leveling role, and talent decompression with the league.

Third, researchers are not sure if sports leagues need more Competitive Balance. This is in part because researchers have a hard time finding empirical evidence about how uncertainty of outcome and thus Competitive Balance relate to game attendance or team revenue (Berri et al., 2007). Berri concluded that the economic significance of the relationship between Competitive Balance and attendance is not appealing; consequently, it is not clear whether spectators truly care about the level of Competitive Balance in a league (Berri et al., 2007). Utilizing Scully's measure of Competitive Balance, the NBA is the most unbalanced sport in comparison with the MLB, NHL, and NFL. However, as Zimbalist pointed out, the popularity of basketball and the rate of increase in revenue is the fastest growing among all these sports; but this cannot be explained by our arguments of Competitive Balance, which would favor the trends of diminishing spectators' support. Even as support is diminishing, it is hard to say if it is because of imbalance, the temporary retirement of Jordan, team and league pricing policies, or general macroeconomic conditions (Zimbalist, 2003; Quirk & Fort, 1992). Zimbalist also argued that Major League Baseball never reached Competitive Balance, and its survival proved that Competitive Balance is not a problem in baseball.

At the same time, due to the complexity introduced by cross ownership and operation intervention, profit claimed by the team owner does not always accurately reflect the team's achievements. For example, the owners value their ballplayers not only for what they produce on the field, but also for what they produce in terms of their media networks and other investments. The team owner does not treat his team as a standalone profit center, but rather as long-term profits of a larger entity (Zimbalist, 2003). For example, the Red Sox lost 13.7 million in 2001, but the former owners wanted to buy the team for over 700 million. The only explanation is either the team is not losing money, or there are substantial nonfinancial returns to the ownership (Zimbalist, 2003). Therefore, it is difficult to evaluate the relationship between Competitive Balance and team achievement by the profit claimed by team.

Moreover, good performance on the field does not always lead to increased revenues, as profit maximation and winning maximation can be conflicting goals for a team. In addition, the economic depression can offset income greatly. When people are busy with low paying second jobs or worry about unemployment, they might be less likely to enjoy the sport. For example, the 1931 Philadelphia Athletics won their third straight American league, but the attendance dropped by 100,000 compared with the previous year, likely due to the Great Depression.

1.4 Research Goals

Corresponding to the debating issues in the research of Competitive Balance, this research will focus on answering these questions: Do measures of Competitive Balance accurately reflect the amount of Competitive Balance? How does Competitive Balance change over time? Do spectators really care about Competitive Balance?

This research will try to answer these questions by examining the records of Major League Baseball games. Since the relationship between Competitive Balance and team league operation are similar for all league members, the results of this study can be generalized to other sports leagues as well.

1.5 Chapter Summary

Studies in psychology have shown that competition facilitates athletic performance, fulfills spectator curiosity and brings eustress to participants. Therefore, Competitive Balance, which leads to unpredictable sporting events, will bring in more spectators as stated by the theory of Uncertainty of Outcome in Sports Economics.

Given that Competitive Balance is an important issue in sports economics, not all researchers favor the current operational rules aimed at increasing Competitive Balance which are instituted by team owners and league commissioners. Moreover, researchers are currently debating whether or not competitive sports need more Competitive Balance and what the optimal amount of Competitive Balance is. By exploring the Competitive Balance in Major League Baseball (MLB) this dissertation aims to provide additional understanding of the debated issues.

The contents of this dissertation are organized as follows: Chapter 1 is an introduction of background information of studying Competitive Balance. After identifying three debating issues in this field, I set up three research objects corresponding to the three debating issues.

Chapter 2 is a literature review of existing research about Competitive Balance. Chapter 3 and 4 are devoted to the first research object: In Chapter3 I construct a set of Competitive Balance measures that aim to address different dimensions of Competitive Balance, and in Chapter 4 I examine the dimensions of the proposed measures. Chapter 5 is about the second research object--analyzing the Competitive Balance by displaying the patterns of Competitive Balance based on the proposed Competitive Balance measures. In Chapter 6 I address the third research object--Checking the hypothesis of uncertainty of outcome (UOH) by exploring the relationship between the Competitive Balance measures and game attendances. Chapter 7 concludes all previous chapters and discusses future research directions. A structure is also shown in Figure 1.1.

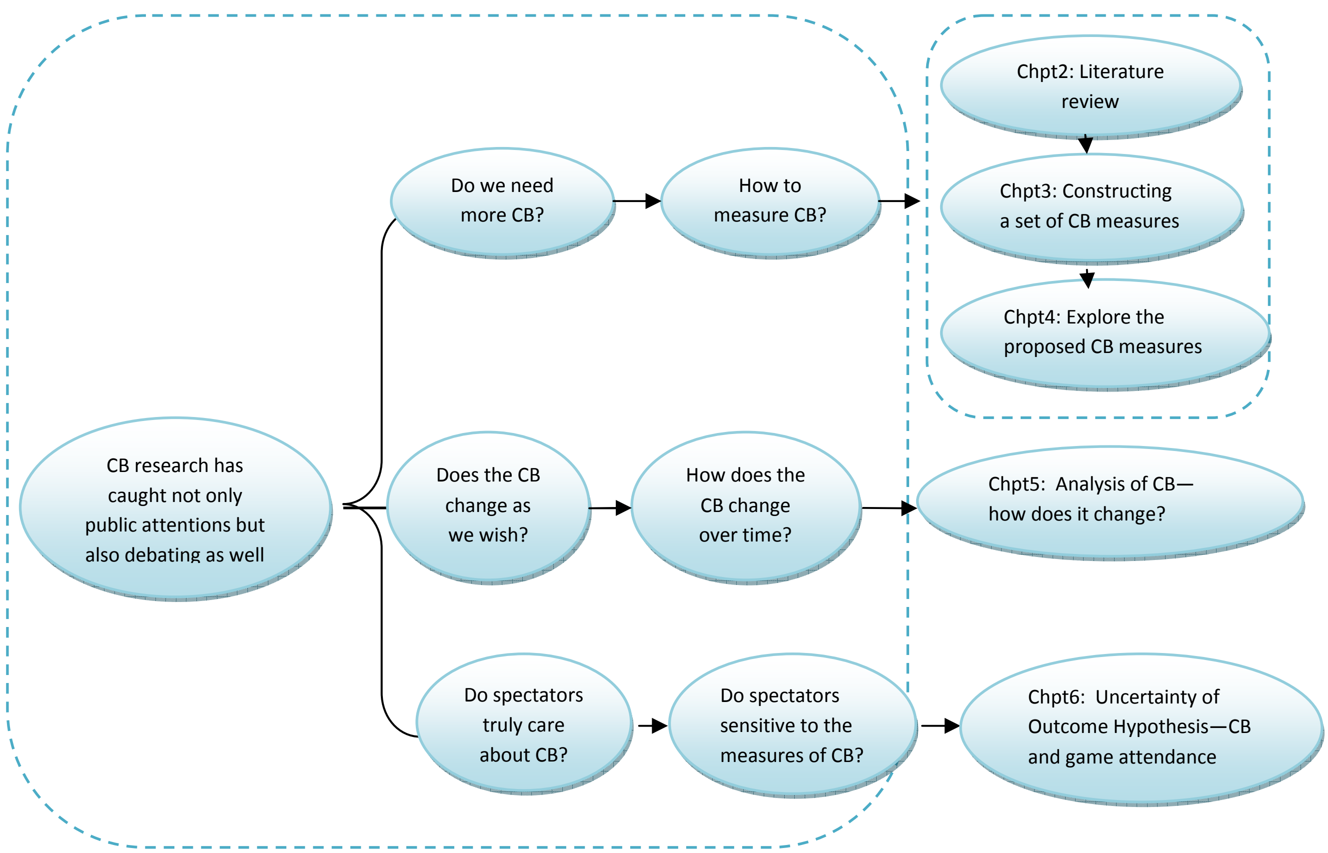


Figure 1.1 Structure map

CHAPTER 2: LITERATURE REVIEW

Knowing what to measure and how to measure it makes a complicated world much less so.

-Levitt & Dubner(2005)

2.1 Definition of Competitive Balance

In the economic sense, competitiveness is a comparative concept regarding the ability and performance of a firm. In sports research, scholars have being debating the appropriate definition and evaluation of Competitive Balance for a long time (Fort & Quirk, 2004; Fort & Maxcy, 2003; Fort, 2003; Humphreys, 2003a; Humphreys, 2003b;Kahane, 2003; Sanderson, 2002; Sanderson & Siegfried, 2003). Fort and Quirk (1992) defined Competitive Balance in a league as "a catch all term that refers to a number of different aspects of competition on the playing field. Essentially, there is more Competitive Balance within a league when there is more uncertainty of outcome in league games". In 2000, the commissioner's Blue-Ribbon Panel on baseball economics representing fans' interests was formed to investigate whether Baseball's current economic system has created a problem of Competitive Balance in the game(Schmidt, 2006). The Blue-Ribbon Panel defined Competitive Balance from the aspect of sufficient revenue redistribution: "…in the context of baseball, proper Competitive Balance should be understood to exist when there are no clubs chronically weak because of MLB's financial structural features. Proper Competitive Balance will not exist until every well-run club has a regularly recurring hope of reaching postseason play" (in Zimbalist, 2003, p35).

Humphreys conceptualized the Competitive Balance as:

"Competitive Balance describes the degree of uncertainty about the outcome of sporting events. Economists posit that uncertainty about the outcome of sporting events plays an important role in determining fans' interest in these events; Sporting events with a high degree of uncertainty of

outcome are said to be competitively balanced, and sporting events with a low degree of uncertainty of outcome are said to be competitively imbalanced" (Humphreys, 2005).

Vrooman (1996) identified three issues in Competitive Balance. According to Vrooman, *"There are three interrelated issues in the conceptualization of Competitive Balance: The dominance of large-market clubs, the closeness of league competition within the season, and the continuity of performance (superior or inferior) from season to season"*. In my opinion, the interrelated issues in the conceptualization are the different dimensions of Competitive Balance. Humphreys and Vrooman's concept will be used as fundamental bases of current research.

However, notice it not easy to identify a way to measure Competitive Balance from its definition, and it is not clear which indicators one should use to capture the dimensions of Competitive Balances. Haan, Koning, & Witteloostuijn (2008) proposed one measure for each of the three dimensions proposed by Vrooman, but made no connection to game attendances.

2.2 Two Approaches to Measure Competitive Balance

Competitive Balance is hard to assess directly in the real world. The common practices are to capture Competitive Balance via variables that may cause its change, I termed as input variables, or the variables changed with Competitive Balance, I termed as output variables. For example, the talent distribution, the coach experiences, and the financial supports are input variables that may lead to the changes of Competitive Balance. On the other hand, variables like number of spectators, the length of the game, and the scores of the game may change corresponding to Competitive Balance, and thus are output variables.

Due to the nature of team work involved, accessing the input variables for Competitive Balance can be very complex. For instance, players' personal records may change under influences of team leadership, star effects, audience reactions and game strategy. Therefore, it is difficult to capture the talent distribution with simple indicators such as successful shots, game score or salary level (Bodvarsson & Brastow, 1998; Depken, 2000; Frick, B, Prinz, J. & Winkelmann, 2003; MacDonald & Reynolds, 1994).

Thus, this study will assess Competitive Balance with output variables, to be specific, how the uncertainty of outcome affects the number of spectators. The limitation inherent with this approach lies in that Competitive Balance is not the only reason fans attend games. The star effect, contest significance, contest legitimacy (Gerrard, 2006), time of the game, market base, weather and location (Lee & Fort, 2005) may all affect to game attendance among other factors. Further, the relationship between Competitive Balance and game attendance may be nonlinear due to the limitation of the facility and the spectator market size. Researchers also noticed that Competitive Balance is not a factor where “more is better” (A. S. Zimbalist, 2002). It is possible that too much Competitive Balance may harm attendance because spectators may feel bored if they know that the outcomes are completely random. Last but not least, there are mediators that intervene in the relationship between Competitive Balance and attendance. For example, attendance may drop if an increase in Competitive Balance is accompanied with an increase in ticket price. For example, increasing the salaries of a weaker team leads to better Competitive Balance, but the salary increase may also lead to higher ticket prices, which may potentially decrease the market demands.

Given the limitations of the output approach a strength it has is that outcome variables are relatively easy to access when compared to the input approach. For example, the length of the game, the attendance, the market size, and the number of starts in the team are all object numbers and they have been accurately recorded. Therefore, this research will focus mostly on the output approach.

2.3 Existing Measures

Researchers developed many measures of Competitive Balance. The existing measures will be discussed in this section based on Vrooman's concept of Competitive Balance. That is, to measure the Competitive Balance by accessing the closeness, concentration, and continuity dimension in sport teams.

2.3.1 Measures of closeness.

Measures of Competitive Balance that focus on the closeness of league competition within a season can be grouped into two categories: the most popular measures Competitive Balance by winning percentage; the alternative measures Competitive Balance by play-off appearances, which capture how far the team advances in the post season.

When using winning percentages to measure Competitive Balance, sports economists typically compare the actual distribution of teams' winning percentages to an ideal spread of winning percentages, which is based on each team winning 50% of its games. The advantage of using this measure is that it is easy to understand and its calculation is straight forward. However, the information in winning percentage is limited as well. For example, is a team that wins 60% of its games an extraordinary competitive team or a just a good team? To answer this question, you may want to compare this team with the average team. But the simple average of winning percentage tells us nothing about the Competitive Balance of a league, because it is always 50%.

The standard deviation of winning percentage is a more advanced measure, also known as the Noll-Scully measure (Quirk & Fort, 1992), developed by economists Roger Noll and Gerald Scully separately, which compares the teams' actual standard deviation of winning percentage with the ideal standard deviation of winning percentages, which is 0.5 over the square root of the number of games each team plays (M. Leeds & von Allmen, 2002,p252; Vrooman, 1995):

$$CB_{ideal}=\sigma_{ideal}\,(Y)=\frac{0.5}{\sqrt{N}}$$

$$CB_{y,actual}=\sigma_{actual}\,(Y)=\sqrt{\frac{\sum_{i=1}^{N}(wp_i-0.5)^2}{N}}$$

**Where N is the number of the games played by a league Y.*

This measure shows by how much each team's winning percentage differs from the average winning percentage. If the Noll-Scully measure is equal to one, then the league has the ideal standard deviation; if the Noll-Scully measure is 1.5, then the league's standardized deviation is 1.5 times larger than the idealized standard deviation.

2.3.2 Measures of concentration.

Several Competitive Balance measures focus on the concentration of game results, such as the Herfindahl-Hirschman Index (HHI), the range of winning percentages or the excess tails of winning percentages. These concentration-measures describe the distribution of Competitive Balance across all teams in a league.

The HHI measures the market share of each team: $\mathrm{HHI}=\sum_{\mathrm{i}}^{\mathrm{N}}(\mathrm{MS_i})^2$; where $\mathrm{MS_i}$ is the percentage of the total wins of i_{th} team and N is the total number of teams (Craig A.Depken, 1999). The lower bound of HHI is $\frac{1}{N}$, because as long as there is a game, there is a winner; and the upper bond is 1, it happens when all wins belongs to one team. HHI is a non-linear transformation of winning percentage and the relationship between the two can be written as:

$$\sigma=\sqrt{\frac{NG^2}{4}\,(HHI-\frac{1}{N})},$$

Where G is the number of games played by all teams and σ is the standard deviation of winning percentage across teams (Depken, 1999).

The Gini coefficient (Lambert 1993, in M. B. Schmidt & Berri, 2001) is another measure of concentration, which is defined as:

$$G_i=(1+\frac{1}{N_i})-\frac{2}{\mu_{xi}{N_i}^2}*(x_{N,i}+2x_{N-1,i}+3x_{N-2,i+\cdots+N*x_{1,i}}) \text{ ,}$$

Where N is the number of teams, $x_{N,i}$ is the winning percentage of team N. μ_{xi} is the average winning percentage of Team i.

The range of winning percentages (highest to lowest win percentages) measures the distance between maximum and minimum winning percentage among the teams: a league with a wider range has less Competitive Balance than a league with a narrower winning percentage range. Excess tail frequencies measures how often extreme winning percentages occur in the league: if the distribution of winning percentage is skewered to left (or right), then there are many teams who have low (or high) winning percentages. If the right or left tail is very long, then there are extremely high or low winning percentages in the league. In both situations, the Competitive Balance is less ideal than an evenly distributed league.

Researchers that use concentration to measure Competitive Balance often focus on post season appearances. Because post season teams often play in large markets, and only one third of the teams in the league advance into the playoffs, these post season teams may not fully represent the distribution of Competitive Balance in a league. We do not have enough theoretical support to assume that the Competitive Balance for post season teams is the same as regular season teams. Neither can we prove that Competitive Balance of large market teams can represents all other teams in a league. Therefore, playoffs data and its measures are not the focus of this study.

2.3.3 Continuity of performance across seasons.

An influential trend in measuring Competitive Balance is to consider the fluctuations over time. Such time elements include baseball seasons, the number of the games, etc.. Empirical evidence shows that the more seasons included in the analysis, the more likely the variance due to time will overweight the result (Eckard, 2001b) . Existing measures in this branch include winning percentages across teams during a year, standard deviation over time (Zimbalist, 2003), and Competitive Balance Ratio (CBR) (Humphreys, 2002). Among all these measures, the CBR is one of the most frequently cited measures. CBR is defined as the average time variation in won-loss percentage for teams in the league by the average variation in won-loss percentages across seasons (B. R. Humphreys, 2002). As we can see from the function, CBR is a portion of the winning percentage of between season variances over within season variances.

$$CBR = c * \frac{Between\ Season\ Variance}{Within\ Season\ Variance}.$$

This measure contains information about the level of Competitive Balance and year-to-year fluctuations in team performance (Humphreys, 2002). Similar research can be found in Maxcy and Mondello (2006), whose measure also captures Competitive Balances over time. Schmidt and Berri (2002) used time series techniques and found that aggregated demands decrease as the Competitive Balance decreases in a league. However, game attendances decreased as the Competitive Balance increases in given season. All the existing measures have their possible limitations, and next section will discuss these limitations and propose the use of new measures.

2.4 A Set of Proposed Competitive Balance Measures

2.4.1 Areas for improving existing measures.

Standard deviation of winning percentage and its related measures have become well accepted measures of Competitive Balance. Many economists like to use this measure together with a regression model to infer the changes in Competitive Balance, and use attendance to evaluate the measure (Dobson & Goddard, 2001; Donihue, Findlay, & Newberry, 2007; Forrest, Beaumont, Goddard, & Simmons, 2005b). Researchers from math or statistics use Bayesian models together with distribution assumptions such as Poisson (for scores, ranks), binomial (for win/loss) to predict the probabilities of game outcome, and simulation methods such as Markov chains (Bukiet, Harold, & Palacios, 1997) are used to generate data and infer the effect of changes in Competitive Balance of a league.

Two pitfalls can be identified in existing research using winning percentages to measure Competitive Balance. First, this method is vulnerable to possible schedule bias. For example, the number of games played among matched pairs of teams in a league differs systematically. The matched pair of teams refers to two teams that are designed to play with each other. For example, the Baltimore Orioles (BAL) & the New York Yankees (NYA) played 18 games in 2008, and as such BAL&NYA is considered one distinct matched pair, hereafter referred to as matched pair. In many leagues, teams only play a subset of the other teams in another league. For instance in MLB, each team in the National League (NL) is scheduled to play against three to six teams in the American League (AL), and these interleague opponents are differ across teams. Therefore, a NL team may play against a subset of strong teams from the AL, or the opposite. The team that is scheduled to play nine times against the weakest team in the AL may have a better winning percentage than the team scheduled to play against the weakest team three times. Similarly, any team that scheduled to play against the weakest teams six times may have a better winning percentage than the team scheduled to play against the strongest team six times. Thus, the game schedule introduces the possibility of overweighed losses or overweighed wins, for that the winning percentages changes systematical by game schedule.

Second, the idealized winning percentages derived from the assumption that each team's record is independent, that league competition is perfectly balanced, and that each game result will be determined as randomly flips of a coin. These assumptions need further examination, for that research in winning percentages focus on the dichotomized game results (win or loss) of all games played in a season, and these results may contain non-random components. An example of a non-random component would be the location of the game, is it an away game or home game? Home field advantage has been shown to be a significant predictor of the game results (Forrest et al., 2005b; Meehan, Nelson, & Richardson, 2007; Stefani, 2008). In addition, it is hard to defend the assumption that all game results are completely independent. As teams in MLB are scheduled to play against the same opponent at least three times, it is reasonable to suspect the game results are dependent for the same matched pairs of teams, or the 162 games played by the same team. Moreover, the MLB schedule is an unbalanced design, because the number of games played by each team is different, some teams played more than 162 games in a season when a tie breaker was needed to determine the rank of the team, and some teams played less than 162 games in a season when no additional games were necessary to determine the rank of the team. Ignoring these structured dependencies and unbalances may cause biased conclusions when using winning percentages to accessing Competitive Balance.

At the same time, using attendance to evaluate the effectiveness of winning percentage as a measure of Competitive Balance can be difficult without controlling for variation in other factors, because Competitive Balance is not the only reason fans attend games. A spectator's interests and awareness of a team's real-time winning percentage may be limited. Instead, spectators may be more familiar with their favorite team's history or favorite stars' statistics.

Score differences can reflect some Competitive Balance information that is neglected by winning percentages. For example, a game score of 10 to 1 is different from a game score of 6 to 5, and it is reasonable to assume the second game is between more balanced teams. Sometimes, score differential is used as a tie breaker. However, score difference is vulnerable to point shaving and team strategy changes. As baseball player Mark Grace once said, "if you are not cheating, you're not trying"(Levitt & Dubner, 2005). After all, the final ranking of a team rarely depends on score differences but its winning percentage. Teams may put less effort toward scoring after the result is locked. As a matter of fact, many games in the regular season do not go a full nine innings, but are still considered a complete game because the result is settled.

Some research calculates a HHI or a Gini coefficient for the regular season to investigate how Competitive Balance varies over time. By doing so, they facing the same pitfalls as using winning percentages, as the HHI is a non-linear transformation of winning percentages. At the same time, research in other sports shows that league rank does not reflect teams' past performance, which is an important factor in predicting game outcome (McHale & Davies, 2008).

2.4.2 A three-dimensional indicator of Competitive Balance.

As such, I propose a new framework of Competitive Balance measures. My theoretical guidance is based on Vrooman's concept about Competitive Balance.

In order to do that, I propose a set of indicators that will address

- The closeness of the games. It should address the difference between:

  Scenario 1: Team A defeats team B by a score of 10-1, and

Scenario 2: Team A defeats team B by a score of 6-5;

Or

Scenario 1: Team A played twice against team B. First time, Team A defeats Team B by a score of 10-1, and second time, Team B defeats Team A by a score of 10-1 , and

Scenario 2: Team A defeats Team B twice by scores of 10-1 and 10-1.

- The change for upset games. This corresponds to the dominance of the team with a higher winning percentage ranking over its opponent with a lower winning percentage ranking. If the higher ranking team wins all the games played with the lower ranking opponent, then the higher ranking team is quite dominate in this matched pair. However, if the lower ranking team wins some of its games with the higher ranking team, which are also referred to as upset games, then the dominance of one team over another is not strong in this matched pair. This indicator is derived from Vrooman's concept about 'the dominance of large market clubs to other clubs' in MLB; however, in current research, the dominance is no longer 'large market clubs', but 'teams with higher winning percentage in a matched pair'. The measure should address the concentration of the regular season, and address the possible imbalanced design in the schedule.
- The consistency of play. This includes two types of consistency: first is the consistency in overall rank in the league based on the rank differences of the teams across seasons. Second, how many upset games the team wins. If Team A ranks 15th in the league, the measure should address the difference between:

Scenario 1: Team A beats the team ranked 1st 2-0 and

Scenario 2: Team A beats the team ranked 30th by 2-0.

The next chapter will describe how the indicators are constructed in detail.

2.5 Chapter Summary

Conceptually, Humphreys defined Competitive Balance as a description of the degree of uncertainty about the outcome of sporting events. Structurally, Vrooman states Competitive Balance includes three aspects: dominance, closeness, and continuity of performance. Technically, this research will use the structure proposed by Vrooman, and measure Competitive Balance from the aspects of uncertainty outcome.

After reviewing existing measures of Competitive Balance in the literature, this chapter identified the possible improvements of the measures of Competitive Balance, and thus proposes a new set of measures. These measures focus on the variables that change with the Competitive Balance, and allow the researchers to examine the dependence among game results, as well as the information in score differences, which has been neglected by winning percentages. Therefore, the new set of measures may reveal some insight into Competitive Balance in MLB, and thus provide a clear observation of the issues under debate.

CHAPTER 3: CONSTRUCTING A SET OF COMPETITIVE BALANCE MEASURES

As mentioned above, the goals of this research are to explore how Competitive Balance changes over time and understand how Competitive Balance affects game attendance. To achieve these goals, one fundamental question needs to be answered first: how should we measure Competitive Balance in a league? This chapter focuses on developing the proposed Competitive Balance measures based on the game structure and data in Major League Baseball (MLB).

3.1 Data Description

The population studied in this research is the Major League Baseball (MLB) in North America. Baseball is the oldest professional sport in the North America, and the records in the MLB date back as early as 1871. Modern MLB contains thirty teams from all over North America, and thus is an ideal setting for exploring Competitive Balance. The MLB competition consists of regular season games, all-star games, post-season playoff games, and the World Series. This research will focus only on the regular season games.

3.1.1 Regular season games.

Major League Baseball's regular season starts in late March or early April each year. Prior to 1969, MLB was comprised of two leagues, the American League (AL) and National League (NL). The NL, world's oldest extant professional team sports league, contained ten teams in 1968 as did the AL. Beginning in 1969, divisions were introduced to MLB when the NL and the AL each expanded to twelve teams, and split into two divisions per league based on location. In 1994, both leagues expanded again and further split into three divisions: East, West and Central. Each season, the teams play half of their games in their host city as the home team, and the rest of their games at their opponent's host city as the visiting or away team.

Occasionally, games are played in a third city, in these cases, one of the teams will be designated as the home team, and another designated as the visiting team.

Since the three division era, each AL team has 18-19 opponents, which includes 13 opponents in the same league and 5-6 opponents from the NL. For the NL teams, each of them has 19-21 opponents, 15 from the NL and four or five teams from the AL. As for inter-league, play opponents are more likely to be spatially close to each other, for example, the AL central division teams are more likely to play against the NL central division teams.

If the teams in a game, referred as matched pair of teams in this research, are from the same league and same division, they are scheduled to play 15 to 19 games against each other. In the AL, there are 26$(C_5^2 + C_5^2 + C_4^2)$ possible matched pairs in same division, and in the NL 35$(C_5^2 + C_5^2 + C_6^2)$ pairs. The scheduled games played within a division summed up to 1073 in 2007.

If the matched pair of teams are from different divisions but the same league, they play six to ten games in the regular season; the NL has 170 ([5*(6+5)+5*(6+5)+6*(5+5)]/2) matched pairs, and the AL has 130 ([5*(4+5)+5*(4+5)+4*(5+5)]/2) matched pairs.

For the matched pairs consisting of inter league teams, they play three or six games against each other. Ten matched pairs are guaranteed to play six games against each year (Wikipedia contributors, July 2008). Of these, four pairs are from the Central division, three pairs are from the Eastern division, two pairs are from the Western division, and one pair is a West & Central combination.

At the end of the regular season, each team has a winning percentage calculated from all the games they have played, and the top four teams are selected from each league to advance to the post season-divisional and championship series and then to the World Series. If two or more teams have same winning percentage, game records between the teams are used to determine the winner. For example, in 2006 the San Diego Padres won the division championship over the Los Angeles Dodgers, based on their 13 wins to 5 loses records in the season against Los Angeles Dodgers. Both teams finished the regular season with the same winning percentage.

Since 1994, each season should have 2430 regular season games, if all games are played (Wikipedia contributors, July 2008). Yet the number of games played in each season is not always 2430. Some seasons had 2431 games, due to a playoff game. A playoff game is used to determine the fourth team who qualifies for the post season series, in addition to the division champions from the three divisions. The forth team is also called the "wild card" team. Some seasons have less than 2430 games; for example, the schedules for 1995 were reduced from 162 to 144, due to the games cancelled during the strike that took place in 1994 and 1995. Also, when postponed games have no influence on the teams' division standings or wild card qualification, the games are often not played. Incomplete games were not counted in the game played by the team. Thus, beside the season 1995, other seasons have had less than 2430 games as well. In fact, the 162 game schedule for each team has dated back as early as 1962; however, both leagues have changed their schedule setting several times during 1962-1994.

By the end of September or early October, the first round of the playoffs begin with the American League division series (ALDS) and National League division series (NLDS). The four top teams in each league play each other, and the two teams who win 3 out of 5 games (best-of-five) will advance to the League Championship Series (NLCS or ALCS). The teams who win the best-of-seven games (since 1985) in the LCS will advance to the World Series. The division series, LCS and World Series together are considered post season games. Because the post season and All Star game are not of interest for this research, they are not included in the data set analyzed in this research.

The regular season data comes from project Retrosheet (http://www. retrosheet. org/boxesetc/index. html), and consists of play-by-play records from the 1871 season to the 2008 season. In each data set, the number of records varies according to the schedule of the season. The dataset only contains the regular games played by each team, it does not include the post season games that were played after the first Sunday in October (or the last Sunday in September). The data consists of three types of statistics: 1) Schedule statistics, which include the time, day/night of the game played, game numbers, etc. 2) Performance statistics, which includes batting statistics, base running statistics, pitching statistics, fielding statistics, and the scores for home team and visiting team; and 3) Demographic statistics, which include the manager's name, players' names and positions, ball park location (ID), among other variables.

3.1.2 Division and city profile information.

Since 1994, both the AL and the NL expanded and each split into three divisions. The division information is essential to understanding the labor market and spectators' attendance, and therefore this data was integrated in the research. The teams' host city population and income was also collected from Census Bureau (http://www. census. gov/popest/datasets. html) , and will be used to understand the game attendance.

3.2 Secondary Data Analysis

From the above description, it is clear that this research will use secondary data analysis, which is defined as the analysis of existing data sets (Sales & Lichtenwalter, 2006). Unlike survey data or simulation data, which is collected from a sample population or generated by the researcher, secondary data comes from a third party, and secondary data researchers have no control over the data structure design or data collection. There are both pros and cons associated with the use of secondary data analysis.

3.2.1 Benefits of using secondary data analysis.

Acquiring secondary data is relatively easy compared to primary data. In addition, the quality and availability of secondary data is improving as technology develops. More organizations are able to provide their data resources to the public and social science data archives are available all over the world via governments, statistic bureaus, scholarly journals, research institutions, universities, libraries and internet. For instance, the U. S. Census Bureau conducts nation-wild surveys periodically and generates comprehensive social data information. To collect data like this through individual researchers would be almost impossible due to the limitations of time and funding. Other organizations do not conduct surveys, but offer various archives of secondary data, such as the Inter-University Consortium for Political & Social Research (ICPSR), Indiana Political Data Archive and Laboratory, Connecticut Social Science Data Archive and others. Many of the datasets are even available online or free in the public library, such as the data used in this research. With the growing concerns about privacy, survey human subjects become more troublesome. As such, the availability of existing data has provided researchers with good data quality, a wild range of selections, and at the same time circumvents the problem associated with the time and financial constraints of data collection.

Moreover, popular secondary data resources make the comparison of research results possible. Sharing knowledge among disciplines is a desirable goal of researchers. For example, the data curation projects conducted by the graduate school of library & information science aim to develop best practices materials for the Library and Information Science and Museum Communities, including a smooth transactions of data and knowledge. Sampling procedures used by individual researchers are often constrained by the resources available to them. For instance the definitions of certain terms may vary among researchers due to their knowledge structures, making comparisons across studies meaningless. By utilizing secondary data resources, even engaged in independent projects, researchers have the advantages of using similar data definitions and sampling frames, and therefore comparing the results to different studies is more applicable. Meanwhile, with the growing familiarity of a dataset, and the expanded knowledge of the research associated with the same dataset, researchers are able to explore more sophisticated and creative methods of research designs with the secondary data.

Last but not least, the secondary data analysis approach can accommodate various research designs that are difficult to implement otherwise. An example is a longitudinal research design, or cross national comparison. In terms of longitudinal research, trend analysis, panel analysis, event history analysis, and time series analysis are frequently used research designs. These designs either require cross sectional data, sometime even cross national, or years of observations on each individual subject. If one does not have the means to collect data over several years of international travels, secondary data analysis may be a good option.

Secondary data is also relevant in meta-analysis, which is defined as a research design that combines the results of several studies that address a set of related research hypotheses. One may argue that meta-analysis is not secondary data analysis (Kiecolt & Nathan, 1985); regardless it is a type of research can be done only using existing data sources.

3.2.2 The challenges of applying secondary data analysis.

The most prominent problem that needs to be addressed here is the validity of secondary data analysis. Validity refers to the extent to which data gives a true measure or description of social reality, or the degree to which a study supports the intended conclusion drawn from the results (Wikipedia contributors, 25 May 2008). Campbell (1965) further divided validity into four aspects: internal validity, external validity, statistical conclusion validity, and construct validity. Internal validity is affected by the flaws intrinsic with the study itself such as because of survey instrument. External validity refers to the extent that one can generate the research findings to a population beyond the studies population, or generating to different research settings. Unbiased samples and large enough sample sizes are essential to the statistical conclusions of validity, and the quality of the measures or scale design are the major concerns for construct validity.

Because users of secondary data do not participate in the construction of the instrument design, and the measures and scales of the existing data are created with different purposes, construct validity may diminish through the usage of secondary data analysis. For example, if one wants to analyze unemployment among athletes, one needs to examine the concepts of unemployment over time. Researchers should also be cautious about the external validity of the second hand data. For example, when one examines the labor price of baseball, the data collected in 1990 may not provide enough information to assess the situations in 2000. Often, the sampling frame of the existing data may not match the target population of the new study, which can lead to validity concerns. For example, data collected in minor league baseball may yield biased information about American baseball players' salaries. Sometime, researchers may find conflicting data in an existing dataset, which indicates the internal validity is problematic. In this case, one should find other data recourses.

In order to get appropriate data structure that is suitable for the research questions developed using a different theoretical frame work, one needs to combine data from a variety of secondary data resources. Consequently, the researcher is now faced with the difficulty of merging the data. The definitions or measures from different sources may be incompatible, the data sets may be inconsistent and overlap or the meanings of the data field might be unclear and not well documented.

Researchers also face constrains given the quality of secondary data. First, in terms of data operation, researchers only have the freedom to condense or simplify the existing data set, but not get into more details. If the dataset codes participants' education as “no education”, “high school”, “college”, then researchers can only analyze the education groups, but are unable to access subgroup information such as “associate degree”. Also, when data sets have errors, researchers have no means to rewind the procedure and correct the problems. Lastly, research using secondary data analysis is constrained by the existing data, if the framework or the concept being applied is too new, secondary data often not available.

As a result, researchers who plan to use secondary data analysis need to follow a procedure to overcome the drawbacks of secondary data analysis, that is: define the research question, identify possible research designs, locate a trustworthy data source, verify the existing data, and then merge, clean or transform the data to fit ones research needs.

3.2.3 Secondary data analysis and current research.

As described in the beginning of this chapter, the secondary data used in this research comes from multiple sources. The data from Retrosheet is a panel data set for Major League Baseball (MLB) spanning over one hundred years. The data set is trustworthy, because its records are consistent with the official set provided by MLB and other sports data resources as well.

Due to the nature of this comprehensive, rich, precise and easy accessible data, Major League Baseball records have been used far beyond sports division. Researchers using sports data on their studies come from the fields of economics, social science, education, statistics, law, business, tourism and health studies. Retrosheet data provide complete documentation for the definition of data fields. In addition, it is easy to manage and is familiar to sports data users. The wild citation of this data source makes it possible to compare research across disciplines. In fact, when examining the citations of the research papers using baseball data, one can find references from various fields.

Moreover, the population of the secondary data is an exact match with the questions of this study. This study focuses on the measure of Competitive Balance in MLB, and the data resources are all about MLB. Thus, conclusions draw from the data set match the research question needs.

More important, the research questions asked in this dissertation require a longitudinal research design. Cross sectional data for each team's performance under different conditions, as well as team performance across seasons, are essential for a design that aims to answer questions such as how Competitive Balance changes over time. Also, the research design requires an exploration of the covariance structure among variables, and the rich sample size available in this setting makes this goal easy to achieve. Therefore, secondary data analysis is the best choice for this research.

Like other types of data sources, this dataset also has limitations. One limitation is that researcher has no means to get additional variables that may improve the research. For example, television coverage and contract values, the club property, and the style of the management (profit or win oriented, etc). Another limitation lies in the price recorded in the database. The data set recorded one ticket price for each game, whereas in reality games have several ticket prices, with the expensive tickets often selling out faster. So the unique prices may not lead to the right conclusions being drawn. Therefore, researchers should be careful when exploring questions related to price.

3.3 Strategies Available to Analyze Longitudinal Data

There are four analytic strategies to choose from when using longitudinal data, as suggested by UCLA's statistic consulting group (UCLA,2010): Regression, repeated measure of ANOVA, Mixed model ANOVA, and Multilevel models. The selection of the appropriate method depends on the specific research question, theoretical assumptions, and data structure.

Regression is flexible in that some teams have more records than other teams; if the measures are taken in multiple time points, regression models tolerate data acquired with unequal spacing of time schedules. Regression assumes that there is no covariance among measures taking in different conditions. However, this assumption is problematic when using raw game data, because intuitively the outcomes of the games between same pair of teams are very likely to be correlated, same for the game played by same team with different opponents. One can condense the multiple observations of one team into one observation, but it should not be the best choice.

Traditional repeated measure of ANOVA assumes that all teams have the same number of waves of data, which is not true in this dataset. For data measured at different time points, ANOVA also assumes all teams are measured at the same schedule, which is not exactly true for the raw game data. Lastly, ANOVA assumes two types of correlation structures between the measures taken at different time schedules: compound symmetric or unstructured. Compound symmetric assumes all between subject covariance are the same, and all the within subject variances are the same. Unstructured covariance matrix has no assumptions for variance and covariance, but it will affect the power of the model due to greatly increased number of parameters need to be estimated.

Mixed models ANOVA allows some teams have more data waves than others; the repeated statement assumes measures are taken at the same time. In addition to the covariance structure provide by repeated ANOVA, mixed models ANOVA offers more choices of covariance structures, such as autocorrelation, which assumes measures taken in close time points are more correlated with measures taken in large time span. The assumption adds one more parameter to the model compared with compound symmetric assumption.

Multilevel modeling accommodates the fact that measures taken at different time points are correlated with each other, and it allows each team to have a different number of observations, in different schedules. Compared with mixed models, it has more choices relating to the covariance matrix, such as Autoregressive Heterogeneous Variances, which allows variances to change over time.

3.4 The Many Dimensions of Competitive Balance

3.4.1 Game importance as measures of team closeness.

If two games have exactly the same scores, and same location, can we say the two games have the same Competitive Balance? I argue that depends on the importance of the game as the closeness of the team may be different.

Because the selection of play-off teams are based on winning percentage in a division, teams most direct competitors are other teams in the same league and division. It is reasonable to assume that a team may try harder if the game is closely related to its standing in the division, thus have better performance than in games irrelevant to its standing.

Previous Competitive Balance measures treat all games as equal and assume the competitiveness of teams is the same across time. In this research, I propose to measure the importance of each game. The importance measure is calculated based on the winning percentage of the team and its division members' winning percentage at the same time. Because the game results are easily available, all teams have accurate information about the importance of the game. The importance of game for each team is different, and it changes game by game.

I denote importance of the game at time t for Team i as imp_{it}, it is the winning percentage difference between Team i and division head and wild card candidates. The wild card candidate is the team that has the largest winning percentage in the league other than the division leaders.

$$imp_{i,t} = \begin{cases} 1-(wp_{i,t}-wp_{div2nd,t})\text{ , } when\ team\ i\ is\ division\ head\ at\ time\ t \\ 1-\min\left(wp_{div1st,t}-wp_i,\ \ wp_{i,t}-wp_{wc2nd,t}\right)\text{ , } when\ team\ i\ is\ the\ wild\ card\ candidate \\ 1-\min\left(wp_{div1st,t}-wp_i,\ \ wp_{i,t}-wp_{wc1st,t}\right),\quad otherwise \end{cases}$$

(p1.1)

In the equation (p1.1)

$wp_{1st,t}$ is the higest winning percentage of the division at game g.

$wp_{2nd,t}$ is the second largest winning percentage of the division at game g.

$wp_{wc1st,t}$ is the largest winning percentage of the league other than division heads.

$wp_{wc2nd,t}$ is the largest winning percentage of the league other than division heads and wild card candidates.

The first measure of game importance reflects the difference in importance for the two teams in the matched pair:

$$impD_{ij,g}= imp_{i,g}-imp_{j,g}$$

This measure aims to capture the difference in the following situation:

Team A beats Team B in a normal game: 10:5;

Team A beats Team B in a wild card competition game: 10:5;

The second measure of game importance captures the total importance of the game for both teams:

$$impS_{ij,g} = imp_{i,g} + imp_{j,g}$$

This measure aims to capture the difference in the following situation:

The Boston Red Sox beats the Minnesota Twins with 1:2, and this game gains the Red Sox the wild card;

The Boston Red Sox beats the Minnesota Twins with 1:2, and this game has no influence on neither the Red Sox's winning percentage rank nor the Twins.

As the wild card did not start until 1994, there is no importance value calculated before this date. For games before 1994, the winning percentage by the time of the game is used to indicate the game importance:

$$impS_{ij,t} = wp_{i,t} + wp_{j,t}$$

$$impD_{ij,t} = wp_{i,t}, - wp_{j,t}$$

3.4.2 More indicators of team closeness.

Traditional research on Competitive Balance focuses only on the result of a game: win or lose. However, I suspect that the variation in winning percentage does not fully reflect closeness between the opponents. Winning percentage is a summary statistic that filters out the opponent information and treats all the wins the same. Competitive Balance measures derived from winning percentages do not typically consider information such as the scores of the game, the length of the game, nor the importance of the game. In this research, I propose to use following three measures to indicate the closeness of the game.

- Score differences

Do the Competitive Balance measures consider the differences of the wins? When using traditional measures such as winning percentages and the measure derived from winning percentages, the answer to this question is "No", because previous measures treat all the wins the same.

$CLO1_{ij,t}$ is a measure aimed at capturing the Competitive Balance reflected by score differences within each game, in other words, the magnitude of the wins. This type differences has been ignored when using winning percentages as a measure of Competitive Balance.

$CLO1_{ij,g}$=Score difference between home team and visting team at game g $= s_{ig} - s_{jg}$ (p1.2)

In equation (p1.2), s_{ig} is Team i's score at game g and s_{jg} is Team j's score at game g. And S_{ig}-S_{jg} is the score difference of Team i and Team j at game g. In addition to game results, which are captured by winning percentage, CLO1 identifies the differences between a close game and a game that one team wins a lot. For example, the following two games have different CLO1 values:

The Chicago Cubs beats the St. Louis Cardinals three times, with scores 6-5, 11-10 and 8-7, respectively;

The Chicago Cubs beats the Pittsburgh Pirates three times, with scores 10-1, 15-3, 10-2, respectively;

- *Score sum*

If the two games have the same score differences, can we say the two games have the same Competitive Balances? I propose to treat wins differently on one more level, because a game result such as 25:20 may indicate a different balance level than 6:1, yet the score differences are equal in both games.

$CLOA_{ij,g}$ is a measure that captures the balance reflected by score differences between the games, and this difference is not captured by winning percentages.

$$CLOA_{ij,g} = \text{sum of scores of Team i and Team j} = \frac{s_{ig}+s_{jg}}{length_{ij,g}} \quad (p1.3)$$

It will capture the difference that is ignored by winning percentage in this example:

The Boston Red Sox beats The New York Yankees three times, with scores 2:1, 3:2 and 4:3, in 100 minutes, respectively;

The Boston Red Sox beats The Baltimore Orioles three times, with scores 11-10, 12-11, 22-21, in 100 minutes respectively;

CLO1 and CLOA together capture the association between the games. McHale and Davies (2008) examined the international football (FIFA) game scores, and found that scores of the opponents may not be correlated, which may be a sign of relatively balanced competition. They also found that some opponents' scores were correlated, where one teams higher score indicated another team's lower score. That, he concludes, may indicate a dominance of one team over another.

Closeness of the teams could also be revealed by the team's previous season winning percentage. This set of measures aimed at capturing the unexpected results of a pair, based on the pair's previous year or current year's winning percentages.

$$DOM2_{ij,year} = wp_{i,year-1} - wp_{j,year-1} \quad (p1.4.1)$$

$$DOM3_{ij,year-1} = wp_{i,year-1} + wp_{j,year-1} \quad (p1.4.2)$$

In above equations $\mathrm{wp}_{\mathrm{i,year-1}}$ is Team i's previous year's winning percentage, and $\mathrm{wp}_{\mathrm{i,t}}$ is Team i's winning percentage at team t. Team i is the assigned home team.

This set of measures aims at capturing the differences between competitor's previous year's winning percentage and ad hoc winning percentage. The ad hoc winning percentage is calculated based on the total game played by time t and total wins by time t.

Both competitors have very high rank (high previous winning percentages) in previous season verse one competitor ranked high verse both ranked low;

3.4.3 Indicator of upset games.

When using HHI to measure Competitive Balance, team dominance is evaluated by examining how many teams have the chance to win league championships. This approach focuses on playoffs and it does not fully capture competitive concentration in the whole league. Team i's winning percentage is an accumulated result, it does not identify where the winning percentage was determined. Some teams accumulate their wins by winning a lot games against a few opponents, whereas some teams accumulated their wins by winning a few games with a lot opponents. I assume that gathering wins from more opponents indicates less dominance than winning the same number of games from fewer opponents. It is especially true when the higher rank teams win all the games against lower ranked teams. I refer to this as rank transformability. To be specific, I propose to measure team dominance by counting the number of wins that the team should win according to its rank, and the number of games that the team is supposed to lose which is also referred as upset games.

A complete league level transformability means when a higher ranked team faces a lower ranked team, the higher ranked team wins the game. For example, team A,B,C,D have a rank order A>B>C>D. Team A will defeat all the other teams when the rank order is transformable to the game result. Similarity, Team B defeats all other teams but Team A, Team C can only defeat Team D, and Team D loses all games.

In this research, the league level rank transformability is defined as the summation of two parts: a) the number teams that have a lower rank than Team i, and lost more than 50% of their games with Team i, and b) the number of teams that have a higher rank than Team i, and win more than 50% of their games with Team i.

- Dominance measure by within pair winning percentage

The purpose of the measure is to capture the unexpected results of a pair, based on the pair's previous winning percentages.

$$\text{DOMA}_{ij,\text{year}} = \frac{\#\text{ of games } lower\ rank\ wins}{games\ between\ ij} = \frac{\sum_{j=1}^{j=g_{ij}} I_{\min \text{rank}(i,j),\max rank\ (i,j)}}{g_{ij}} \qquad \text{(p2.1)}$$

In equation (p2.1), g_{ij} is the number of games played between i and j in the season, and $I_{\min \text{rank}(i,j),\max rank\ (i,j)}$ is the indicator equation of a lower rank team wins.

$$I_{\min \text{rank}(i,j),\max rank\ (i,j)} = \begin{cases} 1, lower\ rank\ team\ wins \\ 0, lower\ rank\ team\ losts \end{cases}$$

Therefore, $\frac{\sum_{j=1}^{j=g_{ij}} I_{\min \text{rank}(i,j),\max rank\ (i,j)}}{g_{ij}}$ is the winning percentage that the lower ranked team gets from the higher ranked team in the games between a matched pair of teams. When the lower ranked team wins all the games, the ranking is not transformable to this pair at all, and the lower ranked team dominates the pair. On the other hand, if the lower ranked team loses all the games, the ranking is fully transformable, and the higher ranked team dominates the pair. When no team dominates in the pair, the winning percentage of either team will be close to 50%. Thus, equation (p2.1) takes values on [0, 1], and a higher value indicates higher ranking transformability, and an absolute value close to 0.5 indicates less dominance of a higher winning percentages team over a lower winning percentages team.

DOMA capture the differences in situations like:

When Team i has higher overall winning percentage (higher rank) than Team j, and Team i wins all the games between Team i and Team j;

When Team i has higher overall winning percentage (higher rank) than Team j, and Team i does not wins all the games between Team i and Team j;

3.4.4 An Indicator of Performance Consistency.

Performance consistency investigates team performance based on its previous record. There are two levels of consistency for each team: between season team performance consistency and within season team performance consistency. Between season performances consistency measures the rank change for a team in a given period. In this research, I will measure the team rank variance in a three year period. The team between season performance consistencies can be denoted as:

- Between season consistency measure

$$CSCA_{ij,\ year} = \frac{\text{std(wp of i in past 3 years)} + \text{std(wp of j in past 3 years)}}{2}$$

$$= \frac{\sqrt{\frac{\sum_{y=1}^{3}(wp_{i,year-y} - \frac{\sum_{y=1}^{3} wp_{i,year-y}}{3})^2}{3}} + \sqrt{\frac{\sum_{y=1}^{3}(wp_{j,year-y} - \frac{\sum_{y=1}^{3} wp_{j,year-y}}{3})^2}{3}}}{2} \qquad \text{(p3.1)}$$

In equation (p3.1), $wp_{i,year-y}$ is the winning percentage of Team i in year-y, y take values 1,2, and 3. And $\frac{\sum_{y=1}^{3} wp_{i,year-y}}{3}$ is the average winning percentages of Team i for the past three years.

The assumption is that when the team's winning percentage changes dramatically, there is less consistency in the team's performances and we get a bigger CSCA value. Thus, bigger CSCA indicates a better inter season balance situation.

The rank is another option to measure consistency. However, due to league expansions and division structural changes, it can be misleading to compare the team rankings across seasons. Therefore, winning percentage is the focus of this measure. Table 3.1 summarizes the number of teams and divisions since 1969.

Table 3.1

Divisions and Teams in Major League Baseball

	Nb. of Teams	Nb. Of Divisions
1969-1976	24	2
1977-1992	26	2
1993-1994	28	2
1994-1997	28	3
1998-Now	30	3

- Pair wise between-season consistency measures—chasing wins?

$$\text{CSCB}_{ij,g}=\left(\frac{\text{\# of consecutive wins between Team i \& } j \text{ by game g}}{games\ played\ by\ team\ i\ \&\ j\ by\ time\ t}\right)=\frac{\text{W}_{ij,g}\sum_{g=1}^{g_{ij}}\text{W}_{ij,g}}{g_{ij}} \qquad \text{(p3.2)}$$

In equation (p3.2), W_{ijt} is a indicator equation of whether the same team has consecutive wins in game g in the games played between Team i&j. The values of W_{ijt} for the first two games are zeros.

$$\text{W}_{ij,g} = \begin{cases} 1, for\ the\ game\ played\ between\ the\ pair, team\ won\ last\ time\ still\ won \\ 0, for\ the\ game\ played\ between\ the\ pair, team\ won\ last\ time\ lost \end{cases}$$

$\sum_{j=1}^{g_{ij}} \text{W}_{ijg}$ is the total consecutive win/lose at game g. g_{ij} is the total number of games played between Team i&j, CSCB is a measure of within season consistency; it focuses on continuous wins or continuous loses in a given period.

- Pair wise between season consistency measures—does last year count?

This is measure of the change of wing percentages of the teams in the game.

$$CSC3_{ij,g}=(wp_{i,g}-wp_{i,year-1})^2+(wp_{j,g}-wp_{j,year-1})^2 \qquad (p3.3)$$

In equation (p3.3), $wp_{i,year-1}$ is Team i's winning percentage of last season.

3.5 Significance of Proposed Measures—the Benefit of Applying Proposed Competitive Balance Measures

The significance of these proposed measures lies in five aspects. First, researchers have acknowledged the multiple dimensions of Competitive Balance for a long time ago. The multiple dimensionalities of Competitive Balance are difficult to be represented with a single measure. This research provides a set Competitive Balance measures that address different dimensions.

Second, existing measures usually have one summarized value for a team/league in a giving season(s), but a summarized value is not good at reflecting real-time changes during a season. Most of the proposed measures in this study are calculated with real-time information, such as the game importance, consecutive wins, score differences, etc. So each matched pair has a set of measures for each game. Therefore the proposed measures capture more dynamic information.

Third, the subject of previous studies was individual teams. For example, for the standard deviation of winning percentages measure, the basic unit of analysis is individual team and the measures of Competitive Balance are a summary statistic of individual team's winning percentages. It is problematic to fit the team-based measures with regression analysis, because traditional OLS regression invokes three assumptions: the dependent variables and independent variables have a linear relationship, the equation errors (residuals) are independently and identically distributed, and homoscedasitic variance across occasions and research objects (Singer & Willett, 2003)(p55). In baseball records, because the same pair of teams is measured on several occasions, any unexplained pair specific time invariant effects in the residuals will create a correlation across games. When ignoring those unexplained time invariant effects, we will find the outcome may have a different precision or reliability at different occasions. The proposed measures are based on matched pair of teams, thus there is no need to assume the independence of the teams in the same game. In addition, we can examine the dependence of games between the same/different matched pairs easily.

Fourth, this study examines the relationship between Competitive Balance and attendance by relaxing the assumption of variance and covariance structures between/within the matched pairs. By constructing growth models, this research allows each matched pair to have their own average attendance and change rates as well as correlated residual structures for games between the same pair. Results show that models with relaxed assumptions have better model fit statistics [chapter 6].

Fifth, this research not only presents the trend of Competitive Balance in the line of Analysis of Competitive Balance (ACB), but also examines the Uncertainty of Outcome Hypothesis (UOH). Analysis of Competitive Balance and Uncertainty of Outcome Hypothesis are two lines existing in Competitive Balance literature (Fort, 2006), and this research contributes to both lines of the literature. This research contributes to the literature of Analysis of Competitive Balance literature by displaying the trends and patterns of Competitive Balance across seasons and locations/divisions, and contributes to the Uncertainty of Outcome Hypothesis literature by inquiring about how the spectators react to Competitive Balance measures.

Last, this study introduces game importance into the matrix of Competitive Balance. It is reasonable to assume that teams are willing to devote more effort to the games that matter than to their standings in a division/league. Score shaving is often the topic not only in Baseball, but other sport games as well. Therefore, we should not treat winning as a simple win, instead, we should consider how the game importance affects the game results. To my knowledge, the current study is the first research that considers game importance as a factor of Competitive Balance. Chapter 6 will show that game importance is an important factor in average game attendance and attendance change rate. Thus game importance should not be ignored, especially when we study Competitive Balance under the Uncertainty of Outcome hypothesis.

3.6 Chapter Summary

This chapter first describes the Major League Baseball game schedule, the data resources, and then discusses the advantages and disadvantages of using secondary data analysis. To fit the character of the data at hand, four possible solutions for longitudinal data analysis are evaluated in this chapter as well.

The last section of Chapter 3 proposes a set of Competitive Balance measures which addresses three dimensions of Competitive Balance: Closeness, Dominance and Consistency.

The closeness measures aimed at capturing how close the teams in each matched pair are in games, in terms of scores, winning percentages, and game importance. Game importance measures aim to the cover the possible score shaving effects that may hide the true Competitive Balance of the game. The dominance measure is based on how many times one team wins over another in games between the matched pair. The consistency measures focus on the consistency of the game results of the matched pair. Table 3.2 displays all the proposed measures.

Table 3.2

A Set of Proposed Competitive Balance Measures

		Notes
Closeness Measures		
Measures based on current game scores	clo1	Same scores, wild card game vs. normal game
	cloA	Big score sum vs. lower score sum, same game length
Measures based on previous wp	dom2	Last year's wp gap
	dom3	Last year's wp sum
Importance base on real-time wp	impS	Importance differences
	impD	Importance for both team
Dominance measure	doma	Ratio of underdog wins
Consistency Measures		
Consecutive wins	cscb	Within pair
Gap between real-time& previous wp	csc3	WP changes from last year for both teams in a pair
Measures based on the variance of wps	csca	STD of three years WPs for each team in a pair

CHAPTER 4: EXPLORING THE PROPOSED COMPETITIVE BALANCE MEASURES

Vrooman (1996) states that, conceptually, Competitive Balance has three interrelated issues: Dominance of large-market clubs, the closeness of league competition within the season, and the continuity of performance from season to season. In Chapter 3 I proposed a set of measures that aim to measure these three dimensions of Competitive Balance. So far, I do not know if the proposed set of measures has three dimensions, or if it is only my wishful thinking. The worst case scenario is that all ten measures are redundant to each other, and they do not provide any additional information other than winning percentages. A natural question to ask therefore is whether the proposed measures are truly multidimensional.

The following research attempted to answer this question in three steps: 1) Detecting the underlining dimensions of the proposed Competitive Balance measures; 2) If the measures are multidimensional, what are the relationships between the proposed measures? and 3) How the proposed Competitive Balance measures related to the existing Competitive Balance measures.

4.1 Detecting the Structure in the Proposed Competitive Balance Measures: Are the Measures Truly Multidimensional?

4.1.1 Research question.

In chapter 3, I proposed ten measures for the three dimensions of CB: Six measures for Closeness, one measure for Dominance and three measures for Consistency. The research goal for this section is to examine if the designed measures capture different dimensions of Competitive Balance as desired.

4.1.2 Method-Principal Component Analysis.

Principal Component Analysis (PCA) is performed to identify both the presence and nature of the multidimensionality of the proposed measures. All the ten measures in the Competitive Balance measures will be rotated to form factors so as to find the dimensional structures of the measures. If the measures are one-dimensional, one factor will be able to represent all the measures. On the other hand, if the results suggest more than one factor to represent the Competitive Balance measures, I can conclude that the measures are multidimensional.

Varimax rotation will be applied in the PCA because my research interest is to check if I can find three or more factors to represent the ten measures, and the varimax rotation provides an easy way to interpret the results for my research interest. The data set for the PCA is the MLB records from 1901 to 2008. After excluding outliers and records with missing values, a total of 127,707 games are used in the analysis.

4.1.3 Results and discussion.

The correlations among the measures are presented in Table 4.1. This result shows that only one pair of Closeness measures (impD & DOM2) have a correlation level higher than 25%. This indicates that the proposed measures are not identical to each other.

Table 4.1

Correlations among the Competitive Balance variables

Variable	CLO1	CLOA	DOMA	DOM2	DOM3	impD	impS	CSCA	CSCB	CSC3
CLO1	1									
CLOA	0.02349	1								
DOMA	-0.00307	0.02688	1							
DOM2	0.13261	-0.00999	0.0073	1						
DOM3	0.00134	-0.03306	-0.02789	-0.00105	1					
impD	0.22462	-0.00454	0.00381	**0.34346**	0.0013	1				
impS	-0.00411	-0.04182	-0.03087	-0.00041	0.13774	-0.00861	1			
CSCA	0.00008	0.00615	-0.10479	0.00116	-0.0518	-0.00055	-0.00075	1		
CSCB	0.12908	-0.01543	0.0244	0.01485	-0.00011	0.16245	0.07807	0.00823	1	
CSC3	-0.00163	-0.00592	-0.03649	-0.00113	-0.01288	0.04465	-0.07047	0.09978	0.19279	1

The eigenvalue of the first four factors are all greater than one, and these four factors explained about 50% the variance in the data (Table 4.2). In practice, when researchers need to make a decision about the minimum number of factors, eigenvalue 1 or 70% of explanatory power often serves as the cut-off point. In this study, the first four factors do not fully represent the information in the ten proposed measures. In order to explain more than 70% of the variation in the data, the number of the factors must increase to seven (Table 4.2). Considering that the original number of variables is ten (the ten Competitive Balance measures), using factor analysis to deduce the dimensions in the future analysis is not unnecessary, because we have to keep at least seven factors to explain 70% of the information in the data. At the same time, the explanation power of the factors are similar--none of the factors explains more than 16% or less than 6% of the variances in the data.

Table 4.2

Eigenvalues of the Correlation Matrix

Number	Eigenvalue	Difference	Proportion	Cumulative
1	**1.55375632**	**0.3408389**	**0.1554**	**0.1554**
2	**1.21291743**	**0.03416434**	**0.1213**	**0.2767**
3	**1.17875309**	**0.09858165**	**0.1179**	**0.3945**
4	**1.08017144**	**0.09121627**	**0.108**	**0.5026**
5	0.98895517	0.07550872	0.0989	0.6015
6	0.91344645	0.01769827	0.0913	0.6928
7	0.89574817	0.05100798	0.0896	**0.7824**
8	0.8447402	0.12881676	0.0845	0.8668
9	0.71592344	0.10033515	0.0716	0.9384
10	0.61558829	_	0.0616	1

The rotated factor pattern Table shows how much the factors correlated with the measures (Table 4.3). The values in the Table 4.3 reflect the unique variance each of the four factors contributes to the variances of the measures. Factor loadings, also known as component loadings in PCA, are the correlation coefficients between the measures and the factors. Ideally, the Competitive Balance measures should load highly on just one factor each, that is, each column in the Table has a value 1, and all other values zero. For example, the ideal loadings of CLO1 on the four major factors can be: 1, 0,0,0, this means factor one can fully represents CLO1.

The ideal situation rarely happens in practice, so researchers need to decide which variables are important for the factor construction based on the loading values. A loading of 0.45 corresponds to about 20% of the variance in the measures being explained by the factor, and is often used as the cut-off value.

Table 4.3

Rotated Factor Patterns

Variable	Factor1	Factor2	Factor3	Factor4
CLO1	**0.55578**	0.13425	0.0056	0.06309
DOM2	**0.73806**	-0.16793	-0.02807	-0.0721
impD	**0.78098**	0.11275	-0.00373	-0.00276
CLOA	0.00637	-0.01548	-0.33267	0.12079
impS	-0.00666	0.02365	**0.72779**	0.02639
DOM3	-0.00397	-0.03542	**0.66849**	0.07569
DOMA	-0.01564	0.09373	-0.20339	**0.72973**
CSCA	-0.01534	0.14372	-0.10661	**-0.70582**
CSCB	0.18076	**0.76754**	0.1612	0.14645
CSC3	-0.05266	**0.73136**	-0.13273	-0.21177

Table 4.3 shows that the Closeness measures are heavily loaded on Factor1 (CLO1, DOM2 and CLOA) and Factor 3 (impS and DOM3); the Consistency measures have the highest loadings on Factor 2 (CSCB and CSC3); and the Dominance measure and one Consistency measures are jointly loaded on Factor 4 (DOMA and CSCA). There are no obvious cross loadings in the rotated Table. This indicates that the factors are not associated with each other. The relative independence of the four factors can be observed in the scatter plots as well, because the scatter plots between the pairs of factors seem reasonably clear (only shown Factor1 & Factor 2, Figure 4.1).

```
                             Factor1
                                1

                               .9

                               .8  F
                          D
                               .7

                               .6
                                    A
                               .5

                               .4

                               .3

                               .2                     I
                                                                F
                               .1                               a
                                                                c
-1 -.9-.8-.7-.6-.5-.4-.3-.2-.1EBG C1H.2 .3 .4 .5 .6 .7 .8 .9 1.0t
                                                     J          o
                              -.1                               r
                                                                2
                              -.2

                              -.3

                              -.4

                              -.5

                              -.6

                              -.7

                              -.8

                              -.9

                               -1

CLO1=A        CLOA=B        DOMA=C        DOM2=D        DOM3=E        impD=F        impS=G
CSCA=H        CSCB=I        CSC3=J
```

Figure 4.1 Factor pattern for Factor1 and Factor2

In Table 4.1.3, there is one Closeness measure (CLOA) has no notable loadings on any of the factors, it could mean that CLOA is not represented by any of the four factors. To verify this suspicious, the Communality Estimates (Table 4.4) are checked:

Table 4.4

Final Communality Estimates

CLO1	**0.33092579**
CLOA	**0.12553675**
DOMA	0.58290054
DOM2	0.57891844
DOM3	0.45387971
impD	0.62266742
impS	0.530973
CSCA	0.53044403
CSCB	0.6692216
CSC3	0.60013098

The communities (squared multiple correlations) indicate the percent of the variance in a variable that overlaps the variances in the factors. It can be used to investigate whether the variables are well defined by the factors solution. From the Table it is clear that the four-factor solution basically ignores the information in CLO1 and CLOA, thus we cannot use the four factors as a substitute for the ten measures.

In this section, we successfully confirmed that the proposed measures are not one-dimensional, because one factor could not represent the information in the ten measures. At the same time, the factors which are formed by the measures are consistent with the suggested theoretical dimensions. For example, the Closeness measures are heavily loaded on Factor1 (CLO1, DOM2 and CLOA) and Factor 3 (impS and DOM3); the Consistency measures have the highest loadings on Factor 2 (CSCB and CSC3). However, to represent all the information in the measures, three factors are not sufficient. Instead, a seven-factor solution is required to represent seventy percent of the variances in the Competitive Balance measures. Therefore, using a three-factor or four-factor solution to study the relationship between the theoretical dimensions of Competitive Balance can be misleading.

4.2 The Relationship between the Theoretical Dimensions of Competitive Balance Using Canonical Correlation Analysis

4.2.1 Research questions.

Vrooman states that Competitive Balance has three interrelated issues: Dominance of large-market clubs, the closeness of league competition within the season, and the continuity of performance from season to season. Section 4.1 confirmed that the proposed measures are multidimensional, and the factors which are formed by the measures are consistent with the suggested theoretical dimensions. However, PCA cannot efficiently reduce the measures into three dimensions as desired. Therefore, to study the relationship between these dimensions, I will use the proposed Competitive Balance measures to form three variables which correspond to the three dimensions of Competitive Balance, and then explore the relationship between the three theoretical dimensions of Competitive Balance via these three variables. In other word, the research objective in this section is to understand the relationship among the three sets of data: the measures for the Closeness, the measure of Dominance, and the measures of Consistency.

4.2.2 Research methods.

This section will explore the correlation among the dimensions using canonical correlation analysis. It is a technique often used in examining the relationship between two multivariate data sets. In canonical correlation analysis, linear combinations of the measures of each dimension (canonical variables) are created such that the correlations between the canonical variables are maximized.

The combined canonical variables are analog to the eigenvectors of PCA. However, the differences between the two are that the new canonical variables are formed exclusively by the measures in each dimension. Since the new canonical variables best represent the measures in each dimension, the exploration of the relationship between the conceptual dimensions will based on the canonical variables. Game records from 1901 to 2008 season are used to perform the canonical analysis.

4.2.3 Results and discussion.

Because the canonical correlations deal with two sets of variables at one time, and the research interest is to check the relationship between the three sets of variables that correspond to the three conceptual dimensions proposed by Vrooman(1996), the canonical procedure is performed three times, one for each pair of dimensions.

4.2.3.1 Examining the relationship between Closeness and Dominance.

Canonical correlation analysis forms a new Closeness canonical variable by linearly combining the Closeness measures, and the same held with the Dominance and Consistency canonical variables.

Table 4.5 provides a decomposition of the canonical variables. The approximate F test shows the first component in the canonical test is significant ($p<0.0001$). Because the first pair of canonical variables for Closeness and Consistency accounts for almost all data variability in these two dimensions (82.4%), additional pairs of canonical variable are not considered in the later analysis. Because Dominance has only one variable (DOMA) there is no need to form canonical variables for this dimension.

Table 4.5

Test of Canonical Correlation

		Eigenvalues of Inv(E)*H CanRsq/(1-CanRsq)				Test of H0: The canonical correlations in the current row and all that follow are zero				
		Eigenvalue	Difference	Proportion	Cumulative	Likelihood Ratio	Approximate F Value	Num DF	Den DF	Pr > F
CLO&CSC	**1**	0.0482	0.0407	**0.824**	0.824	0.9442559	406.27	18	356962	**<.0001**
	2	0.0076	0.0048	0.1291	0.953	0.98978364	129.93	10	252412	<.0001
	3	0.0027		0.047	1	0.99725808	86.75	4	126207	<.0001

The correlation between the canonical variables, also known as canonical correlation, is presented in Table 4.6. The canonical correlation between the first pair of canonical variables for Closeness and Dominance is 0.05, and the correlation between Closeness & Consistency dimension is 0.21, and Consistency & Dominance dimension is 0.11. Neither the correlations of the first pair of canonical variable is greater than 0.30, thus the correlations between the three dimensions are not very strong.

Table 4.6

Canonical Correlation

		Canonical Correlation	Adjusted Canonical Correlation	Approximate Standard Error	Squared Canonical Correlation
clo &dom	1	**0.047216**	0.046798	0.002809	0.002229
clo&csc	1	**0.214471**	0.214339	0.002685	0.045998
	2	0.086574	.	0.002794	0.007495
	3	0.052363	.	0.002807	0.002742
dom&csc	1	**0.110248**	0.110178	0.002764	0.012155

The canonical coefficients, shown in Table 4.7, are analogous to the loadings in factor analysis. The coefficient value of each measure indicates its importance in constructing the canonical variable. Because the measures have different scales, it is best to interpret the standardized canonical coefficients.

Result shows that when studying the relationship between Closeness and Dominance, impS (-0.5632), CLOA (0.5337) and DOM3 (-0.4952) (Table 4.7) are the major contributors in constructing the canonical variable for Closeness. None of the Closeness measures correlated with the Dominance dimension because all values in the first column in Table 4.8 are less than 0.05. When studying the correlation between the Closeness and Consistency dimension, impD (0.7055), CLO1(0.4911) and impS(0.449) are the variables that contribute the most to the Closeness canonical variable, and CSCB is the major contributor to the Consistency canonical variable (Table 4.7). ImpD, CLO1 and impS are more correlated to Consistency dimension than other measures in the Closeness dimensions, and CSCB (0.21) is more close to the Closeness dimension than another two Consistency measures (Table 4.8, column 2). None of the Consistency measures are correlated with the Dominance dimension more than 0.11 (Table 4.2.4, column 3).

Table 4.7

Canonical Coefficients & Correlations

Standardized Canonical Coefficients

	clo1		clo1	clo2*	clo3*		csc1
CLO1	-0.1112	CLO1	0.4911	-0.1311	-0.021	CSCA	-0.9017
impD	0.0501	impD	0.7055	0.6105	-0.0111	CSCB	0.2846
impS	**-0.5632**	impS	0.449	-0.8227	0.0727	CSC3	-0.2969
DOM2	0.1567	DOM2	-0.2362	-0.2087	0.0256		
DOM3	**-0.4952**	DOM3	-0.0591	0.0371	-1.0005		
CLOA	**0.5337**	CLOA	-0.0612	-0.1071	0.0866		
	dom1		CSC1	CSC2	CSC3		dom1
DOMA	1	CSCA	0.0192	-0.1297	0.9965	DOMA	1
		CSCB	1.0191	-0.0011	-0.0082		
		CSC3	-0.1963	1.0048	0.0312		

*The second and third pairs of canonical variable for Closeness & Consistency dimension (grey columns in Table 4.6) is not the focus of this study, because the first pair of canonical variable has explained 82.4% (Table 4.5) of the variances for the two sets of data.

Table 4.8

Correlations between the VAR Variables and the Canonical Variables of the WITH Variables

Correlation between Closeness & Dominance		Correlation between Closeness & Consistency		Correlation between Consistency & Dominance	
	dom1		csc1		dom1
CLO1	-0.0031	CLO1	0.1319		csc1
impD	0.0038	impD	**0.1568**	DOMA	0.1102
impS	-0.0309	impS	0.0934		dom1
DOM2	0.0073	DOM2	0.0154	CSCA	-0.1025
DOM3	-0.0279	DOM3	0.0014	CSCB	0.0244
CLOA	0.0269	CLOA	-0.0144	CSC3	-0.0367
	clo1		clo1		
DOMA	0.0472	CSCA	0.0017		
		CSCB	0.2105		
		CSC3	0.0005		

This section uses canonical correlations to determine the associations between the conceptual dimensions in Competitive Balance. The conceptual dimensions addressed by the proposed measures are only weakly associated (Table 4.5, all values are less than 21%). The association between each individual measure and other dimensions are not strong either (no more than 0.1568, Table 4.8). Therefore, it is safe to conclude that the correlations among the three conceptual dimensions in Competitive Balance are not strong.

4.3 Relating Proposed Measures to Winning Percentage

4.3.1 Research question.

Previous sections showed that the proposed Competitive Balance measures are multidimensional and the correlations among its theoretical dimensions addressed by the proposed Competitive Balance measures are weak. One remaining question is how the proposed measures related to existing Competitive Balance measures. To be specific, is there any information that proposed measures can tell but winning percentage cannot tell?

However, the proposed Competitive Balance measures are not directly comparable with existing Competitive Balance measures because the basic study object of the Competitive Balance measures is the paired teams, and each game generates a set of Competitive Balance measures. Whereas the basic study unit of the existing Competitive Balance measures is the individual team. For instance, winning percentage derived measures focus on teams' winning percentage/standard deviation of the winning percentage, and the HHI measures which teams win championships.

To relate the Competitive Balance measures with the individual team performance, I transform the pair-wise measures to a team-based measure. To achieve this, team capability will be estimated using the proposed measures, and then the rank based on the estimated team capability will be compared with the rank based on existing measures. Because the majority of existing measures are derived from winning percentages, and the wins or losses are based on the score difference, the investigation in this section will focus on the winning percentage and score differences measure in the Closeness dimension (CLO1).

4.3.2 Model construction.

The following section will propose three sets of models to estimate team capability, and then select the model that has the best fit statistics, and then compare the estimated capability rank with the winning percentage rank.

Match based analysis that focuses on winning margin (score difference) has been found in the home advantage literature (Clarke & Norman, 1995; Clarke, 2005; David & Smith, 1994;

Forrest et al., 2005b; Koning, 2000; Nevill & Holder, 1999). One of the most common models used by researchers is:

$$w_{ij}= u_i-u_j+ \boldsymbol{\varepsilon}_{ij} \quad [1]$$

Where:

- w_{ij} is the score difference between Team i and Team j,
- u_i is Team i's capability, and u_j is Team j's capability,
- $\boldsymbol{\varepsilon}_{ij}$ is random error term

Model [1] does not include home advantage, and each team has its own capability. However, home advantage has been shown to be correlated with game results in many studies in MLB and in other sports (Booth, 2005; D. Forrest et al., 2005b; Lapointe, 2004; Meehan Jr. et al., 2007; Stefani, 2008). One may argue that home advantage is not part of Competitive Balance, however we are evaluating team performance based on final results, thus we need to partition out the influence of home advantage in order to get an accurate estimations of team capabilities. So model [1] is expanded to:

$$w_{ij}= h+u_i-u_j+ \boldsymbol{\varepsilon}_{ij} \qquad [2a]$$

Where h is a measure of home advantage of all the teams.

Model [2a] assumes the home advantage for all teams is equal. This assumption can drive more complex models. For example, assuming the home advantage varies by team, then each team will have its own home advantage:

$$w_{ij}= h_m+u_i-u_j+ \boldsymbol{\varepsilon}_{ij} \qquad [2b]$$

Where h_m is a measure of home advantage of m^{th} division.

Another variation of model [2a] is to assume at home advantage varies by division and the teams in the same division have a similar home advantage, then each division has its home advantage:

$$w_{ij}= h_i+u_i-u_j+ \boldsymbol{\varepsilon}_{ij} \qquad [2c]$$

Where h_i is a measure of home advantage of i^{th} team.

The third set of models adds two more variables: game importance for the home team, and game importance for the away team. Research in soccer (Clarke, 2005) suggests that the variances explained by the first and second set of models are very small, and a considerable amount of variances in score margins are not explained by home advantages and team capabilities. Therefore, two more variables—game importance for the home team and game importance for the away team--will be included in the third set of models to improve the model fit, and thus generate a better estimation of team capability.

$$w_{ij}= Imp_i + Imp_j +h+u_i-u_j+\boldsymbol{\varepsilon}_{ij} \quad [w3a]$$

$$w_{ij}= Imp_i + Imp_j +h_i+u_i-u_j+\boldsymbol{\varepsilon}_{ij} \quad [w3b]$$

where

- $Imp_{i,t}$ is game importance of home team at time t
- $Imp_{j,t}$ is game importance of away team at time t
- $length_{ij,t}$ is the length of the game between the teams at time t

Model assumptions:

- $\sum_{ij} \varepsilon_{ij}=0$

- $\boldsymbol{\varepsilon}_{ij} \propto N(0, \sigma^2)$ or $w_{ij} \propto N(\mu_{ij}, \sigma^2)$
- u_i, h_i are constant throughout the season.

To estimate u_i and u_j in the function, assign

$$h_i = \begin{cases} 0 \; when\ team\ i\ is\ away\ team \\ 1, when\ team\ i\ is\ home\ team \end{cases}$$

and

$$u_i = \begin{cases} -1 \; when\ team\ i\ is\ away\ team \\ \;\; 1 \; when\ team\ i\ is\ home\ team \\ \qquad 0 \; otherwise \end{cases}$$

Since the ability of the teams are relative, an arbitrary constrain is added:

$$\sum_i u_i = 0$$

4.3.3 Results.

The 2008 season data for the AL is used in this section. Results are in Table 4.9.

Table 4.9

Regression Results of Model [1] [2] [3], ui,hi,R Square

	M1**	M2a	M2b	M2c	M3a	M3b
Adj R-Sq	0.0215	0.0307	0.0308	0.0302	0.0512	0.0528
Model DF	14	15	28	17	17	30
Error DF	995	994	982	994	993	980
F value	2.59	3.13	2.15	2.85	4.21	2.88
Pr>F	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001

**Model [1] is denoted as M1, model [2a] is denoted as M2a, etc.

Only results for the American League 2008 season is shown in this paper, and all the three sets of models are highly significant ($p<0.001$). The results for other seasons and results for National League are also similar. Thus it is safe to conclude that the margin scores differ by matched pair of teams. The adjusted R square for all the models using winning margin are below 10%, indicating a high variance in the score differences. All the three models are significant compared with the null model, and the estimated u_i and h_i followed normal distributions (Pr > W-Sq >0.2500 for all three models. Normality tests are not shown). A well behaved model should have a residual plot with predicted values randomly scattered in a constant width band about the zero line. The residual plots (refer to Appendix) of all the models are randomly distributed around the zero, and they have no curve trends.

A model comparison is conducted to find out whether adding additional variables to model [1] is necessary. Table 4.10 shows that model [2a] significantly differs from model [1] (F=10.375, P=0.003), but model [2b] does not shows any significant different from model [1] (F=1.679, P=0.118). Therefore, home advantage does affect the score margin (p=0.003 for model [2a]), but the home advantages are not necessarily differ from team to team (p=0.118 model [2b]).

Model [2c] assumes different home advantages for each division. The adjusted R square is not much different from model [2a]. Thus, assuming home advantage varies by division does not improve model fit either.

$$F = [(SSE_a - SSE_b)/\text{change of model df}]/MSE_b$$

Table 4.10

Model Comparison Table

Source	Df	SSE	MSE	F		P
M1	14	19459				
M2a	15	19258	19.37384	F2a|1	10.37481	0.003
M2b	28	19004	19.353	F2b|1	1.679326	**0.118**
M2c	17	19228	19.34443	F2c|1	3.980474	0.026
M3a	17	18812	18.944	F3a|1	11.38432	<<0.001
				F3a|2a	11.77142	0.001
M3b	30	18534	19.55698	F3b|1	2.956106	0.005
				F3b|2b	12.01617	0

In Major League Baseball, because of the configuration of ball parks, one would assume that home advantage should be team specific, and if the model assumes each team has its unique home advantage should perform better than other models. However, the test results show that the variance explained by adding 15 more parameters—one home advantage parameter for all team verses 16 home advantage parameters, one for each team—is not significantly higher than the model assuming one home advantage for all teams. Model comparison results are presented in Table 4.10.

In conclusion, model comparison (Table 4.10) suggests that model [3a] is a better fitting model. It means that, in addition to team capabilities, game importance affects team performance in addition to the team capabilities. Because the influence of overall home advantage and game importance are partialled out from model [3a], its team capability estimates ought to be more reliable than model [1].

4.3.4 How the score difference measure relates to winning percentage rank.

Table 4.11 presents the team rank based on the team capability estimation verses team rank based on the winning percentage in a season. Only the American League's 2008 season results are shown below, however, I also did the same test using data in other seasons for the AL and the NL, and the results are similar to what is shown in Table 4.11.

The ranking estimated by model [1] is mostly consistent with the ranking based on winning percentages, the estimates based on model [3a] are not as consistent as model [1]. This indicates that the existing team rank may not strictly represent a team capability rank, because team performance is clearly influenced by home advantage and game importance, both of them are highly significant in predicting the score difference of games.

Table 4.11

2008 the AL Ranking Estimates Based on Game Score Differences

Team	**2008 AL Team winning percentage**	**Team ranking based on winning percentage**	**Team capability estimation derived from Model [1]**	**Team capability ranking based on Model [1] estimation**	**Rank difference**	**Team capability estimation derived from Model [3a]**	**Team capability ranking based on Model [3a] estimation**	**Rank difference**
ANA	0.61728	1	0.13975	1	0	0.2330	6	-5
TBA	0.59877	2	0.121	2	0	0.3067	5	-3
BOS	0.58642	3	0.04483	4	-1	0.6204	2	1
NYA	0.54938	4	0.05574	3	1	0.3587	3	1
CHA	0.54601	5	0.02028	6	-1	-0.0792	7	-2
MIN	0.53988	6	-0.02743	10	**-4**	-0.1194	9	-3
TOR	0.53086	7	-0.00411	7	0	0.8213	1	6
CLE	0.5	8	-0.01022	8	0	0.3545	4	4
TEX	0.48765	9	0.03154	5	**4**	-0.4229	12	-3
OAK	0.46584	10	-0.01951	9	1	-0.4815	13	-3
KCA	0.46296	11	-0.04362	11	0	-0.6882	14	-3
DET	0.45679	12	-0.10108	13	-1	-0.0884	8	4
BAL	0.42236	13	-0.08191	12	1	-0.4120	11	2
SEA	0.37654	14	-0.12525	14	0	-0.4033	10	4
std					**1.70970083**			**3.55181427**

*bb. AL_m1_w_est reg w

* Theoretically GEE is a better choice than GLM when consider the scores as a ranked variable and the association among games. However the association does not seems to affect the team capability estimations and the results of GLM and GEE are about the same. GEE results are not shown.

Compared with the results in the following section, estimates for the Minnesota Twins (MIN) and the Texas Rangers (TEX) have larger discrepancies between the estimated rank and actual ranks. This is because the current estimate is based on score differences, while the true rank is based on wins/losses. In 2008, the total summation of the net score won by the Texas Rangers was more than that of the Minnesota Twins. However, the number of games won by the Texas Rangers was less than the Minnesota Twins. The Texas Rangers won 43 out of 162 (43/162=0.27) games, and the net score win of all its 162 games is 26. The Minnesota Twins won 88 games out of 163 games (88/163=0.54), and the net score win of Minnesota Twins in its 162 games is negative at -72. Here, net score win refers to the extra score won by one team against its opponent. For example, the Rangers played the Twins and the game result was 3:10, so the Rangers' net win is -7, and the Twins' net win is +7.

When using a logistic model to estimate the team capabilities based on the dichotomized game results (wins/losses), the results are highly consistent with the rank based on winning percentages (Table 4.12). The standard division of rank difference based on binary game results is about half of the standard deviation of the rank difference based on score margin (0.96 verses 1.71).

Table 4.12

2008 the AL Ranking Estimates Based on Wins/Losses

Team	2008 AL Team winning percentage	Team capability estimation derived from Model [1]	Team ranking based on winning percentage	Team capability ranking based on Model [1] estimation	Rank difference	Team capability estimation derived from Model [3a]	Team capability ranking based on Model [3a] estimation	Rank difference
ANA	0.61728	-0.4259	1	1	**0**	-0.4182	1	0
TBA	0.59877	-0.4787	2	2	**0**	-0.6713	3	-1
BOS	0.58642	-0.5128	3	3	**0**	-0.7298	5	-2
NYA	0.54938	-0.6348	4	4	**0**	-0.673	4	0
CHA	0.54601	-0.7779	5	6	**-1**	-1.0477	10	-5
MIN	0.53988	-0.8735	6	8	**-2**	-1.1022	13	-7
TOR	0.53086	-0.666	7	5	**2**	-0.6327	2	5
CLE	0.5	-0.8394	8	7	**1**	-0.7829	6	2
TEX	0.48765	-0.9744	9	9	**0**	-0.8439	7	2
OAK	0.46584	-1.0532	10	10	**0**	-1.0841	11	-1
KCA	0.46296	-1.1685	11	11	**0**	-1.0858	12	-1
DET	0.45679	-1.1918	12	13	**-1**	-1.041	9	3
BAL	0.42236	-1.1872	13	12	**1**	-1.1195	14	-1
SEA	0.37654	-1.4189	14	14	**0**	-0.9707	8	6
				Std	**0.96076892**		std	3.50823208

In general, the team ranking based on winning percentage and team ranking derived from score margin is consistent. However, some discrepancies exist between the winning percentage ranking and the ranking based on estimates. To be specific, the team rank estimation based on dichotomized game results is more consistent with winning percentage ranking than the rank estimation based on score margin. At the same time, rank estimation generated by the simplest model (model [1]) is more consistent with the winning percentage rankings than the estimations generated by the models with game importance and home advantages, given the simplest model does not have the best model fit.

One explanation of the results is that winning percentage ranking is about team performance, but it does not consider the score difference. Therefore, the rank estimation based on the dichotomized game results is more consistent with the winning percentage team ranking than the rank estimation based on score margin.

Meanwhile, winning percentage based on team ranking is about team performance, it does not consider game importance or home advantage. As we can see when I ignore the information in the game importance and the home advantage, the team ranking estimates are more close to the winning percentage rankings, even though we know both the game importance and the home advantage are significant in explaining the score differences of games. Therefore, it is reasonable to question whether winning percentage as a one-dimensional Competitive Balance measure does not capture the multidimensionality of Competitive Balance. The information revealed by game importance and score differences are not shown in the team ranking based on winning percentages.

4.4 Chapter Summary and Conclusions

This chapter examines the measures of Competitive Balance proposed in Chapter 3 from three aspects: 1) Confirms the proposed Competitive Balance measures are multidimensional; 2) Presents the relationship between the dimensions of the Competitive Balance measures; 3) Relates the proposed Competitive Balance measures with the winning percentages.

Results show that the Competitive Balance measures are multidimensional [4.1] and that there are no strong correlations between the proposed measures. In addition, the correlations between the three theoretical dimensions--Closeness, Dominance and Consistency-- are weak [4.2]. The last section of the fourth chapter relates score difference, one of the closeness measures, with winning percentages.

The proposed Competitive Balance measures is not directly comparable to the existing Competitive Balance measures, because the basic research object of the proposed Competitive Balance measures is the matched pairs in a given period, and each game has a set of Competitive

Balance measures (Recall that matched pairs refer to distinct pairs of teams that scheduled to play in a season. For example, Orioles & Yankees is considered a matched pair when including 2008 data into study, because they played against each other at least once in 2008). However, in existing Competitive Balance measures, the basic research units are the individual teams. In order to relate the Competitive Balance measures with the winning percentages of each team, section 4.3 proposed three sets of models to estimate team capability with one of the Closeness measures—score difference.

The team rankings based on winning percentage and the team rankings derived by score margin were consistent largely. Discrepancies between the estimated ranking and winning percentage ranking are due to the fact that winning percentage team rank is not exactly about team capability, but rather about team performance. One common agreement among researchers is that team performance is subject to the influences of external factors such as game location (home/away), so it is reasonable to control the effects of external factors when a researcher aims to estimate team capabilities. However, using the same estimation techniques, the team capability ranking is less consistent with the winning percentage ranking when controlling external factors in the model.

In addition, team winning percentage ranking is a result of binary win/loss outcomes, and thus, does not include the performance differences in terms of score margins. As a result, using the same estimating techniques, the team capability estimations derived from binary results (wins/losses) are mostly consistent with the team winning percentage rankings, whereas the estimations derived from score margins drift apart from winning percentage rankings. On the other hand, when sacrificing the information in score differences, home advantage and game importance, the ranking estimates are more consistent with the winning percentage rankings.

This indicates that winning percentage as a one-dimensional Competitive Balance measure does not capture the multidimensionality of Competitive Balance.

CHAPTER 5: THE STRUCTURE OF COMPETITIVE BALANCE —A MULTIVARIATE APPROACH

5.1 Changes in Competitive Balance Patterns over Time

5.1.1 Research question.

Chapter 4 confirmed that the Competitive Balance measures are multidimensional and the information in the measures of Competitive Balance differs from winning percentages. A natural question that emerges is, what are the patterns revealed by the proposed measures? The following section will address two research questions. First, how do the Competitive Balance measures patterns change over time; and second, how does the Competitive Balance change across locations and leagues?

5.1.2 Research methods.

This chapter evaluates the diversity of the Competitive Balance measures in two steps. First, measure the inter pair similarity/dissimilarity regarding to all the ten proposed Competitive Balance measures. Again, matched pair refers to a distinct pair of teams that played against each other in a season. For example, the Baltimore Orioles (BAL) & Los Angeles Angels (ANA) played nine games in 2008, so the BAL&ANA is considered one distinct matched pair, hereafter referred to as matched pair. The Gower's general similarity coefficient will be used to evaluate attribute similarities between two sets of matched-pairs. In the current circumstance, the attributes are the ten Competitive Balance measures. Gower's general similarity coefficient will compare matched pair_i & pair_j on each of the Competitive Balance measure. It is defined as follows:

$$s_{ij} = \frac{\sum_k w_{ijk} s_{ijk}}{\sum_k w_{ijk}}$$

$$s_{ijk} = 1 - \frac{|x_{ik} - x_{jk}|}{r_k}$$

where:

- r_k is the range of values for the k_{th} measure
- x_{ik} is the k_{th} measure value of $pair_i$
- s_{ijk} denotes the contribution provided by the k_{th} variable, and
- w_{ijk} is usually 1 or 0 depending upon whether or not the comparison is valid for the k_{th} variable;

The SAS distance procedure is used to calculate the matrix of Gower's general similarity coefficient. The dimension of this matrix depends on the number of distinct matched pairs in a given season. For example, in 2008 the matrix has a dimension of 284*284, because there are 284 distinct matched pairs that have played against each other. The values in the correlation matrix range from 0 to 1, with higher values indicate more similarity.

The second step is to calculate the overall similarity of all the matched pairs in a season. Root-Mean-Square-Error (RMSE) is used to evaluate the overall Competitive Balance measures' similarity among the matched pairs in a season.

$$\text{RMSE } (\boldsymbol{\theta})_y = \sqrt{\frac{\sum_{ij=1}^{K_y} (s_{ij} - \hat{s})^2}{K_y}}$$

Where

- K_y is the total number of inter-pair similarities in season y. For example, in the 1941 season, there were 56 matched pairs, and $K_{1941} = C_{56}^2 = 1540$
- s_{ij} is the Gower's general similarity coefficient between matched for matched $pair_i$ & $pair_j$.

- $\hat{s}$ is the mean Gower's general similarity coefficient for all the matched pairs in the season.

RMSE measures the diversity of the structure similarity in a season. The more consistent the Competitive Balance, the smaller the RMSE value will be. If the Competitive Balance measures of the games are all the same, then the value of RMSE is close to zero.

5.1.3 Results.

The plot of average Gower's general similarity coefficient for each seasons from 1941 to 2008 is displayed in Figure 5.1. Because the records before 1941 have too much missing fields, they are excluded from the analysis. A total of 435 matched pairs played during the 70 seasons. Other than the AL and the NL, the Federal League (1914-1915), the Mexican League (1946-1947) and the Continental League (1959-1960) were other major professional baseball leagues that ever existed in American baseball history, but games in those leagues are not included in the current analysis. The Gower's general similarity coefficient plot shows an overall increasing trend, and indicates a growing Competitive Balance similarity among the matched pairs from season 1941 to 2008. The RMSE plot for the variances of the similarity coefficient has a decreasing trend over time (Figure 5.2), which shows that the diversity of the similarities of matched pairs has been decreasing across time.

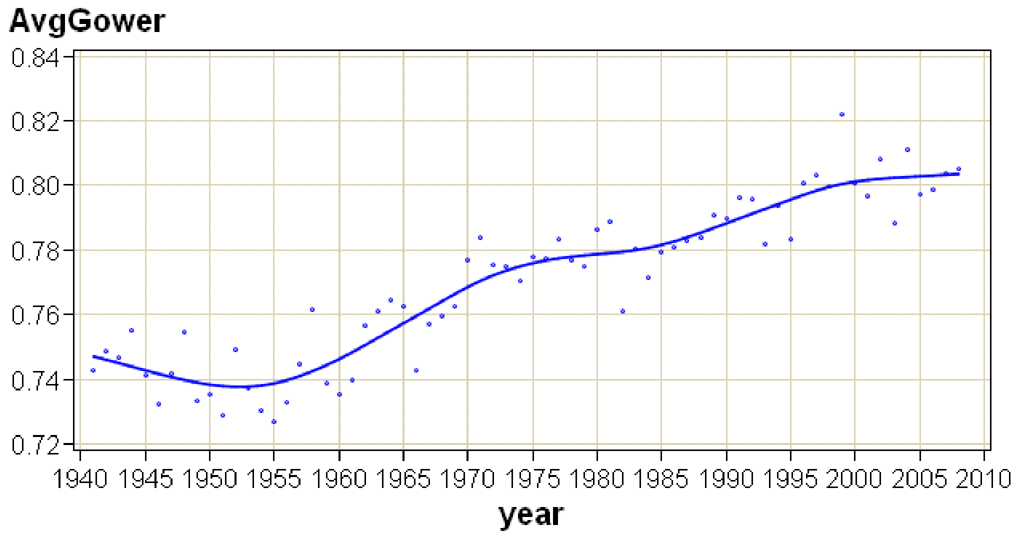
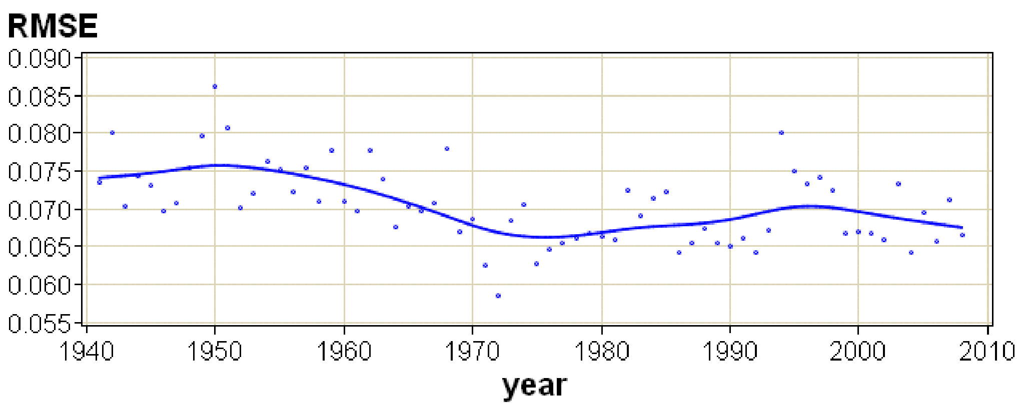


Figure 5.1 Plot of average Gower general similarity by year



*Plot generated by Enterprise Guide (file name: rank compare).

Figure 5.2 Plot of RMSE by year

Association analysis between Competitive Balance structure similarity and average game attendance shows that the average Gower's general similarity coefficient is significantly correlated with annual game attendance (Table 5.1). Since the Gower's general similarity coefficient is based on matched pairs, the average attendance is calculated based on matched pairs' average attendance of the year. The average Gower's general similarity coefficient is not normally distributed, thus the interpretation of R-square is not reliable in the current situation. However, it is safe to conclude that higher baseball annual attendance is significantly associated with higher Competitive Balance similarities among mated pairs in a season. Meanwhile, the count of matched pairs in a season is significantly positively associated with the average attendance.

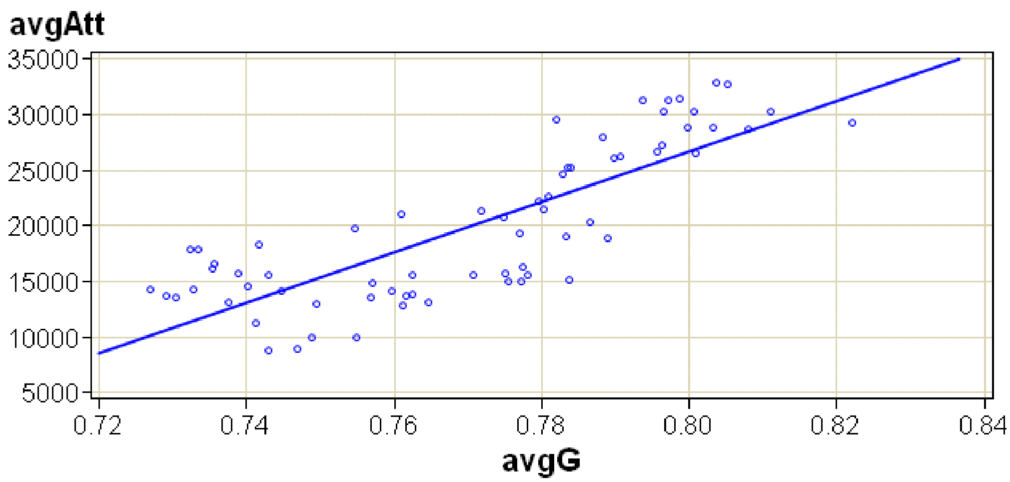


Figure 5.3 Plot of annual game attendance vs. Gower's general similarity coefficient

Table 5.1

Test the Association between Average Game Attendance and Average Gower's General Similarity Coefficient

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	2.10E+09	2.10E+09	128.55	<.0001
Error	66	1.08E+09	16333716		
Corrected Total	67	3.18E+09			

Root MSE	4041.4993	R-Square	0.6608
Dependent Mean	19948	Adj R-Sq	0.6556
Coeff Var	20.26017		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-155072	15444	-10.04	<.0001
avgG	1	227206	20039	11.34	<.0001

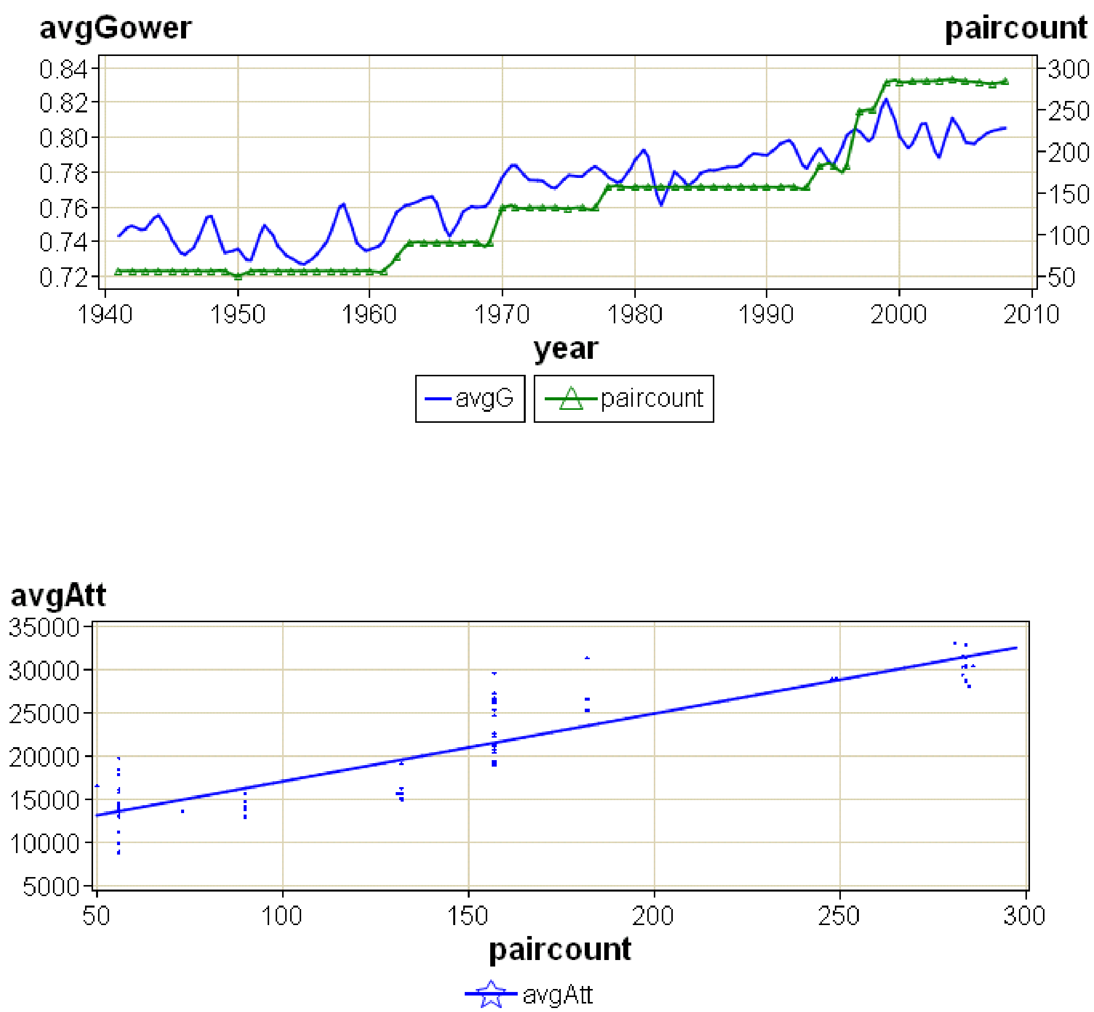


Figure 5.4 Plot of the number of matched pairs vs. Gower's general similarity coefficient

5.2 Changes in Competitive Balance Patterns across Leagues—a Multivariate Approach to Repeated Measures

5.2.1 Research question.

The previous section showed that the structure similarity of Competitive Balance measured by the ten proposed Competitive Balance measures has increased over time. The trend has increased particularly after 1996. It is possible that the observed time trend is due to teams that are from the American League (AL), or the National League (NL), or both. Because research in section 5.1 does not differentiate the league identity of the matched pairs of teams, it is unknown whether these two leagues have the same time trend. As discussed before, the AL and the NL are different in many ways including the number of teams, the number of games in divisions, and policies, etc. , and these differences could lead to the Competitive Balance structural differences. Therefore, the aim of this research is to compare the Competitive Balance measures' profile based on league identity.

5.2.2 Research methods.

Data used in this study came from the game records between 1901 and 2008. The basic unit of analysis object is the matched pairs in these 108 seasons, a total of 435 pairs from 1901 to 2008, and total of 126,214 games, excluding records with missing values, are used in this research. There are three categories of matched pairs: inter league pairs (one team from the AL and one team from the NL), the AL pairs (both teams are from AL), and the NL pairs (both teams are from the NL).

Because each matched pair of teams has played more than one game in this given period, each pair of teams has more than one set of measures. Based on data structure, profile analysis is be used for the study in this section. Profile analysis is a multivariate technique that can handle repeated observations. The GLM procedure for repeated measures is used to generate the profile analysis results. Three tests are conducted to compare the league's Competitive Balance profile: the flatness test, the parallelism test, and a test of the level of the profiles.

5.2.3 Results and discussion.

The results of the profile analysis of the ten measures of Competitive Balance for the three groups (the AL, the NL & the Inter league) are summarized in Table 5.2. Significant tests are shown for the flatness test (MANOVA Test Criteria and Exact F Statistics for the Hypothesis of no measures effect for measures), the parallelism test (MANOVA Test Criteria and Exact F Statistics for the Hypothesis of no measures effect for measure* teams league), and the levels test (Repeated measures analysis of variance of between subject effect for teams league), respectively.

The parallelism test, called *the test of the measure* pair league effect*, shows that the three categories of matched pair of teams--inter league, the NL and the AL-- are significantly different. So the gaps between Competitive measures differ for different pair categories. Various multivariate tests of parallelism all produced highly significant results (Wilks, Pillai's Trace, Hotelling-Lawley Trace & Roy's Greatest Root all have $p<0.001$, only Wilks is shown in the Table 5.2). Flatness and level tests demonstrate that the profiles for the three categories of matched pair are significantly different. Figure 5.5 is a plot of the Competitive Balance profile for the three pair categories. The Figure reflects the proportion of the actual mean value for the purpose of identification. The level values of all the competitive measures are shown in Table 5.2.

From Table 5.2, Table 5.3 and Figure 5.5, it is apparent that the profiles for the AL and the NL pairs deviate from the inter league pairs. Further level tests of the AL pairs verse the NL pairs show the profile these two types of pairs are similar to each other (result not shown).

However, when performing the level test on the NL pairs and the AL pairs in one season, say 2008, the profiles differences for the three categories of pairs are no longer significant (Table 5.4, $p=0.165$). But the interleague pairs are still significantly different from the AL and the NL pairs.

Table 5.2

The Competitive Balance Profile Tests for the Matched Pairs Categorized by Pair's League (1901-2008)

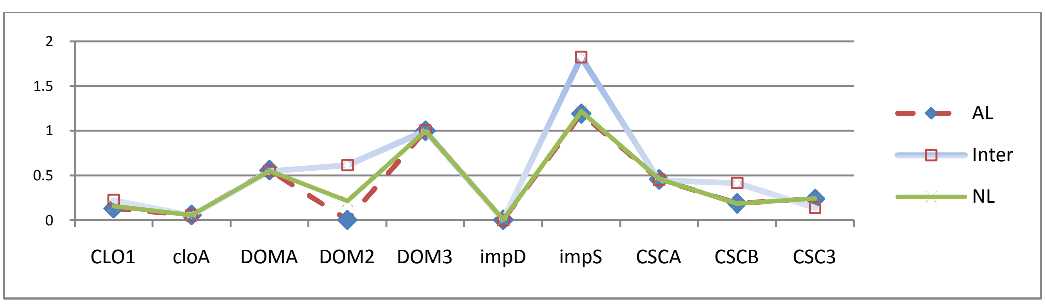
Test	Hypothesis	Statistic	Value	F Value	Num DF	Den DF	Pr > F
Flatness	No measures Effect	Wilks' Lambda	0.0496025	268676	9	126203	<.0001
Parallelism	No (measures*pair category) Effect	Wilks' Lambda	0.9269625	541.96	18	252406	<.0001
Level	No (between subjects) effects	Source	DF	Type III SS	Mean Square	F Value	Pr > F
		Pair category	2	245.5026	122.7513	65.03	<.0001
		Error	126211	238254.31	1.8877		

Table 5.3

Level for the Three Categories of Matched Pair of Teams

Categories of matched pair of teams	N	CLO1		cloA		DOMA		DOM2		DOM3	
		Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
AL pair	61516	0.1312	4.2972	0.0564	0.0268	0.5571	0.1562	0.0000	0.1157	0.9983	0.1040
Inter pair	2920	0.2240	4.4200	0.0543	0.0240	0.5506	0.2783	**0.0006**	0.1022	1.0007	0.1034
NL pair	61778	0.1538	4.1153	0.0556	0.0272	0.5488	0.1539	0.0002	0.1114	0.9969	0.0995

Categories of matched pair of teams	N	impD		impS		CSCA		CSCB		CSC3	
		Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
AL pair	61516	0.0027	0.1775	1.1889	0.3639	0.0455	0.0170	0.1863	0.2950	0.0237	0.0591
Inter pair	2920	0.0012	0.0971	**1.8232**	0.1004	0.0450	0.0182	0.4149	0.4366	0.0140	0.0143
NL pair	61778	0.0031	0.1772	1.2166	0.3839	0.0459	0.0176	0.1843	0.2953	0.0240	0.0587



*Some measure scales are changed in order to fit all measure into one plot.

Figure 5.5 The Competitive Balance profile for the matched pairs categorized by pair league--Means of the levels (1901-2008)

Table 5.4

Level Test of the Pair Categories Effect (the AL and the NL pairs) in 2008

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Category of the matched pair of teams	2	7.277433	3.638717	1.8	**0.165**
Error	2425	4893.7039	2.018022		

In conclusion, the profile analysis for the three categories of matched pair of teams shows that the Competitive Balance profile for interleague matched pairs (one team from the AL and one team from the NL) is significantly different from the matched pairs in which both teams are from the same league. The profile for inter league matched pair of teams is significantly different from the pair of teams that are from the same league. The closeness measure regarding previous winning percentages (DOM2) and overall game importance (impS) for inter league matched pairs are always higher than the pairs in which both teams are from the same league. Recall the definition of DOM2 and impS, the results indicate that the winning percentages in interleague pairs within the season are closer to each other than other matched pairs. And the interleague pairs previous winning percentages are more differ than other pairs.

The profiles of the matched pairs from the same league are not significantly different in some seasons, but differ in general during the period of 1901-2008.

5.3 Changes in Competitive Balance Patterns across Locations

To understand the Competitive Balance differences among the divisions, a profile analysis is also conducted to compare the matched pairs in which the teams came from different divisions. A total of 435 matched pairs for the season 1901-2008 and 90,166 games are used in the analysis after excluding the observations with missing data. When categorized by location, the pairs formed six groups in total: (1) pairs consisting of two west division teams, (2) pairs consisting of two east division teams, (3) pairs consisting of two central division teams, (4) pairs consisting of one west and one east division team, (5) pairs consisting of one west and one central division team, and (6) pairs consisting of one east and one central division team.

The profile analysis of the ten measures of Competitive Balance for the six categories of pairs is shown in Table 5.3. The parallelism test, *called the test of the measure* pair division effect*, shows a significant difference for the six categories. Again, all the multivariate tests of parallelism produce highly significant results--Wilks, Pillai's Trace, Hotelling-Lawley Trace & Roy's Greatest Root all have $p<0.001$, only Wilks is shows in Table 5.3. Therefore, there are statistically significant profile differences among the six categories of pairs.

The profiles of the six categories are plotted in Figure 5.6. Again, the Figure reflects the proportion of the mean value for the purpose of identification. The levels of all the ten measures of Competitive Balance are shown in Table 5.5 and 5.6. It is apparent that, for matched pairs in which both teams are from the West division or both teams are from the East division, the closeness measure for overall game importance (impS) is higher than other pair categories. The results indicate that the games played by these West-West pairs and East-East pairs are more likely to affect the teams’ standings in their divisions/leagues.

Single season profile analysis based on league affiliations no longer generates significantly results in some years (say 2008). The similar test is also performed to find the profile differences by divisions. The result shows that the division effect is still significant when using 2008 data (results not shown).

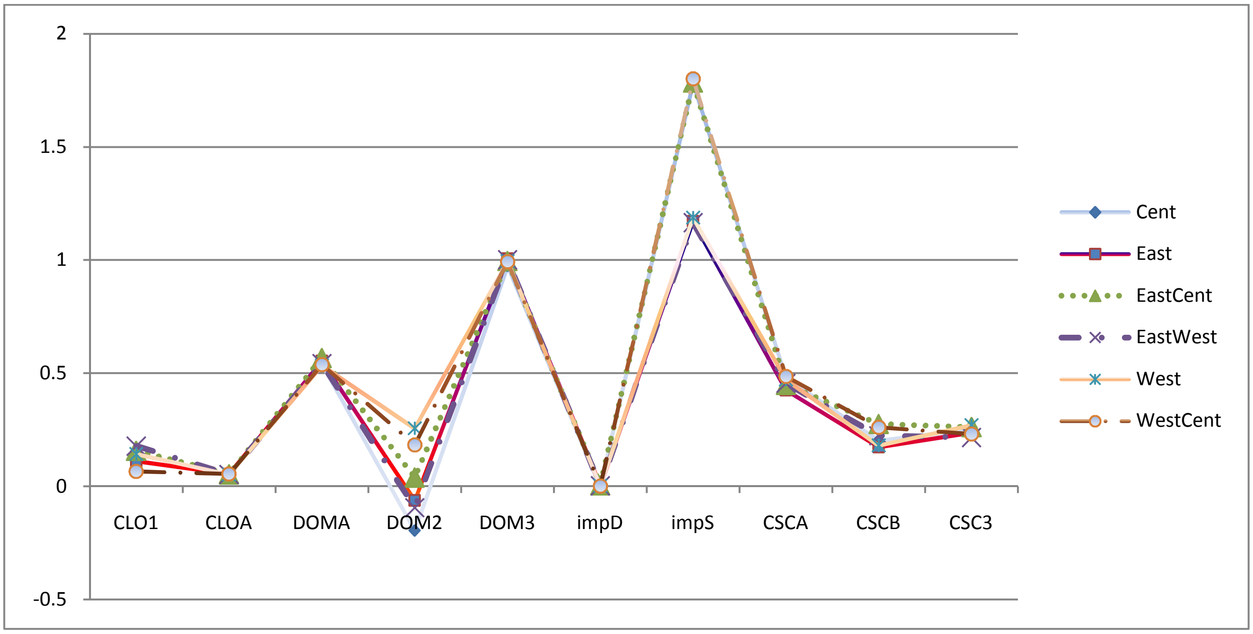


Figure 5.6 The Competitive Balance profile for matched pairs' division- Level means (1901-2008)

Table 5.5

The Competitive Balance Profile Tests for the Matched Pairs Categorized by Pair's Division (1901-2008)

Test	**Hypothesis**	**Statistic**	**Value**	**F Value**	**Num DF**	**Den DF**	**Pr > F**
Flactness	No measures Effect	Wilks' Lambda	0.01328505	743980	9	90152	<.0001
Parallelism	No measures*teamsLg Effect	Wilks' Lambda	0.59178309	1115.12	45	403275	<.0001
Level	No between subjects effects	Source	DF	Type III SS	Mean Square	F Value	Pr > F
		teamsdiv	5	625.2789	125.0558	66.31	<.0001
		Error	90160	170046.1603	1.886		

Table 5.6

The levels for the six types of pair groups

Level of teamsdiv	N	CLO1		cloA		DOMA		DOM2		DOM3	
		Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Cent	6007	0.1110	4.4195	0.0554	0.0238	0.5396	0.1649	-0.0002	0.1059	0.9743	0.0928
East	24291	0.1118	4.1847	0.0549	0.0259	0.5541	0.1472	-0.0001	0.1110	1.0068	0.0978
EastCent	6833	0.1573	4.4646	0.0540	0.0235	0.5677	0.2023	0.0000	0.1072	0.9964	0.1028
EastWest	29370	0.1777	4.1399	0.0528	0.0242	0.5421	0.1733	-0.0001	0.1005	1.0014	0.0955
West	17261	0.1434	4.1199	0.0521	0.0242	0.5350	0.1553	0.0003	0.0996	0.9959	0.0895
WestCent	6404	0.0670	4.4651	0.0557	0.0245	0.5387	0.1950	0.0002	0.0998	0.9931	0.0929
Level of teamsdiv	N	impD		impS		CSCA		CSCB		CSC3	
		Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Cent	6007	-0.0019	0.1510	1.7828	0.1769	0.0481	0.0162	0.2029	0.3166	0.0253	0.0612
East	24291	0.0029	0.1819	1.1721	0.3508	0.0425	0.0168	0.1735	0.2841	0.0238	0.0634
EastCent	6833	0.0054	0.1574	1.7849	0.1657	0.0449	0.0161	0.2760	0.3671	0.0262	0.0658
EastWest	29370	0.0023	0.1554	1.1651	0.3495	0.0455	0.0168	0.2231	0.3258	0.0216	0.0475
West	17261	0.0029	0.1862	1.1878	0.3743	0.0472	0.0175	0.1809	0.2905	0.0271	0.0666
WestCent	6404	0.0021	0.1371	1.8019	0.1668	0.0487	0.0180	0.2622	0.3506	0.0233	0.0537

In conclusion, when categorizing the matched pairs into six groups by their division identifies, the Competitive Balance profile for the six categories are different from 1901 to 2008. For the match pairs in which both teams are from the West division or both teams are from the East division, the closeness measure about overall game importance (impS) is higher than other pair categories. The results show that the winning percentages in west-west or east-east pairs are closer to each other than other matched pairs.

5.4 Chapter Summary and Conclusions

This chapter has two main aims to explore. First, documenting how Competitive Balance changed over time, and second, documenting how the changes relate to pair's league(s) and division(s). The basic unit of analysis is still the distinct pairs of teams that played against each other in a given season. In section [5.1] the correlation similarity coefficient matrix is calculated for all pairs based on the set of Competitive Balance measures, and then the Root-Mean-Square-

Error (RMSE) is used to evaluate the variance of Competitive Balance structural similarities of all the matched pairs across seasons.

I found that the structure of the Competitive Balance becomes more and more similar over time, and the annual game attendance increases with the similarity indicator. The data also revealed that the structural similarity of Competitive Balance increases with the count of matched pairs in a season. Therefore a league expansion is always appears to demonstrate an increasing of structural similarity. For example, the average attendance increased after the AL expanded from 8 to 10 and then again to 12 teams in 1968 and 1969, and increased again in 1994 when the NL expanded to 14 teams in 1993[5.1]. When the interleague game was introduced to MLB in 1997 the attendance increased again. This result is consistent with what Schmidt found (M. B. Schmidt, 2001). Schmidt concluded that the movement toward greater Competitive Balance occurred soon after the two MLB leagues began expanding.

This research found that games between matched pairs which two teams are from different leagues have higher Competitive Balance in the closeness dimension. Profile analysis shows that the competitive measures for interleague matched pairs (one from the AL and one from the NL) is significantly different from matched pairs in which both teams are from the same league. The closeness measures in terms of previous winning percentages and overall game importance for interleague matched pairs are always higher than the pairs in which both teams are from the same league [5.2]. This means that the interleague pairs within season winning percentages are closer than other matched pairs.

Meanwhile, teams that come from different divisions are associated with different Competitive Balance profiles. For the matched pairs in which both teams are from the West division or from the East division, the closeness measures regarding game importance are higher than other pairs [5.3]. This indicates that Competitive Balance is higher when the teams are spatially close to each other in terms of the closeness dimension, because the games played by such pairs are more likely to affect teams' standings in their divisions/leagues.

CHAPTER 6: COMPETITIVE BALANCE MEASURES AND GAME ATTENDANCE —A GROWTH MODEL APPROACH

6.1 How Competitive Balance Measures Relate to Attendance

6.1.1 Research question.

Chapter 5 examined how Competitive Balance changes over time which is also known as the Analysis of Competitive Balance. In the literature, another major approach of studying Competitive Balance is derived from the Hypothesis of Uncertainty of Outcome which assumes that Uncertainties of Outcomes are desired by spectators. Because this study using Humphreys definition of Competitive Balance, which directly links the Uncertainties of outcomes with Competitive Balance, this section will focus on examining the relationship between the proposed Competitive Balance measures with game attendances. Previous empirical studies in the literature revealed mixed results for the relationship between Competitive Balance and attendance in various sports. Some conclude that Competitive Balance is correlated with game attendances, and some conclude otherwise (Borland & MacDonald, 2003; Szymanski, 2003a; Szymanski, 2003b). In addition, Competitive Balance measures used in existing empirical studies often test the relationship between game attendances and Competitive Balance with one-dimensional Competitive Balance measures, thus inevitably missing the whole picture of how different dimensions of Competitive Balance associate with game attendances.

I believe that Competitive Balance is multidimensional, and the relationship between the game attendances and each dimension of Competitive Balance can vary. Thus, before we conclude whether Competitive Balance matters to game attendances, we should measure each dimension of Competitive Balance and study the relationship between the game attendances and all Competitive Balance dimensions. To find out whether it is necessary, this chapter will use empirical data to test the relationship between game attendances and the proposed measures of Competitive Balance.

6.1.2 Research methods—growth curve modeling.

Before deciding which method I should use, a data exploration is conducted to better understand the attendance data. The unit of analysis is matched pairs of teams in a given time period. Figure 6.1 shows a plot of the average annual attendance for each of the 433 matched pairs since 1901. However, due to missing values before 1910, the data used in the analysis focuses on the observations since 1911. As before, records and pairs consist of team(s) outside of the AL or the NL (such as the Federal League) are excluded from the data.

Note from the plot that there is substantial heterogeneity among the matched pairs. The initial average attendance (intercept) varies pair by pair, and the attendance change rate (slope) also differs pair by pair. Therefore, growth curve modeling is used in this section. To understand the relationship step by step, the first part will focus on inter pair attendance differences, and the next step focuses on the inner pair differences across time.

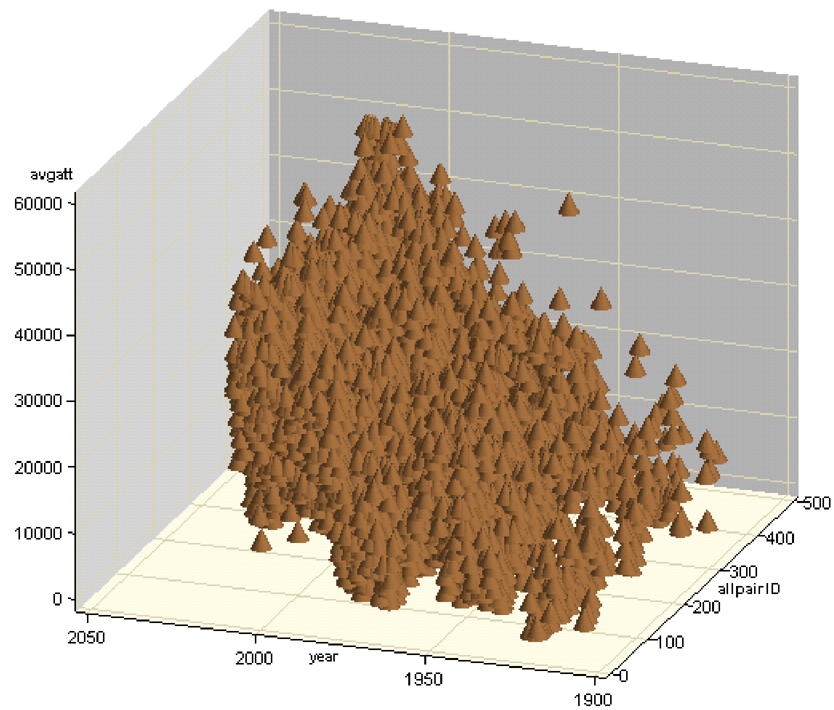


Figure 6.1: Plot of the average annual attendance of each matched pair of teams (1911-2008)

6.1.3 Model development.

The assumption used in constructing the attendance model is that game attendance is determined by constrains such as market population size, income and schedule of the game, as well as motivation to attend (Figure 6.2). By the Uncertainty of Outcome Hypothesis, Competitive Balance can serve as a motivation for spectators. Competitive Balance is measured by the ten proposed measures, and both constrain and motivations variables are included in model shown in Figure 6.2 and model 6.1.

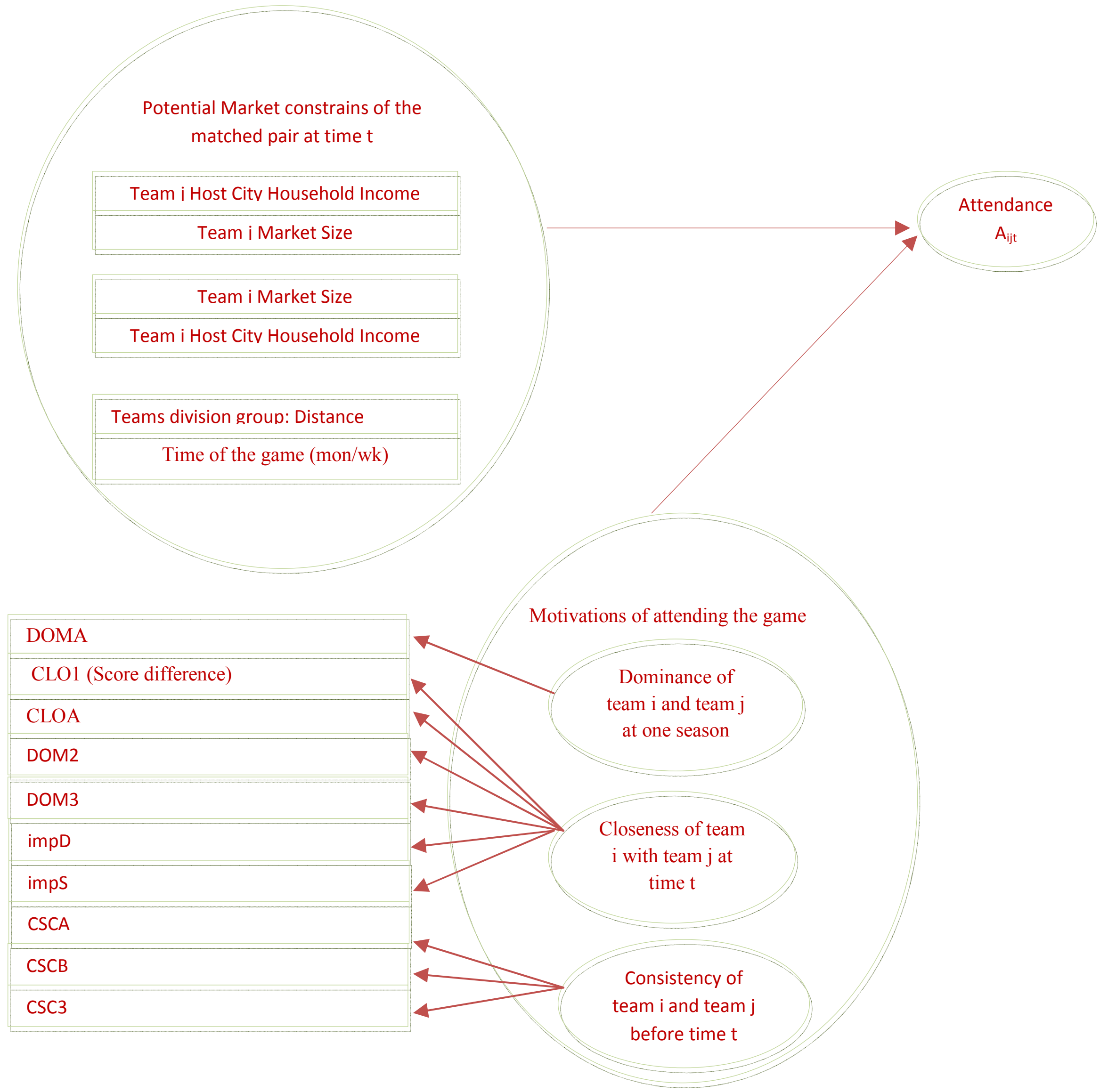


Figure 6.2 Variables that might relate to game attendance

Further exploration of the data shows that the schedules of the matched pairs are different. For example, some pairs only play in selected years/months, and histories of the matched pairs differ with each other as well. Thus the measures occur in irregular time points because each matched pair of teams has a set of Competitive Balance measures that are measured in different months in a season, and distributed differently across seasons. Therefore, the time element in the model should be treated as a random component.

As shown in Figure 6.1, it is more realistic to allow each matched pair to have its own change rate (slope) corresponding to the measures of the Competitive Balance. This allows different pairs to response differently to Competitive Balance measures, as well as other factors included in the model. For example, attendance at games between the ANA & NYA might be more sensitive to score difference than the attendance for other pairs of teams, and a small increase in the score difference can greatly increase game attendance. As Figure 6.1 suggests, it is more realistic to allow each matched pair to have its own average attendance value (intercept). For example, the NYN &CHN's attendance is higher than another other pairs after controlling for all other factors in the model.

To enable the model to handle the requirements discussed above, the conditional pair-wise growth model is selected. Specifically, the time unit is treated as a random component and is included in the first level of the model, and all other variables will be included in the second level of the model. When exploring the data in one season, the time unit will be the counts (labeled as 'counts') of games played between each matched pair in that season. When exploring the data over multiple seasons, the time unit will be the counts of games played since the first game of the matched pair of teams (labeled 'countnow'). Each time the matched pair of teams play a game, there will be a set of Competitive Balance measures, and a set of covariates (such as the date of the game, the league of the teams in the pair, etc.), as well as game attendance.

The model is:

Level 1 (time)

$Att_{pairid,countnow} = \beta_{0,pairid} + \beta_{1,pairid}$ (countnow) $+ \boldsymbol{\varepsilon}_{pairid,countnow}$ [6.1]

Level 2 (pair level)

$\beta_{0,pairid} = \gamma_{0,0} + \gamma_{0,1}$ (month) $+ \gamma_{0,2}$ (day/night) $+ \gamma_{0,3}$ (week of the day) $+ \gamma_{0,4}$ (teamslg) $+ \gamma_{0,5}$ (teamsdiv) $+ \gamma_{0,6}$ (Pophome) $+ \gamma_{0,7}$ (Popvst) $+ \gamma_{0,8}$ (Inchome) $+ \gamma_{0,9}$ (Incvst) $+ \gamma_{0,10}$ (CLO1) $+ \gamma_{0,11}$ (CLOA)$+ \gamma_{0,12}$ (DOMA)$+ \gamma_{0,13}$ (DOM2)$+ \gamma_{0,14}$ (DOM3)$+ \gamma_{0,15}$ (impD)$+ \gamma_{0,16}$ (impS)$+ \gamma_{0,17}$ (CSCA)$+ \gamma_{0,18}$ (CSCB)$+ \gamma_{0,19}$ (CSC3)$+ \gamma_{0,20}$ (year)$+ \gamma_{0,21}$ (year*year)$+ u_{0,pairid}$

$\beta_{1,pairid} = \gamma_{1,0} + \gamma_{1,1}$ (month) $+ \gamma_{1,2}$ (day/night) $+ \gamma_{1,3}$ (week of the day) $+ \gamma_{1,4}$ (teamslg) $+ \gamma_{1,5}$ (teamsdiv) $+ \gamma_{1,6}$ (Pophome) $+ \gamma_{1,7}$ (Popvst) $+ \gamma_{1,8}$ (Inchome) $+ \gamma_{1,9}$ (Incvst) $+ \gamma_{1,10}$ (CLO1) $+ \gamma_{1,11}$ (CLOA)$+ \gamma_{1,12}$ (DOMA)$+ \gamma_{1,13}$ (DOM2)$+ \gamma_{1,14}$ (DOM3)$+ \gamma_{1,15}$ (impD)$+ \gamma_{1,16}$ (impS)$+ \gamma_{1,17}$ (CSCA)$+ \gamma_{1,18}$ (CSCB)$+ \gamma_{1,19}$ (CSC3)$+ \gamma_{1,20}$ (year)$+ \gamma_{1,21}$ (year*year)$+ u_{1,pairid}$

where

$$\boldsymbol{\varepsilon}_{pairid,countnow} \sim N(0, \sigma^2)$$

$$\binom{u_{0,pairid}}{u_{1,pairid}} \sim N\left[\binom{0}{0}, \begin{pmatrix} \widehat{\tau_{0,0}} & \widehat{\tau_{0,1}} \\ \widehat{\tau_{1,0}} & \widehat{\tau_{1,1}} \end{pmatrix}\right]$$

Each continuous variable in the level two is centered at its grand mean, so that the interpretation of the fixed effects is more straightforward. Random effects are that components in the model that are related to the id of the matched pair of teams ($u_{0,pairid} + u_{1,pairid}$* countnow)

- $\beta_{0,pairid}$ is the average attendance for each matched pair
- $\beta_{1,pairid}$ is the average change rate (slope) with counts of games for each matched pair
- $\tau_{0,0}$ variance among random intercepts
- $\tau_{1,1}$ variance among random slopes
- $\tau_{1,0}$ Covariance between slopes and intercepts
- CLO1—CSC3 are the proposed Competitive Balance measures
- Teamslg is the league affiliations of the teams in the matched pair. The matched pairs are categorized into three groups by their league identities (same as chapter 5). When both teams are from AL, then teamslg

takes the value AL, when the teams are from the AL and the NL, team league takes the value 'Inter', otherwise it takes the value NL.

- Teamsdiv is the division affiliation of the teams in the matched pair. The matched pairs are categorized into six groups by their division identities (same as chapter 5). Teamsdiv has six values, and they are the six possible combinations of the teams' location. For example, East-East, East-West, etc. .
- Pophome/PopVst is the host city population of the home/visiting team in the matched pair. The value is abstracted from the U. S. Census Bureau's 2000 census data.
- IncHome/IncVst is the host city's average household income of the home/visiting team in the matched pair. The data also came from the U. S. Census Bureau's 2000 census data.
- σ^2 is the within pair variance. In section 6.1, this is a single value.

6.1.4 Results and discussion.

The model 6.1 is used to answer research questions related to inter-pair differences in game attendances and Competitive Balance. The object of constructing the model is to allow both initial attendance and attendance change rate to vary for each matched pairs. Each matched pair has multiple observations in a given period, and the count of the games is the index of the multiple observations.

A total of 433 pairs and 81,194 games are used in the analysis after excluding the records with missing values (a lot of old games do not have game attendance). Teams that only existed before the three division era are excluded from the analysis. The output is displayed in the Tables 6.1 to 6.4. The set of variables in the model significantly explained the variation in game attendances (Table 6.1). To find out how much variance is explained by the Competitive Balance measures, results based on two models are compared side by side (Table 6.2. Only the covariance parameter estimates are shown). The results on the left side generated by a model without the Competitive Balance measures, on the right side are results generated by the model with the Competitive Balance measures. By including the proposed Competitive Balance measures, the unexplained residual becomes smaller (σ^2 =93642339 verse σ^2 =1.07E+08), and the variance component of the change rate decreases from 2280.96 to 1296.43 (Table 6.2). The estimated τ values tell us that there is a variation in both the intercepts and slopes that potentially could be explained by a level 2 (pair level) covariate: $\begin{pmatrix} \widehat{\tau_{0,0}} & \widehat{\tau_{0,1}} \\ \widehat{\tau_{1,0}} & \widehat{\tau_{1,1}} \end{pmatrix}$=$\begin{pmatrix} 37884377 & -185891 \\ -185891 & 2280.96 \end{pmatrix}$. The total variance reduction rate is (2280.96 -1296.43)/ 2280.96 =43.16%. In other words, the proposed measures improved the fit of the change rates by reducing 43.16% of variance.

Table 6.1

Significant Test

Null Model Likelihood Ratio Test		
DF	Chi-Square	Pr > ChiSq
3	7564.46	<.0001

Table 6.2

Model Comparison

Covariance Parameter Estimates (RMEL)					
Before adding the Competitive Balance measures			After adding the Competitive Balance measures		
UN(1,1)	pairID	37884377	UN(1,1)	pairID	25408530
UN(2,1)	pairID	-185891	UN(2,1)	pairID	-114185
UN(2,2)	pairID	2280.96	UN(2,2)	pairID	1296.43
Residual σ^2		1.07E+08	Residual σ^2		93642339

When adding the Competitive Balance measures into the model, the simulation converges after four iterations. The Estimated Genetic Correlation Matrix shows the correlation between the *average attendances of matched pair (intercepts)* and the *attendance change rates of matched pairs (slopes).* They are negatively correlates at a level of -0.6291 (Table 6.3). The covariance tells us how the average attendance status and change rate are related. In current case, higher average attendances are accompanied by lower change rate.

Table 6.3

Correlations between Average Attendances of Matched Pair and the Attendance Change Rates

Estimated G Correlation Matrix				
Row	**Effect**	**allpairID**	**Col1**	**Col2**
1	Intercept	1	1	-0.6291
2	countnow	1	-0.6291	1

Note from Table 6.4, six of the ten measures of the Competitive Balance show significant effects on the attendance change rates, and seven of the ten Competitive Balance measures show significant effects on average attendance. For the closeness measures, game importance measures (impS and impD) are all significantly correlated with game attendance. This confirms that fans prefer to attend games with overall high importance (impS), and games with more importance differences (impD) can attract more fans. Attendance is more sensitive to overall game importance (impS) than the differences of importance level (impD). Pairs with 1.0 impS difference have change rates that differ by 9.78, and pairs which differ by 1.0 with respect to impD have change rates that differ by 1.2.

The average game attendance is positively correlated with the pair's previous year's winning percentages, hereafter referred as previous winning percentages (DOM2 and DOM3), but the change rate is negatively correlated with these measures. The parameter estimate of -3.51/ -3.66 indicates that the pairs which differ by 1.0 with respect to the Dom2/Dom3 have change rates that differ by -3.51/-3.66 as the number of games between the pair increases. For example, if two teams in a matched pair both have high previous winning percentages, then the average game attendance of this pair is higher than the pairs without two high previous winning percentages, but its attendance change rate is lower than pairs without two high previous winning percentages. It is true for a matched pair with a big gap between previous winning percentages. This result is consistent with previous findings: average game attendance is negatively associated with attendance change rate.

Neither the pair's initial attendance, nor the pair's attendance change rate is sensitive to the measures which are related to game scores (CLO1 & CLOA). Results indicate that the game attendance is not affected much by the score difference of current game, nor does runs scored in the game.

The dominance of one team over another in the matched pairs has no significant effects on the game attendance. Whether the weak team wins or the stronger team wins does not show any effect on the average attendance, and it has no effect on change rate (Table 6.4, DOMA). When I replace upset game ratio (DOMA) with the absolute value of the upset game ratio (abs(DOMA-0.50)), its effects on average attendance and attendance changing rate are not significant either.

While game attendance remains constant with the change of dominance measure, the consecutive wins of one team in a pair has a negative effect on the change rate (CSCB). CSCA is another consistency measure that is based on the pair's three-year winning percentages. CSCA has a significant positive effect on the average game attendance, but the attendance declines a little bit as the number of games between the pair increase. CSC3 is a measure based on two years of winning percentages changes, and it has the same effects on attendance as CSCA.

In terms of average attendance (fixed effects), day time game attendance is lower than night time attendance, and inter league and the NL pairs have better attendance than the AL pairs. The locations of the matched pairs affect on game attendance as well. Games which have matched division teams such as East-West, West-West, and East-East are more likely to have better average attendance than other types of matched pairs. Tuesdays and Wednesdays have lower attendance than other days of the week.

Figure 6.3 displays the predicted (blue line) and observed attendance based on the parameter estimates generated by the model. Model 6.1 provides estimated results for average game attendance for each matched pair as well as attendance change rate, sometimes known as growth rate, of each matched pair. The attendance change rate indicates how attendance changes as the number of games played by the pair increases. Hereafter, I refer to this within pair attendance change rate as change rate. In the data there are hundreds of matched pairs, and each matched pair has its own average attendance and change rate. It will be too much to report all the results, so I randomly selected two pairs, and plotted the predicted and observed game attendance according to the pair's estimated average game attendance and attendance change rate. The trends of observed and predicted values are consistent.

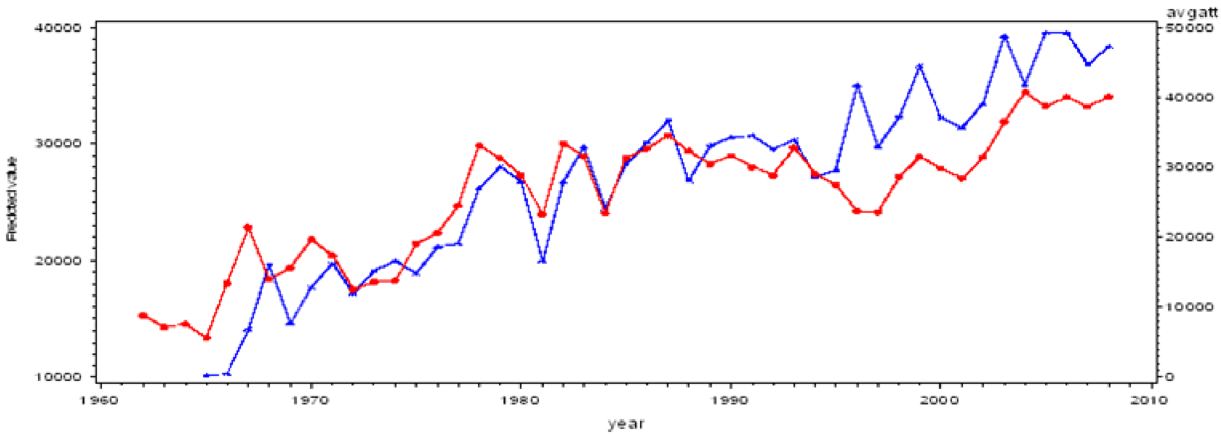
Table 6.4

Results of Model 6.1

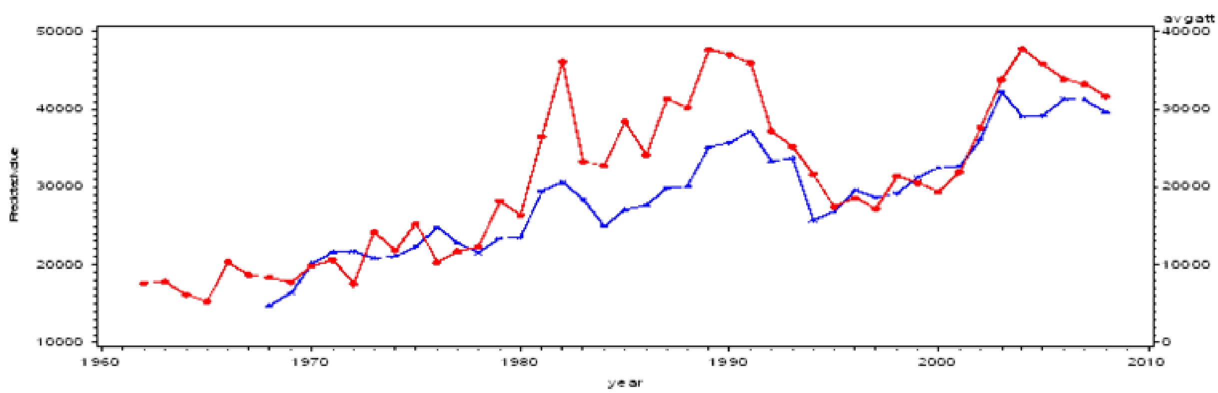
Solution for Fixed Effects									
Effect	3. Day of week	13. Day/night indicator	teamsLg	teamsdiv	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept					13590	720.67	432	18.86	<.0001
countnow					12.9862	5.5551	8.10E+04	2.34	0.0194
VAR3	Fri				4422.5	157.37	2192	28.1	<.0001
VAR3	Mon				1183	175.18	2192	6.75	<.0001
VAR3	Sat				7913.48	159.05	2192	49.75	<.0001
VAR3	Sun				5635.26	179.06	2192	31.47	<.0001
VAR3	Thu				521.72	172.99	2192	3.02	0.0026
VAR3	Tue				-108.79	159.02	2192	-0.68	0.494
VAR3	Wed				0	.	.	.	.
VAR13		Day			-560.51	123.36	411	-4.54	<.0001
VAR13		Night			0	.	.	.	.
teamsLg			AL		-4996.65	694.71	16	-7.19	<.0001
teamsLg			Inter		1299.93	656.41	16	1.98	0.0651
teamsLg			NL		0	.	.	.	.
teamsdiv				Cent	-1590.71	520.38	135	-3.06	0.0027

teamsdiv				East	1494.56	511.59	135	2.92	0.0041
teamsdiv				EastCent	-1419.15	498.72	135	-2.85	0.0051
teamsdiv				EastWest	2750.79	428.73	135	6.42	<.0001
teamsdiv				West	1807.77	459.1	135	3.94	0.0001
teamsdiv				WestCent	0	.	.	.	.
cyear					31517	2332.04	8.10E+04	13.51	<.0001
cysq					-7.8668	0.5869	8.10E+04	-13.4	<.0001
cmonth					153.35	28.9234	8.10E+04	5.3	<.0001
cpophome					0.000272	0.000036	8.10E+04	7.56	<.0001
cpopvst					0.000083	0.000036	8.10E+04	2.31	0.0207
cincvst					0.1462	0.02844	8.10E+04	5.14	<.0001
cinchome					0.01949	0.02846	8.10E+04	0.68	0.4934
cclo1					14.4553	10.4929	8.10E+04	1.38	0.1683
ccloa					2441.09	1826.72	8.10E+04	1.34	0.1814
cdoma					-90.105	262.73	8.10E+04	-0.34	0.7316
cdom2					16995	433.6	8.10E+04	39.2	<.0001
cdom3					28019	540.23	8.10E+04	51.87	<.0001
cimpd					4711.52	288.66	8.10E+04	16.32	<.0001
cimps					7250.4	287.89	8.10E+04	25.18	<.0001
ccsca					26681	2881.25	8.10E+04	9.26	<.0001
ccscb					-670.08	146.01	8.10E+04	-4.59	<.0001
ccsc3					24653	850.8	8.10E+04	28.98	<.0001
countnow*VAR3	Fri				0.397	0.1913	8.10E+04	2.08	0.0379
countnow*VAR3	Mon				0.3151	0.2113	8.10E+04	1.49	0.1358
countnow*VAR3	Sat				0.2785	0.1955	8.10E+04	1.42	0.1543
countnow*VAR3	Sun				1.6175	0.2158	8.10E+04	7.49	<.0001
countnow*VAR3	Thu				-0.2715	0.2135	8.10E+04	-1.27	0.2036
countnow*VAR3	Tue				0.00844	0.1939	8.10E+04	0.04	0.9653
countnow*VAR3	Wed				0	.	.	.	.
countnow*VAR13		Day			-1.0571	0.1506	8.10E+04	-7.02	<.0001
countnow*VAR13		Night			0	.	.	.	.
countnow*teamsLg			AL		26.9212	5.3097	8.10E+04	5.07	<.0001
countnow*teamsLg			Inter		14.2293	14.1512	8.10E+04	1.01	0.3146
countnow*teamsLg			NL		0	.	.	.	.
countnow*teamsdiv				Cent	-20.9824	4.3638	8.10E+04	-4.81	<.0001
countnow*teamsdiv				East	-21.7177	4.338	8.10E+04	-5.01	<.0001
countnow*teamsdiv				EastCent	-8.6937	5.3797	8.10E+04	-1.62	0.1061
countnow*teamsdiv				EastWest	-23.8145	4.3061	8.10E+04	-5.53	<.0001
countnow*teamsdiv				West	-12.2261	4.296	8.10E+04	-2.85	0.0044
countnow*teamsdiv				WestCent	0	.	.	.	.
countnow*cyear					-55.656	3.3321	8.10E+04	-16.7	<.0001
countnow*cysq					0.01409	0.000841	8.10E+04	16.76	<.0001
countnow*cmonth					-0.1155	0.03533	8.10E+04	-3.27	0.0011
countnow*cpophome					1.03E-06	0	8.10E+04	Infty	<.0001
countnow*cpopvst					9.52E-07	0	8.10E+04	Infty	<.0001
countnow*cinchome					-0.00117	0.000236	8.10E+04	-4.97	<.0001
countnow*cincvst					-0.00116	0.000236	8.10E+04	-4.91	<.0001
countnow*cclo1					-0.00585	0.01308	8.10E+04	-0.45	0.6547
countnow*ccloa					0.08131	2.1856	8.10E+04	0.04	0.9703
countnow*cdoma					-0.3747	0.3382	8.10E+04	-1.11	0.2678
countnow*cdom2					-3.507	0.5395	8.10E+04	-6.5	<.0001

countnow*cdom3					-3.6617	0.6588	8.10E+04	-5.56	<.0001
countnow*cimpd					1.2055	0.3238	8.10E+04	3.72	0.0002
countnow*cimps					9.7882	0.421	8.10E+04	23.25	<.0001
countnow*ccsca					-2.9651	3.3881	8.10E+04	-0.88	0.3815
countnow*ccscb					-0.718	0.1895	8.10E+04	-3.79	0.0002
countnow*ccsc3					-8.3453	0.9898	8.10E+04	-8.43	<.0001



Average predicted (blue) and observed attendance (ANA & BOS)



Average predicted (blue) and observed attendance (ANA & OAK)

Figure 6.3 Predicted & observed attendance for random selected pairs

In conclusion, average attendance is negatively associated with change rate. For the ten Competitive Balance measures, six of them are significantly associated with the game attendance.

Adding Competitive Balance measures greatly improved the estimates of attendance change rate of matched pairs. As shown in section 6.1, it explains 43% of the covariance of the attendance change rate of different pairs.

Attendance is not sensitive to score differences as score related measures (clo1,cloa) have no effect on attendances. Game importance measures based on dynamic winning percentages (impD, impS) have a positive effects on average attendance as well as change rate. Closeness measures derived from the previous year's winning percentages (dom2, dom3) have positive effects on average attendance, but a negative effect on change rate. A bigger difference between real-time winning percentage and previous winning percentages (CSC3) associates with a higher average attendance, but a lower change rate.

Consecutive wins (csc3) have negative effects on both average game attendance and game attendance change rate. Dominance of one team over another team in the matched pairs has no significant effects on game attendance. Dominance measured by lower rank team wins (DOMA) has no effect on game attendances. This indicates that fans don't really care if the weak team wins, but the possibility of lower rank team wins.

6.2 Exploring the Structure of the Variance Covariance Matrix within Matched Pairs of Teams in a Single Season

6.2.1 Research question.

In model 6.1, heteroscedasticity was introduced to the error covariance matrix by fitting the time element in the random portion of the model (game count is used as the time element). Thus, each pair of teams not only has its own average game attendance, but also its own change rate. However, the residual of the same pair of teams (within pair residuals) over time is assumed to be independent after controlling the time element (game count). That is, the residuals of the same pair of teams (after controlling for the linear effect of game counts) are assumed to be independent, and it is assumed to follow a normal distribution ($\boldsymbol{\varepsilon}_{\text{pairid,countnow}} \sim N(0, \sigma^2)$). This assumption may not be realistic, especially when the games of the same pair are played in the same season.

Therefore, the goal of this section is to study the relationship between the Competitive Balance measures and game attendance under a more relaxed assumption by introducing within-subject heterogeneity. That is, to allow the residuals of the same pair (after controlling for the game count effect) to be correlated through the within pair error variance covariance matrix ($\boldsymbol{\varepsilon}_{\text{pairid,countnow}} \sim N(0, \Sigma)$).

First, I select a within pair error variance-covariance matrix for the model. Because it is reasonable to assume that for a given pair of teams, the consecutive games are more likely to be strongly correlated than games played far apart. For instance, the second game between the Minnesota & Texas is more likely to be associated with the third game between the Minnesota & TEXAS than the 9^{th} game. Therefore an autoregressive with a lag of one (*AR(1)*) is chosen as the form of the within pair variance-covariance matrix. To make sure that *AR(1)* is a better choice, *Compound Symmetry* and *Unstructed* variance-covariance matrix forms are tested and compared side by side with AR(1) in Table 6.5.

The goodness of fit statistics of the three forms of error variance-covariance matrixes are presented in Table 6.5. Results show that the model assuming AR(1) error variance-covariance structure converges much faster than the model using another two types of variance-covariance structures(test procedures are not shown). In addition, AR(1) provides a better fit of the data(Table 6.5), for it has a smaller value for all the four goodness of fit statistics. Therefore, I have supporting evidences to increase the complexity of **Σ** by adding non-zero off-diagonal elements.

Table 6.5

Choose within Pair Variance-Covariance Matrix Structure

Assumption	AIC	BIC	-2RLL	AICC	Iteration
AR(1)	**48579.0**	**48593.6**	**48571.0**	**48579.0**	**3**
Compound Symmetry	48840.9	48859.2	48830.9	48840.9	7
Unstructured	54638.4	55342	54252.4	54672.8	* Unable to make hessian positive definite

6.2.2 Model constructions.

The model is:

Level 1 (time)

$Att_{pairid,count} = \beta_{0,pairid} + \beta_{1,pairid}(count) + r_{pairid,count}$

Level 2 (pair level)

$\beta_{0,pairid} = \gamma_{0,0} + \gamma_{0,1}(month) + \gamma_{0,2}(day/night) + \gamma_{0,3}(week\ of\ the\ day) + \gamma_{0,4}(teamslg) + \gamma_{0,5}(teamsdiv) + \gamma_{0,6}(Pophome) + \gamma_{0,7}(Popvst) + \gamma_{0,8}(Inchome) + \gamma_{0,9}(Incvst) + \gamma_{0,10}(CLO1) + \gamma_{0,11}(CLOA) + \gamma_{0,12}(DOMA) + \gamma_{0,13}(DOM2) + \gamma_{0,14}(DOM3) + \gamma_{0,15}(impD) + \gamma_{0,16}(impS) + \gamma_{0,17}(CSCA) + \gamma_{0,18}(CSCB) + \gamma_{0,19}(CSC3) + u_{0,pairid}$

$\beta_{1,pairid} = \gamma_{1,0} + \gamma_{1,1}(month) + \gamma_{1,2}(day/\ night) + \gamma_{1,3}(week\ of\ the\ day) + \gamma_{1,4}(teamslg) + \gamma_{1,5}(teamsdiv) + \gamma_{1,6}(Pophome) + \gamma_{1,7}(Popvst) + \gamma_{1,8}(Inchome) + \gamma_{1,9}(Incvst) + \gamma_{1,10}(CLO1) + \gamma_{1,11}(CLOA) + \gamma_{1,12}(DOMA) + \gamma_{1,13}(DOM2) + \gamma_{1,14}(DOM3) + \gamma_{1,15}(impD) + \gamma_{1,16}(impS) + \gamma_{1,17}(CSCA) + \gamma_{1,18}(CSCB) + \gamma_{1,19}(CSC3) + u_{1,pairid}$

where

$$\boldsymbol{\varepsilon}_{pairid,countnow} \sim N(0, \Sigma)$$

$$\begin{pmatrix} u_{0,pairid} \\ u_{1,pairid} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \widehat{\tau_{0,0}} & \widehat{\tau_{0,1}} \\ \widehat{\tau_{1,0}} & \widehat{\tau_{1,1}} \end{pmatrix}\right]$$

$\beta_{0,pairid}$ is the average attendance of each pair;

$\beta_{1,pairid}$ is the average change rate with the counts of games between the pair;

$\tau_{0,0}$ is the variance among random intercepts

$\tau_{1,1}$ is the variance among random slopes

$\tau_{1,0}$ is the covariance between slopes and intercepts.

Because the within pair correlation is more substantial within a season, the data used in the following research is limited to games in 2008. All the 243 matched pairs and 2428 game records are included in the analysis. The basic research unit is still matched pairs in a season. The count of the games played by each matched pair of teams is used as the time element in the model. The structure of the within-pair error covariance matrix is specified using the *repeated* statement in the mixed procedure. The *game count* is also used as a predictor in the attendance model, thus it is used as a class variable and continuous variable in the same model. As in section [6.1], each pair has its own average game attendance and change rate.

6.2.3 Interpreting the results of the conditional growth model.

The simulation converges after three iterations. A rapid convergence could be due to the perfectly balanced data set, and low degree of collinearity among the covariates. The estimate of the autoregressive correlation in the covariance parameter estimates Table is 0.4392, which means the residual correlation within the same matched pairs (after controlling all the variables) is 0.4392. Clearly the correlation of the within pair residuals are quite strong.

Table 6.6 presents the estimates of fixed and random effects. The upper part of the Table showss the relationship between the covariates and average game attendance. For example, the overall game importance (impS) has an estimate of -3827.66 (Table 6.6, line 32) and a standard error of 2107.18. Note that the standard error is almost equal to the estimate itself, it is fair to conclude that there is no relationship between average game attendance and overall game importance (imps) after controlling all other variables in the model. In terms of slope, the overall game importance (impS) shows a significant effect. The parameter estimate of 2667.27 ($p<0.001$) (Table 6.6, line 67) indicates that pairs which differ with 1.0 in overall game importance (impS) have attendance change rates differing by 2667.27. The effect of game important difference (impD) is similar to overall importance (impS), it is significantly associated with the change rate, but not the average attendance. Closensess measures which are based on previous winning percentages (DOM2, DOM3) have positive effects on average attendance, but have no effect on the attendance change rate. Like what was shown in section 6.1, the closeness measures that based on game scores show no effect on attendance.

Again, similar to the results in section 6.1, the dominance measure (DOMA) has no effect on average attendance and the change rate. When I replace upset game ratio (DOMA) with the absolute value of the upset game ratio (abs(DOMA-0.50)), the results are the same. A larger difference between real-time winning percentage and previous winning percentages (CSC3) is associated with a higher average attendance, but a lower change rate. The other two consistency measures have no effect on attendance.

For the 2008 season, day time attendance is not significantly different from night time attendance. The average weekday attendance is lower than weekend attendance. The inter league pairs and the NL pairs have better attendance than the AL pairs. Games for the Center-Center division combination have better attendance than West-Center division combination.

Table 6.6

Results of Model 6.2

	Solution for Fixed Effects-Initial status									
	Effect	3. Day of week	13. Day/night indicator	teamsLg	teamsdiv	Estimate	Standard Error	DF	t Value	Pr > |t|
1	Intercept					31015	1210.4	273	25.62	<.0001
2	count					1392.48	591.77	2095	2.35	0.0187
3	VAR3	Fri				6595.83	777.78	1141	8.48	<.0001
4	VAR3	Mon				1539.11	797.69	1141	1.93	0.0539
5	VAR3	Sat				9132.95	820.45	1141	11.13	<.0001
6	VAR3	Sun				6058.82	927.49	1141	6.53	<.0001
7	VAR3	Thu				-466.57	738.23	1141	-0.63	0.5275
8	VAR3	Tue				951.21	616.63	1141	1.54	0.1232
9	VAR3	Wed				0	.	.	.	.
10	VAR13		Day			434.18	537.64	261	0.81	0.4201
11	VAR13		Night			0	.	.	.	.
12	teamsLg			AL		-6699	863.5	273	-7.76	<.0001
13	teamsLg			Inter		-41.4843	1248.43	273	-0.03	0.9735
14	teamsLg			NL		0	.	.	.	.
15	teamsdiv				Cent	4350.43	1422.44	273	3.06	0.0024
16	teamsdiv				East	-56.2159	1517.24	273	-0.04	0.9705
17	teamsdiv				EastCent	1319.3	1280.43	273	1.03	0.3038
18	teamsdiv				EastWest	-3029.83	1308.4	273	-2.32	0.0213
19	teamsdiv				West	-1003.23	1590.11	273	-0.63	0.5286
20	teamsdiv				WestCent	0	.	.	.	.
21	cpophome					0.000997	0.000069	2095	14.36	<.0001
22	cpopVst					0.000211	0.000069	2095	3.06	0.0022
23	cinchome					0.1575	0.04945	2095	3.18	0.0015
24	cincVst					0.04367	0.05119	2095	0.85	0.3937
25	cmonth					1158.18	240.38	2095	4.82	<.0001
26	cclo1					58.3371	52.0765	2095	1.12	0.2627
27	ccloa					-8202.75	8525.95	2095	-0.96	0.3361
28	cdoma					906.2	1771.11	273	0.51	0.6093
29	cdom2					24168	3398.83	2095	7.11	<.0001
30	cdom3					55479	5178.99	273	10.71	<.0001
31	cimpd					-1770.97	1585.78	2095	-1.12	0.2642
32	cimps					-3827.66	2107.18	2095	-1.82	0.0694
33	ccsca					-19994	21638	273	-0.92	0.3563
34	ccscb					-88.003	779.58	2095	-0.11	0.9101
35	ccsc3					41720	5862.12	2095	7.12	<.0001

36	Solution for Fixed Effects –Growth rates									
37	Effect	3. Day of week	13. Day/night	teamsLg	teamsdiv	Estimate	Standard Error	DF	t Value	Pr > |t|
38	count*VAR3	Fri				-359.2	99.0597	2095	-3.63	0.0003
39	count*VAR3	Mon				-300.56	101.58	2095	-2.96	0.0031
40	count*VAR3	Sat				-303.61	101	2095	-3.01	0.0027
41	count*VAR3	Sun				-170.45	108.11	2095	-1.58	0.115
42	count*VAR3	Thu				118.84	92.3416	2095	1.29	0.1982
43	count*VAR3	Tue				-150.2	75.6547	2095	-1.99	0.0472
44	count*VAR3	Wed				0	.	.	.	.
45	count*VAR13		Day			-135.23	67.2831	2095	-2.01	0.0446
46	count*VAR13		Night			0	.	.	.	.
47	count*teamsLg			AL		207.49	95.6182	2095	2.17	0.0301
48	count*teamsLg			Inter		-0.4556	382.14	2095	0	0.999
49	count*teamsLg			NL		0	.	.	.	.
50	count*teamsdiv				Cent	-27.7226	208.5	2095	-0.13	0.8942
51	count*teamsdiv				East	79.2667	214.1	2095	0.37	0.7112
52	count*teamsdiv				EastCent	138.89	250.9	2095	0.55	0.5799
53	count*teamsdiv				EastWest	170.75	248.05	2095	0.69	0.4913
54	count*teamsdiv				West	28.1828	213.33	2095	0.13	0.8949
55	count*teamsdiv				WestCent	0	.	.	.	.
56	count*popHome					-3.03E-07	8.72E-06	2095	-0.03	0.9723
57	count*popVst					-9.45E-06	8.67E-06	2095	-1.09	0.2755
58	count*IncHome					-0.00669	0.006169	2095	-1.09	0.2779
59	count*IncVst					-0.00079	0.006406	2095	-0.12	0.9013
60	count*month					-141.51	33.6618	2095	-4.2	<.0001
61	count*cclo1					-7.9753	6.3613	2095	-1.25	0.2101
62	count*ccloa					1039.11	1042.67	2095	1	0.3191
63	count*cdoma					50.4188	272.32	2095	0.19	0.8531
64	count*cdom2					489.45	403.84	2095	1.21	0.2257
65	count*cdom3					-931.9	665.28	2095	-1.4	0.1614
66	count*cimpd					1227.13	325.81	2095	3.77	0.0002
67	count*cimps					2667.27	472.06	2095	5.65	<.0001
68	count*ccsca					4696.54	3120.35	2095	1.51	0.1324
69	count*ccscb					-12.7742	194.79	2095	-0.07	0.9477
70	count*ccsc3					-8628.74	2681.66	2095	-3.22	0.0013

* The prefix ‘c’ of each variable shows that the variables are centered before they enter the model.

6.3 Chapter Summary and Conclusions

In this chapter I investigated the relationship between game attendance and the set of Competitive Balance measures. Using the number of games between the matched pairs as the time element, I use a growth model to study the relationship between attendance and the variables of interest. The advantage of the growth model is that it allows each matched pair to have its own initial attendance and change rate.

The Competitive Balance measures explain over 43% of the variance in the attendance change rate [6.1]. A side by side comparison for section 6.1 and 6.2 is present in Table 6.7.

Table 6.7

Summary Tables for the Multi-season and Single Season Studies

		Results of [6.1]		Results of [6.2]	
		Average	Growth	Average	Growth
Measures based on current game scores	clo1, cloA	not sign.	not sign.	not sign.	not sign.
Measures based on previous wp	dom2, dom3	**positive**	**negative**	**positive**	not sign.
Importance measures base on current wp	impS,impD	**positive**	**postive**	not sign.	**postive**
Dominance measure (rate of low rank wins)	DOMA	not sign.	not sign.	not sign.	not sign.
Consecutive wins	Cscb	**negative**	**negative**	not sign.	not sign.
Gap between current wp and previous wp	csc3	**positive**	**negative**	**positive**	**negative**
Measures based on the variance of wps	Csca	**positive**	not sign.	not sign.	not sign.
League, div, population, income, year, ysq, month, week, day/night,count		Sign.		Sign.	

Results shows that measures based on current game scores do not affect game attendance. The two closeness measures--clo1, cloA—seem have no significant effects on game attendances in either multiple seasons, nor in a single season.

Fans are sensitive to previous winning percentages. Measures composed from pair's previous winning percentages—dom2, dom3—have positive effects on game average attendance in multi-season study as well as single season study. The average attendance favors the games between very unbalanced pair—the matched pair with a large gap between teams' previous winning percentages (dom2). The increasing of the gap between the pair's previous winning percentage indicates a higher game attendance in both single season and multi-season. At the same time, the average attendance of the strong pair—the matched pair consists of two teams with high previous winning percentages—is higher than the pair that both teams have lower winning previous percentages (dom3). The results are true for both single and multi-season.

However, the game attendance change rate is the opposite of average game attendances for the strong pair in the multi-season study. The increasing rate of attendance is higher for less stronger pair as the number of games played between the two teams accumulates. Similarly, the attendance change rate for the very unbalanced pair is lower than other matched pairs in the multiple season study.

Research confirms that fans care more about important games. In other words, fans like to go to games that can affect teams' rank. The game importance measures are derived on the how current game outcome can affect a team's stand in terms of winning percentages within its division/league. Both important measures have significant positive effect on average game attendance as well as change rate in multi-season, and it has the maximum magnitude effects among all other factors in single season attendance change rate. The correlation coefficients of impD and impS are the largest of all significant factors.

Fans don't respond vigorously to dominance measure either. Higher wining rate of the underdog in the matched pair cannot serve as an indicator of a higher game turn-out. Similarly to score measures, the dominance measure which is composed by the number of games underdog win has no significantly effect in either cross season data check, nor the single season data check.

If we consider the higher number of underdog wins as an indicator of higher Competitive Balance, this result seems to work against UOH. However, when combing the results of consecutive wins, it makes perfect sense. Consecutive wins (CSCB) impose negative influences on average game attendance as well as attendance change rate in cross season observations. Thus, the explanation can be that fans care more about the variation of the game results, but not the strong or weak team wins all the games. The consecutive wins measure does not have significant results on single season observations.

Fans are sensitive to the fluctuation of team performance as well. The gap between teams' current and last year's winning percentages have positive effect on average winning percentages in single and multi-season study (CSC3), but negative effects on attendance change rate. Also, the fluctuation of team's performance of last 3 years has positive effect on average game attendance in multi-season study.

Time of the game, location and league identity are show significant different across their levels/categories. For example, Inter-league pairs have better average attendance than the NL and the AL pairs. And games which have matched divisions such as East-West, West-West, and East-East are more likely to have better average attendances than other types of matched pairs.

These results support Davis' (2009) study which assumes fans response differently to teams win, and estimates separate attendance models for each Major League Baseball team. Davis used generalized autoregressive conditional heteroskedasticity (GARCH) models and finds that winning is an important determinant of attendance, and finds that interleague games have higher attendance.

CHAPTER 7: CONCLUSION

7.1 Review of Research Objects

Over the past several decades, Competitive Balance has become an increasingly prominent topic in the economics of professional sports in general and of Major League Baseball in particular (Mizak, Neral, & Stair, 2007). The literature on Competitive Balance is large and contentious because researchers disagree about how to properly measure Competitive Balance and disagree about the proper focus of research in this area (Humphreys, Watanabe, 2010). Given that Competitive Balance is multidimensional (Sanderson, 2002), many existing studies focus on only one dimension, and the commonly used measures are focus on individual teams or leagues. This research proposes a set of measures that address different dimensions of Competitive Balance and focus on matched pair in a season or seasons, and present the patterns of Competitive Balance with the proposed measures, and study the relationship between Competitive Balance measures and game attendances.

7.2 Procedures and Findings

Season 2008 is used for all the single season studies across this dissertation. Even if not reported, researcher did the same test on other seasons as well, and the results support the conclusions derived by 2008 data. For example, in section 4.3, only the AL 2008 results are reported, but the studies of other seasons generate similar results as well.

Factor analysis shows that the proposed Competitive Balance measures are multidimensional. The proposed closeness measures capture team capabilities in Major Baseball League. Team capability estimation base on complex model has better fit but more discrepancy with team winning percentage ranking. Compared with winning percentage, score difference not only considers games win/loss outcomes, but also how teams differ. This research found that team rankings based on winning percentages are more about team performances, while the rankings derived from score differences focus more on team capabilities. The ranking estimations using score difference are resistant to random influences such as the schedule of the games as well as opponent arrangements, because these estimations are based on score differences using matched pair of teams as the basic research objects.

The increased annual game attendance is accompanied by the improved Competitive Balance structural similarity over time. This research uses Gower's general similarity coefficient and Root-Mean-Square-Error (RMSE) as an integrated indictor of the Competitive Balance structural similarity among the matched pairs, as well as the diversity of the similarity structure. The Gower's general similarity coefficient has an increasing trend since 1941, and is strongly correlated with annual game attendance. The trend of Competitive Balance structural similarity reflects the league expansions in baseball history. The attendance increased after the AL and the NL's expansions, as well as after the introduction of inter-league games the in 1997.

When matched pairs are spatially proximate, they are more compatible in closeness dimension. Research has found that games played by two west division teams or two east division teams have higher closeness values than games played by other division combinations. Also, Profile Analysis shows that the Competitive Balance profiles for interleague games (one team from the AL and another team from the NL) are significantly different from intra-league games (players in the game belong to the same league). The interleague games are often played by two teams that have higher winning percentages in the previous year.

The proposed Competitive Balance measures explained over 43% of the variance in the attendance change rate. Yet not every dimension of the Competitive Balance measures has a strong correlation with game attendances.

The measures of consistency clearly show fans respond vigorously to changes. If the teams in a matched pair have big differences in current and previous winning percentage, this pair has higher average attendance and lower change rate in multi-season study. The larger fluctuation of three years winning percentages comes with higher average attendances in multi-season study. In addition, the consecutive wins work against both average game attendance and attendance change rate in multi-season/single season study.

The measure of dominance says that fans are not specific about whether or not the underdogs win the games. Empirically, we all observed that the legendary of New York Yankees attract great amount of attendances by win all its game.

The measures in closeness tell different story. Game important measures tell us that fans care about games that can potentially change the stand of the team in its division/league. Games between stronger pairs have higher average attendance and lower change rate in multi-season study. Both score difference and the score sum in a game have no significant effects on current game attendances.

Interleague games attract more attendances than intra-league games. This is consistent with previous findings that interleague game players often have higher winning percentages. The matched pairs that has different league/location identifies differ significantly across different groups/levels.

7.3 Contributions to Researchers and Practitioners

7.3.1 Contributions to researchers.

Levitt, an economist and a professor in University of Chicago, states that Economics is, above all, a science of measurement. He believes that if researchers look at data in the right way, then they can explain riddles that otherwise might have seemed impossible (Levitt& Dubner,2005). This is exactly what this research is trying to achieve. As Sanderson (2002) pointed out in his paper ‘The many dimensions of Competitive Balance’, there has not emerged a uniform, one-size-fits-all approach, or set of rules to resolve in its different dimensions. This study confirmed the multidimensionality of Competitive Balance and proposed measures for different dimensions. By examining the dimensions of Competitive Balance together, this research provides a more whole picture of how different dimensions of Competitive Balance associate with game attendance. By doing so, the current research helps to solve the riddle of whether the Competitive Balance has been increasing or decreasing, or the riddle of whether fans care about Competitive Balance. As mentioned in the introduction chapter, researchers often have conflicted views about this issue. For example, Mizak (2007) measured the degree of competitiveness by measuring the turnover in standings from one year to the next, and found that the Competitive Balance has declined since the 1990s in both the AL and the NL, whereas Sherony and his colleagues (Sherony, Haupert, & Knowles, 2000) measured the variance of winning percentages and concluded that both the AL and the NL became more competitive over time. Eckard (2001a) found that the Competitive Balance decreased in the AL, but improved in the NL. This research helps to clarify the increase -decrease confusion by showing that Competitive Balance is multidimensional, and that it is hard to represent Competitive Balance by one single measure. When using different measures, researchers come to different conclusions about Competitive Balance because they measure different aspects of Competitive Balance. Corresponding to the multidimensionality, the research concludes that fans responded differently to different dimensions in Competitive Balance. Fans are more sensitive to measures that relate to change of performances. In summary, this research shows the multidimensionality of Competitive Balance is hard to represent by a single measure.

Another contribution of this research is that it proposes a matched pair approach in studying sports that involves more than one team in a game. Using matched pair as the basic research unit helps to avoid pitfalls in analysis association relationships. In previous research, individual teams always acts as the basic research unit. When using teams as the basic research unit it violates the assumptions of OLS, which requires an independently identically distributed sample (iid.) and assumes that the regression residual is normally distributed with a zero mean. Sample independency is also preferred by other types of regressions that are often used to check the correlations among variables. For example, in the studies about winning percentages, the research object is an individual team's winning percentage. However, in MLB, all the matched pairs have repeated measures because the paired teams often played against each other more than once in a given season. Therefore, using team-based measures in a simple OLS regression or logistic regression is problematic because any unexplained pair-specific time-invariant effect in the residuals will create a correlation across occasions. When ignoring unexplained time invariant effects, we will find the outcome may have a different precision or reliability at different occasions. However, there is no need to assume the independence of the teams in the same game when using matched pair as the basic research object. In addition, we can examine the dependencies of the games of the same/different matched pairs easily.

Rather than use the conventional logistic or OLS regression, this study applied growth models to study the relationship between Competitive Balance and attendance. This approach further relaxes the assumption of variance and covariance structures between/within the matched pairs. Compared with conventional models, the current research model provides a better model fit. By constructing growth models, this research allows each matched pair to have their own average attendance and change rates as well as correlated residual structures for games played by the same matched pair.

Proposed measures are no longer a single summarized value. Most of the proposed measures in this study are calculated with real-time information, such as the game importance, consecutive wins, score differences, etc. They are capable of reflecting real-time changes during a season. In addition, to my knowledge, measures like game importance are not considered in existing studies of Competitive Balance, and this research shows that game importance deserve serious attentions.

By using matched pair approach and dynamic Competitive Balance measures and adding more flexibility in an estimating procedure, this research helps to compensate the influences of a biased game schedule. By considering the game importance and real-time as well as previous winning percentages of the matched pair, the estimates and conclusions are less affected by the game arrangement itself and the parameter estimates are more reliable than treating all wins equal and completely ignoring dependences among games. Therefore, conclusions regarding Competitive Balance measures and game attendances are more reliable as well.

7.3.2 Contributions to practitioners.

In terms of suggestions to practitioners, one of the most well-known solution packages is provided by the Blue Ribbon Panel on Baseball economics in 2000 (Levin, Mitchell, Volcker & Will, 2000). The Blue Ribbon Panel on Baseball economics is a group of commissioners who aimed to investigate how to set up a healthy economic environment of Baseball. Panel members found a strong correlation between high payrolls and success on the field in Major League Baseball, and proposed to reform the Baseball industry so that each team's success on the field would be determined by the skill of the players and people who conduct the business. The solutions provided by the panel focus on finance, draft and franchise relocation, and they suggest to use $\frac{\text{payroll Quartile I}}{\text{payroll Quartile IV}} \approx \frac{2}{1}$ as an indicator of a durable Competitive Balance.

Like the Blue Ribbon Panel, current research is also aimed at investigating Competitive Balance issues in Baseball, but the solutions suggested in this section will focus on team performance and schedule rather than finance issues such as tax, revenue, or fund redistribution. In addition, the Blue Ribbon Panel report mentioned that "some people suggest that the industry, from a competitive perspective, would be better off eliminating its weakest two franchises", and the panel member suggest there is no immediate need for contraction for new franchise if the recommendations outlined in this report are implemented. Unlike the Blue Ribbon Panel's suggestions, I found that league expansions are always accompanied by an increasing of Competitive Balance structural similarity and an increasing of game attendance. Thus, expanding the league can be another solution.

Unlike the Blue Ribbon Panel, current research uses game attendance as an indicator of Competitive Balance instead of payroll ratio. And the suggestions provided by this research are not focused on adjusting the financial situation of the franchise in MLB, but on how to attract more fans to the game. By looking at the parameter estimates giving by chapter 6, this research is trying to display which kind of game are attractive to fans. Practitioners therefore can allocate their resources to the dimensions that are closely correlated with game attendances.

Instead of simply agreeing or disagreeing that fans prefer uncertainty of outcome, this research reveals that fans prefer changes in games, whether expected or unexpected. For instance, research results say that focusing too much on making the underdogs win all their games may not be an effective way of getting better game attendance. On the other hand, decreasing consecutive wins and increasing game importance may be more relevant to game attendances.

Scheduling of the game affects game attendances. Arranging more inter league games or adjusting matched pairs' division combinations can help Major League Baseball gain more attendances. Referring to the results in chapter 6, time and location can also be used to adjust the attendance. The National Football League uses a pre-determined formula to arrange its regular season schedule. According to this formula, teams play with the opponents that finished in the same place in their own divisions as themselves. For example, Browns in AFC North division finished in the 4^{th} place, then Brown will play with Bills, Jaguars and Chiefs, who also finished in the 4^{th} place in their divisions (Table 7.1 Source: Wiki, 2010). If this research results also applies to attendances in football league, than the games played by opponents that finished in the same place in different division should attract more fans than the games play by two random opponents when controlling for all other factors.

Table 7.1

An Example for the National Football League Game Schedule

	AFC East	AFC North	AFC South	AFC West
1st Place	Dolphins	Steelers	Titans	Chargers
2nd Place	Patriots	Ravens	Colts	Broncos
3rd Place	Jets	Bengals	Texans	Raiders
4th Place	Bills	Browns	Jaguars	Chiefs
	NFC East	NFC North	NFC South	NFC West
1st Place	Giants	Vikings	Panthers	Cardinals
2nd Place	Eagles	Bears	Falcons	49ers
3rd Place	Cowboys	Packers	Buccaneers	Seahawks
4th Place	Redskins	Lions	Saints	Rams

* Schedule for Browns. Yellow teams play with Brown once, and blue teams play with Browns twice.
Source: http://en.wikipedia.org/wiki/Regular_season_(NFL)

However, findings such as game importance and team performance changes are more helpful in terms of understanding why game attendance is higher or lower, but it does not provide direct solutions about how to make attendance higher. For example, the game schedule is arranged before the season starts, and the game importance is a dynamic value that changes game by game. In addition, the data shows that league expanding is associated with game attendance increases, but this study does not answer questions such as will league expanding always lead to game attendance increase.

7.4 Discussion and Future Research

First I would like to discuss the generalization of this research. In this dissertation, I proposed a set of Competitive Balance measures. Using a matched pair as a basic unit of research, I analyzed Competitive Balance, and checked the Uncertainty of Outcome Hypothesis. The matched pair approach can be generalized to any sports involving more than one team in a game. This approach relieves the violation of independent sample assumption. Moreover, the multidimensionality of Competitive Balance exists in many kinds of sports, thus the measures developed in this study as well as the methods used in analysis can be applied to other sports genre.

Although I have a strong feeling that fans' responses to measures such as game importance and consistency are similar in other professional sports, I would investigate the profiles of a specific sport genre before I generalize the detailed findings about Competitive Balance, because the data profile for different sports genre can vary. In the future, I would like to apply my measures to other sports genres in order to make a generalization with more confidence.

Another issue I would like discuss is about measure construction under two circumstances. When I try to capture fans' response, I would use information that fans already know or fans can predict before they go to a game; when I try to analyze the Competitive Balance in a game, I do not have to consider fans' information awareness. For example, if I want to know the closeness of a matched pair in a game, I use current game scores to form the closeness measure. However, if I want to know how fans respond to score differences, I should use scores in previous game(s), because fans do not know what a game's scores are until it finishes. Again, the development of previous score difference as a Competitive Balance measure is on my future research agenda.

Next, I would like to discuss the Dominance in this study. In this research, I use upset game ratio (DOMA) in a matched pair to measure how a stronger team dominates a weaker team. Here strong team refers to the team with higher previous winning percentage in its league. Notice in this research the concept of Dominance is not necessarily equivalent to Vrooman's

Dominance--the "dominance of large-market clubs" in a league, although the large-market clubs are more likely to have higher winning percentages. Consecutive wins (CSCB) is another measure that counts number of wins of one team in a matched pair. However, CSCB does not focus on large market team's dominance but rather examines how often a result repeats itself. An alternative choice is to name DOMA as "upset game ratio" or frame DOMA as a consistency measure. No matter which dimension I frame DOMA and CSCB, the results regarding these measures will not change.

The renaming thought above leads to my next discussion regarding Vrooman's three dimensions concept framework. Current research takes the cue from Vrooman's theoretical frame work about the three dimensions of Competitive Balance. However, it is possible that categorizing Competitive Balance dimensions as Closeness, Dominance, and Consistency are not the only solutions. This research shows that Competitive Balance is multidimensional, and Closeness, Dominance, and Consistency are important factors. However, the proposed Competitive Balance measures have more than three dimensions. Needless to say, the measures in the proposed Competitive Balance measures are subject to revisions and expansions.

With no doubt, the set of measures proposed in this research only capture a fraction of the dimensions of Competitive Balance and the set is ready to be expanded and revised. One future direction is to develop measures using detailed performance data such as the number of home runs, home line scores, visiting line scores, distribution of number of games in a given season etc. In addition, when data is available, one can integrate more information into a study such as ticket prices, players wage levels over years, the size of the stadium, team history, the media coverage etc. These variables are not included in current research due to the availability of the historic data.

Last, I would like to talk about sample size and result significance. Models in Chapter 6 involve hundreds of parameter estimates, such as average attendance and attendance rate for each matched pair, as well as parameters for the variance-covariance structures and coefficients for other covariates. Without the large sample size, the model in Chapter 6 will not work, let along deriving any significant results. The significant levels in the results tell us how likely a researcher is to get a result by chance. Therefore, the more data we have, the more confidence a researcher can conclude that the results are true and not just random chance. And, a researcher with a lot of observations can derive more accurate parameter estimates than with a few observations.

REFERENCES

Amabile, T. (1996). *Creativity in Context*: *Update to the Social Psychology of Creativity.* Boulder, CO: Westview.

Berri, D., Schmidt, M., & Brook (2007). *The Wages of Wins* (1st ed.). Stanford, CA: Stanford University Press.

Bodvarsson, O., & Brastow, T. (1998). Do employers pay for consistent performance? Evidence from the NBA. *Economic Inquiry, 36(1)*, 145-160.

Booth, R. (2005). Revenue sharing. In Levinson, D., & Christensen, K. (Eds.) *Berkshire Encyclopedia of World Sport* (pp.1262-1264). Great Barrington, MA: Berkshire Publishing Group.

Borland, J., & MacDonald, R. (2003). Demand for sport. *Oxford Review of Economic Policy, 19*(4), 478-502.

Branscombe, R., & Wann, L. (1991). The positive and self concept consequences of sports team identification. *Journal of Sport & Social Issues, 15*, 115-127.

Branscombe, R., & Wann, L. (1994). Sport psychology. In Magill F. (Ed.) *Magill's Survey of Social Sciences: Psychology* (pp.2363-2368). Pasadena, CA: Salem Press.

Bukiet, B., Harold, R., & Palacios, L. (1997). A markov chain approach to baseball. *Operations Research, 45*(1), 14-23.

Campbell, T. (1965). Variation and selective retention in socio-cultural evolution. In Barringer, H., Blankstein, G., & Mack, R. (Eds.) *Social Change in Developing Areas: A Reinterpretation of Evolutionary Theory* (pp.19-49). Cambridge, Mass: Schenkman.

Canes, E. (1974). The social benefits of restrictions on team quality. In Noll, G. (Ed.) *Government and the Sports Business* (pp. 81–114).Washington, DC: Brookings Institution.

Clarke, R., & Norman, J. (1995). Home ground advantage of individual clubs in english soccer. *The Statistician, 44*(4), 509-521.

Clarke, S. (2005). Home advantage in the australian football league. *Journal of Sports Sciences, 23*(4), 375-385.

Craig D. (1999). Free-agency and the competitiveness of Major League Baseball. *Review of Industrial Organization, 14*(3), 205-217.

David, A., & Smith, M. (1994). The home-court advantage: How large is it, and does it vary from team to team? *The American Statistician*, 48(1), 22-28.

Davis, M. (2009) Analyzing the relationship between team success and MLB attendance with GARCH effects. *Journal of Sports Economics, 10*(1), 44-58.

Depken C. (2000). Wage disparity and team productivity: Evidence from Major League Baseball. *Economics Letters, 67(1)*, 87-92.

Dobson, S., & Goddard, A. (2001). *The Economics of Football*. Cambridge, England: Cambridge University Press.

Donihue, M., Findlay, D., & Newberry, P. (2007). An analysis of attendance at Major League Baseball spring training games. *Journal of Sports Economics, 8*, 39-61.

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

Eckard, W. (2001b). Free agency, Competitive Balance, and diminishing returns to pennant contention. *Economic Inquiry, 39*(3), 430-444.

El-Hodiri, M., & Quirk, J. (1971). An economic model of a professional sports league. *Journal of Political Economy, 79*(6), 1302-1319.

Fishman, P. (2003). Competitive Balance and free agency in Major League Baseball, *American Economist*, 47(1), 1-8.

Forrest, D., Beaumont, J., Goddard, J. A., & Simmons, R. (2005a). Home advantage and the debate about Competitive Balance in professional sports leagues. *Journal of Sports Sciences, 23*(4), 439-445.

Forrest, D., Beaumont, J., Goddard, J., & Simmons, R. (2005b). Home advantage and the debate about Competitive Balance in professional sports leagues. *Journal of Sports Sciences, 23*(4), 439-445.

Fort, R. (2003). Thinking (some more) about Competitive Balance. *Journal of Sports Economics, 4*(4), 280-283.

Fort, R., & Maxcy, J. (2003). "Competitive Balance in sports leagues: An introduction". *Journal of Sports Economics, 4*(2), 154-160.

Fort, R., & Quirk, J. (2004). Owner objectives and Competitive Balance. *Journal of Sports Economics, 5*(1), 20-32.

Frick, B, Prinz, J., & Winkelmann, K. (2003). Pay inequalities and team performance: Evidence from the North American major leagues. *International Journal of Manpower, 24*, 472-488.

Gerrard, B. (2006). Competitive Balance and the sports media rights market: What are the real issues? In Jeanrenaud, C. & Kesenne, S. (Eds.) *The Economics of Sport and the Media* (pp.26-36). Northampton, MA: Edward Elgar.

Humphreys, B. (2002). Alternative measures of Competitive Balance in sports leagues. *Journal of Sports Economics, 3*(2), 133-148.

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

Humphreys, B. (2003b). Comments on "Thinking about Competitive Balance". *Journal of Sports Economics, 4*(4), 284-287.

Humphreys, B. (2005). Definition (Competitive Balance). In Levinson, D., & Christensen K. (Eds.) *The Berkshire Encyclopedia of World Sport* (pp. 347), Great Barrington, MA: Berkshire Publishing Group.

Humphreys, B., & Watanabe, N. (Forthcoming). Competitive Balance. In Kahane, L., & Shmanske, S. (Eds.) *The Oxford Handbook of Sports Economics* (Vol I & II). New York, NY: Oxford University Press.

Kahane, L. H. (2003). Comments on "Thinking about Competitive Balance". *Journal of Sports Economics, 4*(4), 288-291.

Kiecolt, K., & Nathan, L. (1985). In Kiecolt K., & Nathan L. (Eds.) *Secondary Analysis of Survey Data* (1st Ed.). Beverly Hills, CA: Sage.

Knobloch-Westerwick, S., David, P., Eastin, S., Tamborini, R., & Greenwood, D. (2009). Sports spectators' suspense: Affect and uncertainty in sports entertainment. *Journal of Communication, 59*(4), 750-767.

Koning, R. (2000). Balance in competition in dutch soccer. *The Statistician, 49*(3), 419-431.

Koning, R., & Witteloostuijn, A. (2008) Competitive Balance in national European soccer competitions. In Albert, J., & Koning R. (Eds.) *Statistical Thinking in Sports* (pp.63-75). Boca Raton, London: Taylor & Francis Group.

Lai, L., Ng, F., & Yung, P. (2008). The coase theorem and a coasian construction economics and management research agenda. *Construction Management & Economics, 26*(1), 29-46.

Lapointe, J. (2004). NHL is displaying its Competitive Balance. *New York Times, 153*(52840), D5-D5.

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

Leeds, M., & von Allmen, P. (2002). *The Economics of Sports (1st Ed.)*, Boston, MA: Addison-Wesley.

Leeds, M. (2008). Book review: The bottom line: Observations and arguments on the sports business, by Andrew Zimbalist. Philadelphia, PA: Temple University press, 2006. *Journal of Sports Economics,* 9(1), 106-109.

Levin, C., Mitchell, J., Volcker, A., & Will, F. (2000). *The Report of the Independent Members of the Commissioner's Blue Ribbon Panel on Baseball Economics.* New York, NY: Major League Baseball.

Levitt , S., & Dubner, S. (2005). *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything*. New York, NY: William Morrow.

Litman, A., & Jimerson, L. (2004). The measure of curiosity as a feeling-of-deprivation. *Journal of Personality Assessment,* (82), 147-157.

Litman, J. A. (2005). Curiosity and the pleasures of learning: Wanting and liking new information. *Cognition & Emotion, 19*(6), 793-814.

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological,* (116), 75-98.

MacDonald, N., & Reynolds, O. (1994). Are baseball players paid their marginal products? *Managerial and Decision Economics, 15*, 443-457.

Maxcy, J. (2002) Rethinking restrictions on player mobility in Major League Baseball. *Contemporary Economic Policy, 20*(2), 145.

Maxcy, J., & Mondello, M. (2006). The impact of free agency on Competitive Balance in North American professional team sports leagues. *Journal of Sport Management, 20*(3), 344-364

McHale, I., & Davies, S. (2008). Statistical analysis of the effectiveness of the FIFA world rankings. In Albert, J., & Koning R. (Eds.) *Statistical Thinking in Sports* (pp.77-109). Boca Raton, London, New York: Taylor & Francis Group.

Medema, S. G., & Zerbe, R. O. (2000). Educating Alice: Lessons from the coase theorem. *Research in Law and Economics,19* (1), 69-112.

Meehan, Jr., Nelson, R., & Richardson, V. (2007). Competitive Balance and game attendance in Major League Baseball. *Journal of Sports Economics, 8*(6), 563-580.

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. *Economics Bulletin, 26*(1), 7.

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

Nevill, A., & Holder, R. (1999). Home advantage in sport: An overview of studies on the advantage of playing at home. *Sports Med., 26*(4), 221-236.

Palomino, F., & Rigotti, L. (2000, 01/01). The sport league's dilemma: Competitive Balance versus incentives to win. Message posted to http://129.3.20.41/eps/io/papers/0012/0012003.pdf

Quirk, J., & Fort, D. (1992). In *Pay dirt: The Business of Professional Team Sports*. Quirk, J., & Fort, D. (Eds.), Princeton, NJ: Princeton University Press.

Rottenberg, S. (1956). The Baseball players' labor market. *Journal of Political Economy*, 64, 242–258.

Sales, E. & Lichtenwalter, S. (2006). Secondary analysis in social work research education: Past, present, and future promise. *Journal of Social Work Education, 42*(3), 543.

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

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

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

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

Schmidt, B. (2006). On the evolution of competition: An application of nonlinear tests. *Applied Economics,38*(1), 1-12.

Schmidt, B. & Berri, J. (2002). The impact of the 1981 and 1994–1995 strikes on Major League Baseball attendance: A time-series analysis. *Applied Economics, 34*(4), 471-478.

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

Sherony, K., Haupert, M., & Knowles, G. (2000). Competitive Balance in Major League Baseball. *A Journal of Baseball History and Culture, 9*(1), 225-236

Singer, J., & Willett, J. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press.

Sloan, R. (1989). The motives of sports fans. In Goldstein J.(Ed.) *Sports, Games, and Play: Social and Psychological Viewpoints* (2nd Ed., pp.175-240). Hillsdale, NJ: Erlbaum.

Sloane, P. J. (1971). The economics of professional football: The football club as a utility maximiser. *Scottish Journal of Political Economy,17,*121–145.

Stefani, R. (2008). Measurement and interpretation of home advantage. In Albert, J. & Koning, R. (Eds.), *Statistical Thinking in Sports* (pp.203-216). Boca Raton, London: Taylor & Francis Group.

Szymanski, S. (2003a). Incentives and Competitive Balance in team sports. *European Sport Management Quarterly, 3*(1), 11-30.

Szymanski, S. (2003b). The assessment: The economics of sports. *Oxford Review of Economic Policy, 19*(4),467-477.

Triplett, N. (1898). The dynamogenic factors in pacemaking and competition. *American Journal of Psychology, 9*, 507-533.

UCLA. Introduction to SAS. UCLA: Academic technology services, statistical consulting group. from http://www.ats.ucla.edu.proxy2.library.uiuc.edu/stat/sas/notes2/ (Accessed Nov.24, 2007).

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

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

Wikipedia contributors. (May 2008). *Validity (statistics)*, July 27, 2008, from http://en.wikipedia.org/wiki/Construct_validity

Wikipedia contributors. (July 2008). *Major League Baseball Season*, July 28, 2008, from http://en.wikipedia.org/wiki/Major_League_Baseball_season#Division_games

Winfield, D., & Levin. (2007). In Winfield D. (Ed.), *Dropping the Ball: Baseball's Troubles and How We Can and Must Solve Them*. Scribner, New York: Charles Scribner's Sons.

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

Zimbalist, A. (2003). *May the Best Team Win: Baseball Economics and Public Policy*. Washington, D.C.: Brookings Institution Press.

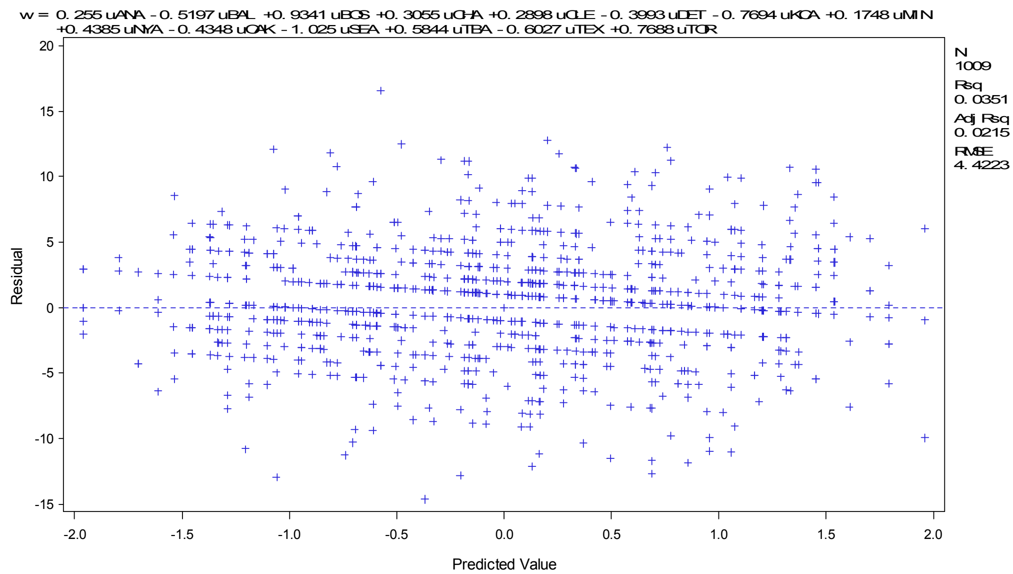
Zuckerman, M. (1979). *Sensation Seeking: Beyond the Optimal Level of Arousal*. Hillsdale, NJ: Lawrence Erlbaum Assocs., Inc.

APPENDIX

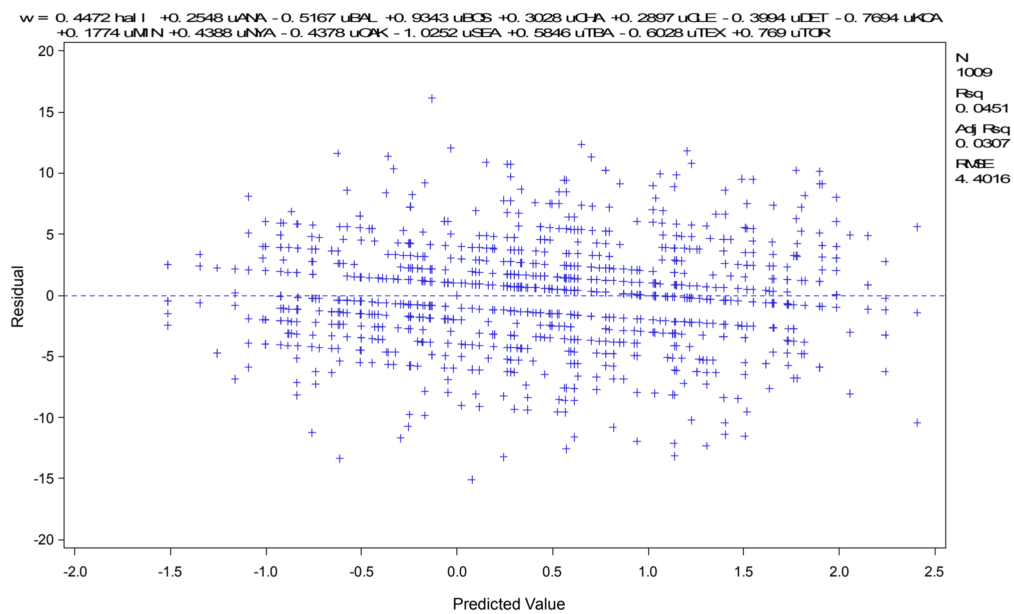
I. Appendix for Chapter 4

Data set used for the following plots: 2008 AL

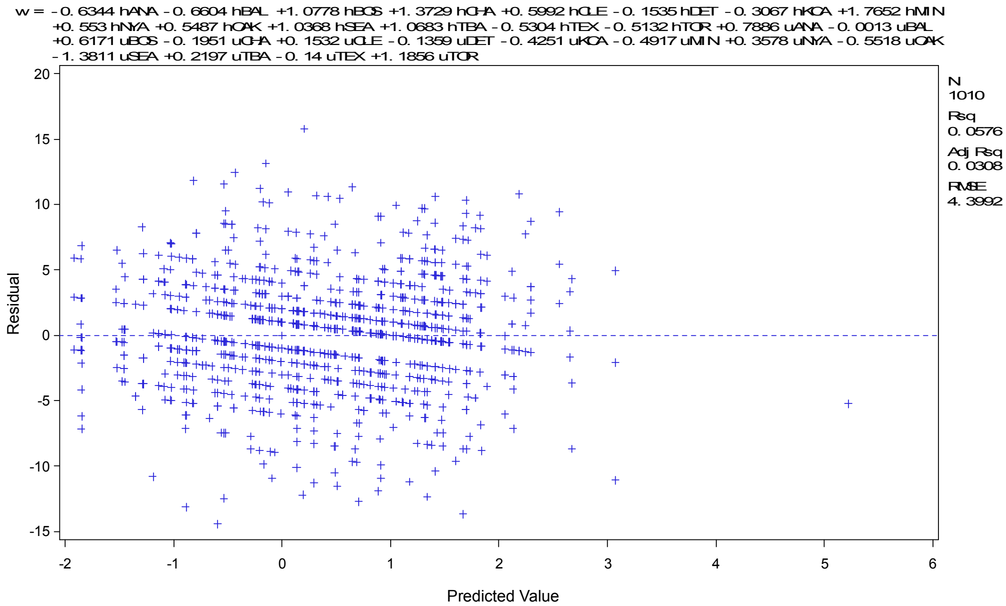
Residual plots for M1 (model 1)



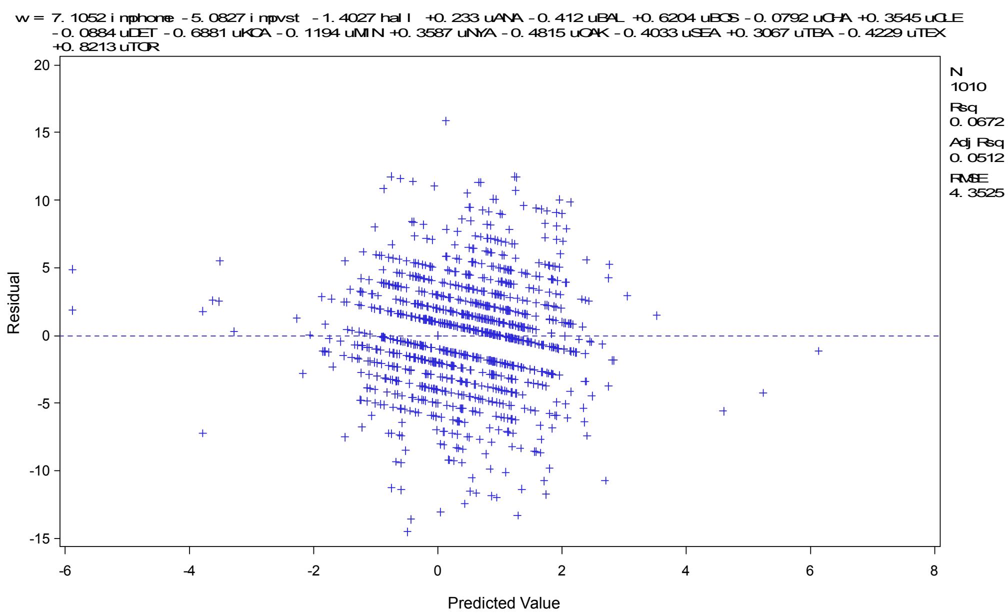
Residual plots for M2a



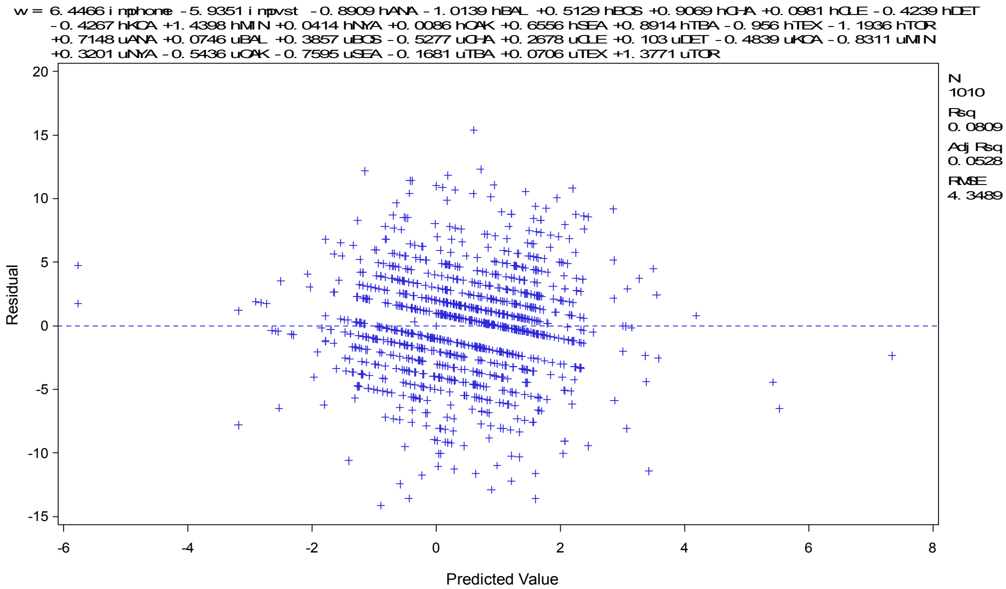
Residual plots for M2b



Residual plots for M3a



Residual plot of M3



II. Appendix for Chapter 5

year	Average matched pair attendance of the year	paircount	RMSE	avgGower
1941	8824.02	56	0.073442	0.74294
1942	9932.71	56	0.079989	0.74891
1943	8928.16	56	0.070434	0.74697
1944	9884.96	56	0.074461	0.75501
1945	11239.27	56	0.073174	0.7413
1946	17892.43	56	0.069824	0.73235
1947	18375.73	56	0.070707	0.74166
1948	19779.04	56	0.075368	0.7547
1949	17870.55	56	0.079662	0.73334
1950	16509.3	50	0.08618	0.73568
1951	13735.41	56	0.080762	0.7291
1952	12955.18	56	0.070053	0.7494
1953	13169.87	56	0.072151	0.73749
1954	13521.9	56	0.076278	0.73045
1955	14304.64	56	0.075156	0.72684
1956	14313.76	56	0.07235	0.73276
1957	14184.01	56	0.07548	0.74477
1958	13706.12	56	0.070941	0.76155
1959	15757.64	56	0.077708	0.73878
1960	16120.24	56	0.070976	0.73542
1961	14602	56	0.06968	0.74008
1962	13544.6	73	0.07779	0.75679
1963	12841.88	90	0.07391	0.76122
1964	13043.57	90	0.067574	0.76462
1965	13884.93	90	0.070436	0.76253
1966	15634.91	90	0.069687	0.74291
1967	14798.24	90	0.070736	0.75699
1968	14106.46	90	0.077903	0.75957
1969	15631.89	90	0.066879	0.76243
1970	14938.1	132	0.068755	0.77722
1971	15127.73	132	0.062466	0.78375
1972	14987.08	132	0.058423	0.7754
1973	15650.36	132	0.068387	0.77498
1974	15621.88	132	0.070607	0.77063
1975	15588.26	131	0.062748	0.77816
1976	16246.64	132	0.06475	0.7774
1977	19100.71	132	0.065518	0.78323
1978	19344.46	157	0.066036	0.77692
1979	20740.43	157	0.066819	0.77483
1980	20384.68	157	0.066413	0.7866
1981	18947.51	157	0.065961	0.78894
1982	21103.88	157	0.072375	0.76095
1983	21420.07	157	0.069046	0.78023
1984	21301.35	157	0.071337	0.77174
1985	22248.3	157	0.072369	0.7795

1986	22614.9	157	0.064196	0.78082
1987	24623.74	157	0.065432	0.78278
1988	25267.62	157	0.067422	0.78393
1989	26325.26	157	0.065558	0.79073
1990	26179.1	157	0.065082	0.78969
1991	27202.13	157	0.066227	0.79629
1992	26676.16	157	0.064252	0.79567
1993	29553.93	157	0.067153	0.78201
1994	31249.52	182	0.080132	0.79363
1995	25237.94	182	0.075044	0.78359
1996	26527.03	182	0.073356	0.8009
1997	28846.64	248	0.074129	0.80318
1998	28907.46	249	0.072573	0.79984
1999	29352.33	283	0.066727	0.82216
2000	30284.78	283	0.066953	0.80051
2001	30310.02	284	0.066779	0.79653
2002	28652.38	284	0.065905	0.80795
2003	28032.5	285	0.073392	0.78821
2004	30331.34	286	0.064191	0.81113
2005	31303.89	284	0.069488	0.79712
2006	31492.88	283	0.065783	0.79876
2007	32911.62	281	0.071305	0.80359
2008	32736.09	284	0.066627	0.80522