EMPIRICAL ESSAYS IN SPORT MANAGEMENT

by

Steven H. Salaga

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Kinesiology)
in The University of Michigan
2012

Doctoral Committee:

Professor Rodney D. Fort, Chair
Professor Charles C. Brown
Associate Professor Jason A. Winfree
Assistant Professor Dae Hee Kwak

ACKNOWLEDGEMENTS

I would like to thank my entire doctoral dissertation committee for offering to provide to their time and expertise in assisting me through the dissertation process. I truly appreciate all the support you have given over the course of my years at Michigan.

TABLE OF CONTENTS

LIST OF TABLES

Table

List of Figures

Figure

ABSTRACT

Chair: Rodney D. Fort

This dissertation contains three separate empirical essays in sports economics with each chapter focusing on a key area of investigation in the discipline – consumer demand, the labor market for players and competitive balance. The first chapter analyzes secondary market demand for National Football League (NFL) attendance through the utilization of personal seat license and season ticket rights sales data. The use of this unique data avoids the venue capacity constraint which has dominated previous NFL attendance estimations. The analysis uncovers strong consumer preferences for high-quality seating locations, clear differences in demand between NFL markets with respect to short-term team quality and evidence that personal seat licenses are depreciable assets.

The second chapter evaluates the relationship between training and employment outcomes in the context of the North American professional baseball labor market. Using historical Major League Baseball (MLB) Draft data, the study examines labor market outcomes as measured by probability of reaching MLB and MLB career duration. Logistic regression models show that players drafted out of four-year institutions have significantly higher probabilities of reaching MLB while hazard modeling illustrates that players drafted directly from high school have significantly longer careers once they

reach MLB. Findings from this specialized labor market support in part the positive theorized relationship between accumulated training and employment outcomes.

The final chapter examines the historical behavior of competitive balance in college football and also evaluates Rottenberg's invariance proposition (IP) in response to three key institutional changes in the sport's business structure. Results illustrate increasing levels of game uncertainty, but otherwise relatively little change in balance over time. Based on the sport's supreme popularity, these results raise the question of how important competitive balance truly is to the long-term financial viability of NCAA conferences. Additionally, the use of time series techniques uncovers mixed support for the IP with only slight evidence suggesting any structural changes in balance in response to the events identified. Time series techniques reinforce the finding that balance has been relatively stable over the history of the sport.

Chapter 1

Introduction

This dissertation contains three separate empirical essays in the area of sports economics. Each is fully contained within its given chapter. Each chapter focuses on a key area of investigation in the field of sports economics – consumer demand, the labor market for players and competitive balance. All three chapters are briefly described below.

The first chapter analyzes secondary market demand for National Football League (NFL) attendance through the utilization of personal seat license and season ticket rights sales data. While estimation of demand for NFL football has been largely ignored in the professional sport demand landscape, the availability of this data has created an opportunity to revisit NFL attendance demand. The utilization of this data avoids the venue capacity constraint which has dominated previous NFL attendance estimations. The analysis uncovers strong consumer preferences for high quality seating locations, clear differences in demand between NFL markets with respect to short-term team quality and evidence that personal seat licenses are depreciable assets.

The second chapter evaluates whether or not the economic theory regarding the positive relationship between accumulated training and employment outcomes holds in the context of the North American professional baseball labor market. Through the use of historical Major League Baseball (MLB) Draft data, the study examines selection into the labor market and labor market outcomes as measured by probability of reaching MLB and MLB career duration. Logistic regression models show that players drafted out of four-year institutions have significantly higher probabilities of reaching MLB while hazard modeling illustrates that players drafted directly from high school have significantly longer careers once they reach MLB. Findings from this specialized labor market show that the theorized relationship holds in only one of the two labor market outcomes measured.

The final chapter contains two sections. The first section examines the historical behavior of competitive balance in Bowl Championship Series (BCS) college football conferences through the presentation of four measures of balance. The findings illustrate an increase in game closeness over time and little change in both the correlation and distribution of team winning percentages and conference championships. This suggests relatively little change in balance over the history of the sport. Furthermore, individual BCS conferences have historically been dominated by a single team or a very small group of elite programs. Based on the supreme popularity of the sport, these results raise the question of how important competitive balance really is to the long-term financial viability of NCAA conferences.

The second section of the final chapter conducts a long-run analysis of the behavior of competitive balance in response to three key events which altered the business structure of college football. This is achieved through empirical tests of Rottenberg's invariance proposition (IP). The analysis uncovers mixed support for the IP by providing only slight evidence that balance was altered following the key events identified. Specifically, time series techniques show that nine of fifteen competitive balance metric series are stationary without break points. This suggests competitive balance has been relatively stable over time and that the three events outlined have not significantly influenced balance in the majority of cases. On the other hand, three of the nine total break points identified match approximately with one of the three key events presented. In each case the identified break point is followed by a subsequent enhancement of balance. This provides evidence supporting the hypothesis that the institutional changes identified are associated with structural changes in competitive balance.

CHAPTER 2

Personal Seat Licenses, Season Ticket Rights and National Football League Demand

2.1 Introduction

Demand analysis is a cornerstone of the sports economics literature. Numerous studies have examined the primary demand drivers in professional team sports, but the majority of this work has focused on professional baseball and European football (Borland & Macdonald, 2003). A relative lack of work has been completed on North American professional football (Fizel, 2006), especially in recent years when a large percentage of games are sold-out. This study analyzes secondary market personal seat license (PSL) and season ticket right (STR) sales data to estimate demand for the National Football League (NFL).

PSLs and STRs grant their owners the rights to purchase NFL tickets at face value for a particular seat in the stadium. Typically these rights are for the duration of the stadium. Since the face value of a ticket is typically below the actual market price, the face value is not sufficient to accurately gauge demand. However, sale prices for PSLs and STRs on the secondary market allow for the ability to capture the true long run demand for NFL attendance.

Beyond traditional single-game ticket sales, which have previously dominated the secondary market, PSLs and STRs have made a substantial entrance into this marketplace the explosion of this secondary market (Drayer & Martin, 2010) has created an opportunity to revisit the estimation of demand for NFL football. Despite the relative importance of the NFL, few studies have examined the demand for league attendance. This is likely because a large percentage of NFL games become sold-out. Due to the prevalence of these sell-outs, attendance at the vast majority of games is truncated due to venue capacity constraints. Subsequently, a statistical truncation issue is present when attempting to estimate demand for professional football.

Based on the current ticket prices charged by NFL franchises, in the vast majority of markets there is a higher quantity demanded for attendance than what NFL teams have chosen to accommodate based on their seating capacities. Numerous studies have verified that professional sports franchises price in the inelastic portion of demand, and this literature is summarized and expanded upon by Fort (2004). Due in part to inelastic pricing, the "venue capacity" variable has dominated the majority of previous demand estimations. Therefore, estimating NFL demand using secondary market PSL and STR sales offers a way around the truncation quandary. Since the secondary market is now so extensive, enough data exists to properly estimate demand for the NFL without having to wrestle with the truncation issues which have previously dominated the statistical estimation of game-day attendance.

This chapter contributes to the literature on professional sports demand in a number of ways. First, a brief theoretical model of the secondary PSL and STR market is provided. Also, the choice of dependent variable eliminates the venue capacity constraint which has traditionally been encountered when using game-day attendance as a proxy for NFL demand. The ability to avoid this statistical issue revisits a line of investigation which has been relatively dormant since Noll's (1974) seminal contribution. The study also utilizes a unique data set and investigates a market which has been seen little empirical investigation in the academic literature. The result is an unobstructed view of the factors which drive attendance demand in the NFL. Specifically, the estimations show that consumer willingness to pay for the live NFL product differs significantly based on seat quality – explicitly in regards to closeness to midfield and the field of play. More importantly, the findings illuminate clear differences in demand between PSL and STR markets. Namely, significant differences exist between PSL and STR markets in the relationship explaining team quality and demand. Consumers in PSL markets exhibit strong sensitivity to changes in team quality while fans in STR markets do not. This is an important finding considering that the positive relationship between team quality and consumer demand is one of the most consistent findings in the sports economics literature. Lastly, the empirical models provide evidence that the value of PSLs decline in value over time – suggesting they are depreciable assets. This result is supported by several negative coefficients on the year of sale indicators in multiple models.

The chapter proceeds with a background section on PSL sales in the NFL, a literature review, a theoretical treatment of PSL and STR sales in the secondary market, outline of the data and statistical models, estimation results and conclusions.

2.2 Background

There are a very limited number of contributions investigating the role of PSLs in the revenue generation landscape of professional sports franchises (for example, Noll &

Zimbalist, 1997; Fort, 2003; Howard & Crompton, 2004; Reese, Nagel & Southall, 2004). Furthermore, there is no empirical work examining PSLs or STRs in either the primary or secondary market. Based on the overall lack of attention paid to the topic, a brief overview appears to be warranted.

PSL sales in the NFL are regulated by the league's Collective Bargaining Agreement (CBA). Specifically, franchises are only able to sell PSLs in conjunction with the construction of a new venue or the renovation of an existing venue. One significant advantage of implementing a PSL program is that revenue collected from PSL sales is eligible for exemption from the NFL revenue sharing formula. If franchises successfully apply for a league "waiver" they are permitted to retain revenues generated from PSL sales. Under the alternative, which is simply increasing the face value price of season tickets, franchises would be forced to share a portion of these increased revenues under "gross receipts" as is specified in the league's revenue sharing formula (NFLPA, 2006). Consequently, for the franchise that is unable to secure public funding, the ability to implement a PSL sales program allows the team to easily fund the private construction of a venue.

Since PSL sales programs are regulated by the NFL CBA, franchises are able to generate additional capital through a PSL sales program and are then able to parlay those PSL revenues into an upgraded venue with enhanced revenue production capabilities. The ability to construct or renovate venues stocked with luxury boxes and premium seating is a vital mechanism needed to generate revenues from the demand existing in each team's local market. Since 1995, 24 of the 32 NFL franchises have either renovated or constructed new venues. Fifteen of those clubs have used the sale of PSLs in order to generate capital to fund these ventures. Those which did not incorporate a PSL sales program were either the recipients of publically funded stadiums or used the sale of stadium naming rights to cover the franchise's private contribution towards venue construction. Table 2.1 shows the growth of PSLs in the NFL.

Over the past two decades, PSL sales have become a significant source of revenue for NFL franchises. Though access to team financial statements are unavailable, revenues collected from the implementation of PSL programs are undoubtedly substantial. In 1993, the Carolina Panthers were the first NFL franchise to utilize a modern PSL sales program, which raised an estimated $100 million in after tax revenue (Ostfield, 1995). More recently, Dallas Cowboys owner Jerry Jones charged a PSL fee ranging from $16,000 to $150,000 for each of his 15,000 club seats in the new Cowboys Stadium (Sandomir, 2009). Current PSL revenues are so substantial that fees collected by the Cowboys in association with the construction of their new venue are estimated at $720.4 million (Vrooman, 2010). Clearly, PSL fees represent a sizeable portion of "gate receipt" revenues collected by many NFL franchises. Though the case of the Cowboys may not represent the league norm, it does provide a frame of reference for the potential revenue a PSL program may generate.

Traditionally, single-game ticket sales have dominated the secondary market, but over the past two decades, PSL and STR sales have become more common. In the primary market, PSLs and STRs are clearly two distinct products. PSLs require an upfront fee which grants the consumer the ability to purchase season tickets from the franchise each year[1]. STRs represent the right to purchase season tickets directly from

[1] Franchises utilizing PSL programs typically set the vast majority of their venue capacity as "PSL seats" which require an upfront PSL fee and are then sold as season tickets. Traditionally, a very limited

franchises that do not implement formal PSL programs and require no upfront fee. For both products, a failure to pay the franchise the face value price for season tickets results in the loss of the PSL or STR, while continuation of payment on a yearly basis maintains the consumer's status as a PSL or STR holder. In the case of a PSL, the consumer is purchasing an asset which grants the ability to purchase season tickets for an extended period of time. As an STR holder, the same right is conveyed to the consumer without the upfront fee.

While there are differences between these two products in the primary marketplace, the distinction between these commodities is eliminated in the secondary market. In the case of both products, the asset being transferred from consumer to consumer is simply the ability to purchase season tickets directly from the franchise. The rights to NFL season tickets are a valuable piece of property (Reese, et al., 2004) and a consumer purchasing either a PSL or STR in the secondary market has equivalent options[2]. First, the consumer can simply choose to purchase season tickets from the franchise each year and maintain their PSL or STR rights. An alternative option is to resell the PSL or STR asset on the secondary market, where a profit or loss may be realized based on market conditions. A final and highly unlikely option is for the consumer to decline their option of purchasing season tickets from the franchise, in which case the secondary market resale value of the PSL or STR would be forfeited along with

number of low quality seats are sold as "non-PSL seats" available for purchase without a PSL either prior to the season or on game day.

[2] In order for STRs to be sold in the secondary market, the issuing NFL franchise must allow their current season ticket holders the opportunity to sell or "transfer" their season ticket rights. NFL franchises place varying restrictions on the resale of both STRs and PSLs (Reese, et al., 2004) and these team policies have been altered over the time period examined.

PSL and STR ownership rights. Therefore, in the secondary marketplace, a PSL or STR sales transaction represents a purchase of the rights to NFL attendance.

It is important to note, however, that there are differences between pricing mechanisms in the primary and secondary market. In the case of a primary market PSL sale, the consumer has the opportunity to purchase a PSL from the franchise that is typically at a pre-determined fixed price based on seat quality and other components associated with individual franchise demand. In the secondary market, pricing represents a dynamic process in which transaction prices fluctuate based on numerous determinants. Both buyers and sellers have the ability to view recent sales transactions resulting in secondary market prices changing more fluidly based on existing market conditions. This distinction is one component which makes this data unique. Previous literature has estimated NFL attendance, which represents an estimation of a fixed quantity. Instead, we are able to estimate price which has the ability to fluctuate based on market conditions.

While previous empirical work has aided in our understanding of demand for professional football attendance, examining secondary market PSL and STR sales may allow for the opportunity to evaluate a more representative purchase decision by the consumer. NFL franchises typically sell the vast majority of their seating capacity in the form of advance season tickets, which largely eliminates the game-day “walk up” consumer. Therefore, examining the purchase decision in advance, which is the case when a consumer purchases a secondary market PSL or STR, may more accurately reflect the true purchase decision regarding NFL attendance. When a consumer acquires a PSL or STR, they are committing to the purchase of not only a single game, but ten total (two preseason and eight regular season) home games per season as well as the rights to purchase season tickets into the future and the option to sell the asset.

Furthermore, there has been a lack of empirical work inspecting the sale of PSLs and STRs[3]. At this point in time, there does not appear to be any published research analyzing either primary or secondary market data on either item. Accordingly, this chapter uses secondary market PSL and STR sales data to estimate demand for the rights to NFL attendance. Over 3,800 secondary market sales transactions occurring from 2005 to 2009 are used in the analysis[4].

2.3 Literature on NFL Demand

Empirical research on demand for professional team sports has received a significant amount of attention in the sports economics literature. In empirical demand studies, game-day attendance has traditionally been used as a proxy for demand (Krautmann & Hadley, 2006). The earliest research on attendance demand in professional team sports was completed by Demmert (1973) and Noll (1974). Demmert's work examined professional baseball attendance while Noll highlighted the similarities and differences between the four major North American team sports. Both of these early empirical pieces paved the way for future research in the area.

On the whole, the attendance demand literature has largely focused on Major League Baseball and European football. There are also studies that estimate attendance in the National Basketball Association and the National Hockey League. However, despite

---

[3] Previous work including Noll & Zimbalist (1997), Fort (2003) and Howard & Crompton (2004) has identified the sale of PSLs as a growing trend in professional sports, but there has been a lack of empirical investigation regarding the topic.

[4] Every effort is made to develop a comprehensive sample of NFL PSL and STR sales transactions over the examination period. However, the sample utilized here does not necessarily reflect the entire secondary market and this is one limitation of the study.

the relative importance of the NFL, few studies have examined the demand for league attendance (Fizel, 2006). In part because of the league's relatively short regular season schedule, the NFL has traditionally played to capacity or near-capacity crowds. Noll's seminal work (1974), which was the first to estimate demand for pro football, noted that this capacity constraint explained almost all of the interteam differences in game-day attendance. Outside of stadium capacity, short-term team quality was the only other variable providing explanatory power despite including many of the traditional control variables and demand drivers. Noll also noted that since almost all sales were for advanced season tickets, short-term team quality was a fundamentally important factor in determining attendance.

Following Noll's contribution, Welki and Zlatoper (1994) examined 1991 NFL attendance data. The authors used a Tobit model to account for the upper bound censoring associated with fixed seating capacities. Welki and Zlatoper found that higher ticket prices decreased attendance, while a higher quality home team boosted game-day gate figures. In agreement with Noll's earlier findings, this study uncovered a negative relationship between per capita income and game-day attendance. Lastly and not surprisingly, the quality of the game matchup, and in particular, the quality of the home team was associated with increased attendance.

Welki and Zlatoper (1999) later completed a similar analysis with older data collected from the 1986 and 1987 NFL seasons. The results were comparable to their first study. This work again illustrated an inverse relationship between ticket prices and attendance, as well as a positive relationship between home team quality and attendance.

Also of interest was the finding that divisional matchups and contests played on non-Sundays resulted in attendance increases.

Putsis Jr. and Sen (2000) followed by analyzing NFL attendance demand in association with the league’s blackout rule. The authors expanded on the previous literature by using Tobit models to estimate demand for both individual game tickets and season tickets. Coinciding with previous empirical work, Putsis Jr. and Sen (2000) uncovered that demand for NFL attendance was inelastic as most NFL teams could easily increase ticket prices. The effect of income on demand was ambiguous as income was positively associated with season ticket demand but negatively tied to single game demand. Also in agreement with previous research was a positive relationship between team quality and demand as teams that reached the playoffs and had high winning percentages in the previous season saw increases in demand for attendance.

Coates and Humphreys (2007) estimated demand for the NFL, MLB, and NBA, but also incorporated the costs of ancillary attendance items specified by the Fan Cost Index (FCI). Using a generalized method of moments estimator, the authors concluded that deriving demand for the NFL was fundamentally different as compared to the other leagues. Ticket price and the FCI were found to be insignificant in determining demand in the NFL, but not so in MLB and the NBA. The authors claim that this between-league variation can be attributed to the fact that the NFL sees such a large percentage of their contests played to capacity.

More recently, researchers have explored estimating demand for the NFL product through the utilization non-attendance based response variables. Nagel and Rascher (2007) used franchise merchandise sales to pinpoint between-team variation in the demand for licensed products. Alternatively, Tainsky (2010) used television ratings as a proxy for NFL demand and uncovered strong consumer preferences for both home and visiting short-term team quality. In agreement with the findings of both Noll (1974) and Welki and Zlatoper (1994), an inverse relationship between personal income and demand was also revealed.

While each of the empirical pieces mentioned above has enhanced our understanding of the primary demand determinants in professional football, there are still some lingering issues with estimating demand for the sport. The two primary obstacles when estimating attendance in the NFL are the prevalence of advance season ticket sales and the venue capacity constraint.

As both Noll (1974) and Putsis Jr. and Sen (2000) previously identified, the decision on purchasing a season ticket package, which is achieved through ownership of a PSL or STR, is much different than the decision on purchasing a single game ticket. Specifically, when estimating demand, there are differences in many of the factors that would influence the purchase decision of a single game ticket as opposed to a full season package (Fizel & Bennett, 1989; Putsis Jr. & Sen, 2000). For example, game-specific factors, such as variables capturing weather conditions, opponent team quality, and temporal variables such as whether the game is played on a weekend or non-weekend are not appropriate for inclusion when estimating demand for PSLs and STRs.

Secondly, previous empirical research estimating demand for game-day NFL attendance has found that venue capacity has been the dominating variable. With this capacity constraint dictating the results, a lack of significance in other covariates has routinely been discovered. The work of Noll (1974), Putsis Jr. and Sen (2000), and

Coates and Humphreys (2007) previously identified this issue and the consensus is that there are fundamental differences between estimating attendance demand for the NFL and the other North American leagues.

The current work uses secondary market PSL and STR sales as a way to account for the two concerns outlined above. The use of PSLs and STRs as representative of overall demand for NFL football allows for the avoidance of the capacity constraint. By using secondary market sale price as the dependent variable, there is no upper limit truncation on attendance as is traditionally seen when using game-day attendance as the response variable. This choice of dependent variable also accounts for the fact that advance season ticket sales are now undoubtedly the norm in the NFL[5]. By analyzing secondary market PSL and STR sale prices, we are able to estimate demand for season ticket access based on market conditions at the time of the sale as opposed to the day of the game. Based on what has been previously noted regarding the combination of the high percentage of advanced season ticket sales, the negligible walk-up game-day customer and the explosion of the secondary market, it is a reasonable assumption that this is more representative of how consumers gain access to NFL season tickets.

2.4 Theoretical Model

In this market, fans are charged both a fixed upfront fee in the form of a PSL and then must also purchase season tickets at their face value. But while the PSL might technically be a two-part pricing mechanism, given that PSL sales do not increase the quantity of tickets sold in any way, the traditional economic model of two-part pricing falls short when evaluating this product in either the primary or secondary market. In

[5] Documentation of season ticket sales dominating over single-game ticket sales in the NFL is referenced as far back as Noll (1974).

other words, this type of two-part pricing is not a way to allow teams to capture consumer surplus that could not otherwise be captured with an increase in ticket price. This is clear given the propensity of sell outs for NFL games regardless of PSL sales. This section shows the economic intuition behind secondary market PSL sales.

Since the possession of a PSL or STR represents the right to buy season tickets for the lifetime of the stadium, the value of the PSL (or STR) can be written as

$$PSL_M = \sum_{t=1}^{N} \frac{G\left(E\left[P_{M,t} - P_{FV,t}\right]\right)}{(1+i)^t} \tag{1}$$

where PSL_M represents the market price for the PSL, *G* is the number of home games in a season, $E\left[P_{M,t} - P_{FV,t}\right]$ is the expected difference between the market price (P_M) and the face value (P_{FV}) of one ticket for one game in year *t*, *i* is the discount rate, and *N* is the number of seasons that the PSL gives the owner the right to purchase the tickets. The PSL only has a value if the expected market value of the ticket exceeds the face value. Furthermore, in most, if not all cases, there is no explicit agreement by league franchises as to what the face value of the tickets will be in the future. Yet, it appears to be the case that there is an implicit agreement that ticket prices will not increase so that it equals market demand.

Since it is not the focus of the chapter, this model does not describe why teams issue a PSL as opposed to simply increasing ticket price. However, there is a clear incentive for teams to issue PSLs because this revenue is not shared under the NFL collective bargaining agreement, while ticket revenue is shared.[6] So, it seems as though

[6] Presumably the league does this to encourage the construction and renovation of stadiums, which is the only time franchises are allowed to issue a PSL. This is discussed in further detail in the future research section at the conclusion of the chapter.

it is in the best interest of the franchise to capture as much of the ticket revenue through the sale of PSLs as possible.

Equation (1) illustrates the relationship between the demand for PSLs and the demand for tickets. Upon making the simplifying assumptions that there is a constant difference between the market price and the face value price of the ticket, and the PSL is a perpetuity, then the market price can be written as

$$PSL_M = \frac{G}{i}\left[P_M - P_{FV}\right] \tag{2}$$

Consequently, under these assumptions, demand for PSLs has a direct linear relationship with the demand for tickets. Hence, estimating PSL sale prices is analogous to estimating the market price of tickets. As is a common assumption in the sports economics literature, supply is fixed by the number of seats in the stadium. Likewise, the number of PSLs and STRs are limited by the number of seats in the stadium. Figure 2.1 shows the relationship between supply and demand in this market.

As stated earlier, this study estimates secondary market PSL sale prices and not the primary market PSL price. First, in the secondary market, there is no distinction between PSLs and STRs. Second, since the sellers in the secondary market are not actually producing the PSLs, they do not comprise the supply curve. Instead, the PSL market is analogous to a secondary bond market, where the buying and selling of PSLs reflects the dynamic changes in demand of the buyers and former buyers (current sellers). In other words, PSL sales are a consumer to consumer transaction where the consumers have changed their willingness to pay. Therefore, it is appropriate to use stadium capacity as the supply.

2.5 *Functional Form of Demand*

Following in line with the seminal work of Demmert (1973), Noll (1974), and Schofield (1983), the empirical modeling approach used in this study incorporates demographic, temporal, team specific, stadium specific and economic variables in an attempt to isolate the drivers of secondary market demand. As previously stated, NFL attendance demand studies have traditionally used game-day attendance as the dependent variable where the quantity demanded is specified based on the following function:

$$Q = f(P_r, I, P_o, S, T),$$

where Q represents the quantity purchased, P_r represents the price of the product, I represents the income of consumers in a market, P_o represents market population, S represents substitute entertainment options available to consumers, and T represents the tastes and preferences of consumers.

Instead of estimating a quantity as is done with attendance, this work instead uses price as the dependent variable, which necessitates a reorganization of the demand function. Taking into account additional factors which have the ability to influence secondary market sale prices, the functional form of demand in this study is specified by the following equation:

$$P_r = f(Q, D_e, F, S_s, T_e),$$

where P_r represents the secondary market sale price of the product, Q represents the quantity available, D_e represents demographic and economic factors in a specific market, F represents franchise specific factors, S_s represents stadium and seat location factors, and T_e represents temporal factors. The tastes and preferences of consumers in a specific market will be captured by the franchise, stadium and seat location specific variables outlined in the econometric models highlighted below.

2.6 Empirical Models

Three fixed effects models and three random effects models are used to estimate demand for NFL PSLs and STRs. The fixed effects models include team indicator variables for each franchise represented in the data set. The random effects models differ from the fixed effects models in that the team indicator variables are omitted from the estimations. Both the fixed effects models and random effects models include: 1) pooled PSL and STR observations, 2) PSL observations only, and 3) STR observations only. Following in line with the early empirical work of Demmert (1973), Noll (1974), and Schofield (1983) and the functional forms outlined above, these six models utilize demographic, temporal, team specific, stadium specific and economic variables in an attempt to isolate the drivers of demand in this secondary marketplace.

The general form of the random effects model follows:

LOGSEATPRICE = β_0 + β_1 YEAR06 + β_2 YEAR07 + β_3 YEAR08 + β_4 YEAR09 + β_5 LOGROW + β_6 SEATQUAL1 + β_7 SEATQUAL2 + β_8 SEATQUAL3 + β_9 SEATQUAL4 + β_{10} LOGSTADIUMAGE + β_{11} LOGSTADIUMCAPACITY + β_{12} TYPEPSL + β_{13} LOGTICKETPRICE + β_{14} AISLE + β_{15} WIN3 + β_{16} LOCALUNEMPLOYMENT + β_{17} LOGLIST + β_{18} NOLIST + β_{19} LOGPOPULATION + β_{20} LOGINCOME + β_{21} DOME + ε

The general form of the fixed effects model follows:

$$\text{LOGSEATPRICE} = \beta_0 + \beta_1 \text{ YEAR06} + \beta_2 \text{ YEAR07} + \beta_3 \text{ YEAR08} + \beta_4 \text{ YEAR09} + \beta_5 \text{ LOGROW} + \beta_6 \text{ SEATQUAL1} + \beta_7 \text{ SEATQUAL2} + \beta_8 \text{ SEATQUAL3} + \beta_9 \text{ SEATQUAL4} + \beta_{10} \text{ LOGSTADIUMAGE} + \beta_{11} \text{ LOGTICKETPRICE} + \beta_{12} \text{ AISLE} + \beta_{13} \text{ WIN3} + \beta_{14} \text{ LOCALUNEMPLOYMENT} + \beta_{15} \text{ LOGPOPULATION} + \beta_{16} \text{ LOGINCOME} + \beta_{17:38} \text{ TEAM INDICATORS} + \varepsilon$$

Non-indicator variables are logged to help alleviate any heteroskedasticity problems. Table 2.2 lists the variables and provides a brief explanation of each.

2.7 Data Description

The data used in this analysis are secondary market PSL and STR sales transactions occurring between 2005 and 2009. 3,821 observations were collected in total. All transactions were gleaned from either www.seasonticketrights.com or www.ebay.com. The transactions from the former location were exclusively fixed-price transactions, while data from the latter were either fixed-price or timed auction style transactions. Only actual sales transactions are included in the sample, as many PSLs and STRs were listed for sale on these secondary market locations over the examination period, but were not sold at the listed asking price.

The dependent variable, *LOGSEATPRICE,* is the logged per seat sale price of the total PSL or STR sales transaction. In certain cases where ancillary items, such as yearly parking passes were included in the sale, the value of the sale price was adjusted accordingly. Other explanatory variables that are specific to the transaction were also collected from the aforementioned websites. These variables include *AISLE* and *LOGROW*. *AISLE* is an aisle seating indicator and *LOGROW* is the logged row number associated with the season ticket location. A lower row number represents a seat location closer to the field of play within a section and therefore a higher quality seat location.

SEATQUAL1, *SEATQUAL2*, *SEATQUAL3*, *SEATQUAL4* and *SEATQUAL5* are indicator variables included to capture consumer preferences for specific seating locations within a venue. Information on the exact seating locations for each variable is available in Table 2.2. While *LOGROW* attempts to capture seat quality as proxied by row location within a specific section, these variables account for the overall quality of the seating location in terms of closeness to both midfield and the field of play.

Stadium related data, such as *LOGSTADIUMAGE*, *LOGSTADIUMCAPACITY*, *LOGTICKETPRICE* and *DOME* were collected from each team's official website. *LOGSTADIUMAGE* represents the logged age of the stadium. *LOGSTADIUMCAPACITY* represents the logged number of seats in the venue and controls for the supply of PSLs and STRs available in the secondary market. *LOGTICKETPRICE* is the logged per game season ticket price (face value) of the seat. Ticket prices associated with specific seating locations were not able to be obtained for the Pittsburgh Steelers and the Washington Redskins. Subsequently, average ticket prices for both standard and premium seating locations (collected from Team Marketing Report) were used for each franchise[7]. *DOME* is an indicator variable equal to one if the NFL franchise plays their home games in a dome or retractable roof venue.

TYPEPSL is an indicator variable coded equal to one if the sale is a PSL transaction and coded zero if the sale is a STR transaction. *TYPEPSL* may affect secondary market sale price since this variable compares secondary market sales for

[7] The empirical models were estimated both with and without the Redskins and Steelers data in order to determine if using average ticket prices for these two franchises would significantly impact the estimation results. The differences in the results were negligible.

franchises with PSL programs against those with STR programs. Franchises with PSL programs were able to utilize this sales strategy due to significant demand for their product. Alternatively, clubs not implementing PSL programs fall into one of three possible scenarios. These franchises 1) have either constructed or renovated new venues and decided not to use a PSL sales program, 2) did not have enough season ticket demand to justify implementing a PSL sales program, or 3) have not built a new venue or renovated their current venue since 1993.

LOGLIST is a variable specifying the logged length of each franchise's season ticket waiting list. This information was collected from numerous sources, including official team websites, www.forbes.com, and from telephone correspondence with NFL franchise ticket sales employees. The length of a team's season ticket waiting list could affect secondary market sale prices as a longer waiting list to acquire season ticket rights directly through the franchise may spur fans to search for purchase options in the secondary market. *NOLIST* is an indicator variable coded equal to one if a team does not have a waiting list to purchase season tickets directly from the franchise.

WIN3 represents the cumulative team winning percentage over the three seasons prior to the PSL or STR sale.[8] A higher three-year winning percentage would represent a higher quality on-field product. If preferences for high team quality hold, differences in three-year team winning percentage should affect secondary market sale price.[9]

---

[8] Other short-term and long-term measures of team quality were tested, and 3-year win percentage was selected because it showed the most explanatory power.

[9] Level of game uncertainty variables and visiting team quality variables are not included in these models because the consumer is purchasing access to an entire season ticket package, representing a different purchase decision as compared to a single game ticket, in which the quality of the visiting team would be a significant determinant on the attendance decision.

YEAR05, *YEAR06*, *YEAR07*, *YEAR08*, and *YEAR09* are indicator variables representing the year in which the PSL or STR was sold.[10] Secondary market sales in 2005 were used as the measurement baseline. Twenty-two team indicator variables are included in the fixed effects models, representing each of the NFL franchises in the data set.[11] Each variable is labeled using the home city or state of each franchise. The Baltimore team indicator is excluded from the estimation as it is used as the baseline for the comparison of other franchises.

LOCALUNEMPLOYMENT is the local Metropolitan Statistical Area (MSA) unemployment rate during the month of the PSL or STR sale. This variable was gathered from the United States Bureau of Labor Statistics (http://www.bls.gov/) and was included to measure the effect of the health of the local economy on secondary market sale prices. A higher local unemployment rate during the month of the sale would represent a weaker local economy, making it reasonable to infer that fluctuations in rates could alter secondary market sale prices.

LOGPOPULATION is logged MSA population during the year of the sale and was collected from the United States Census Bureau (www.census.gov). Because population is one of the five primary demand shifters, a larger population in a defined metropolitan area would suggest that there are more people willing to pay for NFL season tickets at every price point. *LOGINCOME* is logged MSA per capita personal income during the year of the sale and was gleaned from the U.S. Bureau of Economic Analysis

[10] Various linear and non-linear trend variables were tested and indicator variables based on the year of sale were selected because they controlled for more variation in the dependent variable.
[11] A small number of observations for the following franchises were eliminated from the data set: Arizona, Kansas City, New York Giants, and New York Jets. Observations from both New York franchises were withheld because these sales represented "an option to purchase" a PSL from the franchise prior to the opening of their new venue in 2010. These sales did not represent actual secondary market PSL transactions. The Kansas City and Arizona observations were removed because of questionable sales terms.

(http://bea.doc.gov/).[12] If PSLs or STRs are normal goods, a MSA with a larger per capita level of income would suggest an increase in secondary market sale price.

Table 2.3 shows the summary statistics for all variables. One benefit of this data is the ability to examine the total cost of attendance for the consumer separated out between the PSL or STR fee and the face value price of season tickets. If we assume a discount rate of 10%, and given that there are 10 games per season, our sample shows that on average 19.1% of the total market price of attendance is paid in the form of a PSL or STR (11.4% if we use a 5% discount rate). This percentage represents the average percent of the total discounted payment a fan pays for tickets through a PSL or STR. However, if we sum up the total value of the PSLs and STRs in the sample and compare them with total discounted market price, PSLs/STRs represent 22.5% of the total value (12.7% at a 5% discount rate). The maximum value in our sample is 77.6% (63.5% if we use a 5% discount rate) of the total discounted market price of attendance paid through a PSL. However, the possibility exists that our numbers may be skewed as this sample does not account for the entirety of the secondary market.

2.8 Results

Table 2.4 outlines the demand estimation results for the fixed and random effects models using the pooled PSL and STR observations. Robust standard errors were specified in all six models based on significant Breusch-Pagan / Cook-Weisberg and White (1980) test results. The fixed effects model for the pooled data produces a R^2 value of .7561. Twenty-four of the thirty-eight covariates are significant at the 0.01 level with

[12] Because 2009 MSA per capita income was not available at the time of analysis, all 2009 observations use 2008 income data. This causes *LOGINCOME* to be collinear with the team indicator variables in the STR fixed effects model.

three other covariates also showing significance at traditional levels. *TYPEPSL*, *LOGSTADIUMCAPACITY*, *LOGLIST*, *NOLIST* and *DOME* were withheld from the model since these variables do not vary within each team and are therefore perfectly collinear with the team indicator variables.

A key finding from this model illustrates strong consumer preferences for high quality seating locations. Specifically, *LOGROW*, or the row location associated with the PSL or STR sale is highly significant as seats closer to the field of play are associated with higher secondary market sale prices. Additionally, *SEATQUAL1*, *SEATQUAL2*, *SEATQUAL3 and SEATQUAL4* are all highly significant. This suggests significant differentiation in demand based on seating location and strong consumer preferences for seating locations located in the lower level and on the sidelines. The results from these variables support the theory that fan preferences for higher quality seat locations are a primary driver of demand on the secondary market. Additionally, this illustrates significant differences in consumer willingness to pay based on seating location.

Team winning percentage over the previous three years also has a large impact on the dependent variable, as a one unit increase in win percentage produces a 228.38% [exp(1.189) = 3.2838 => 3.2838 – 1 x 100 = 228.38%)] increase in secondary market sale price. This finding supports previous empirical work, including Noll's (1974) seminal piece on pro football demand, which noted that consumer preferences for short-term team quality was the largest factor impacting attendance outside of stadium capacity.

The coefficient on *LOCALUNEMPLOYMENT* illustrates the effect of economic conditions on the demand for NFL football. The significant inverse relationship between unemployment rates in a market and PSL and STR sale prices suggests that all else equal, secondary market sale prices of PSLs and STRs are reduced when the local MSA unemployment rate increases.

In examining the team indicators, sixteen of the twenty-two were significant at traditional levels against the baseline of Baltimore. Cleveland, Tennessee and Chicago were the PSL franchises that were shown to be the most statistically similar to the baseline. Interestingly, two of those three franchises are relatively new to their respective cities. Baltimore and Tennessee are two of the most recent NFL franchises to relocate and Cleveland was awarded a franchise following the Browns move to Baltimore in 1996.

The pooled data random effects model produces a R^2 value of .6936. Fifteen of the twenty-one covariates in this model are significant at the 0.01 level. Many of the results found in the pooled fixed effects model are mirrored in the random effects model and will not be rehashed here. Despite this, the pooled random effects model does provide relevant information. Specifically, *TYPEPSL* has a large impact on determining secondary market sale price. If the transaction is a PSL sale as opposed to a STR sale, secondary market sale price is expected to increase by 726.48% [exp(2.112) = 8.2648 => 8.2648 – 1 x 100 = 726.48%)]. This finding makes intuitive sense for a few reasons. First, in the primary market, PSLs have an acquisition fee tied to them when they are purchased directly from the franchise while STRs do not. This means that if a PSL becomes available on the secondary market, all else equal, the seller is likely to both set and receive a higher sale price as compared to a STR, because of the higher perceived value of the product. Secondly, franchises which enacted PSL programs in the first place, did so because of sufficient demand for NFL football in their market. Those which allow season ticket rights transfers through STRs either did not have sufficient demand to sell PSLs or likely chose not to sell PSLs in exchange for the ability to alter ticket prices based on demand fluctuations (for example, the New England Patriots). Regardless, this finding demonstrates that there are significant differences in demand for the rights to attendance for PSL franchises as compared to STR franchises.

This model also highlights the relationship between PSLs, STRs and the face price of season tickets. The negative and significant coefficient on *LOGTICKETPRICE* illustrates that as the face value price of a season ticket increases, the price a consumer is willing to pay for the corresponding PSL or STR decreases, which is consistent with our model.

The pooled random effects model also illustrates a strong positive relationship between *LOGPOPULATION* and secondary market sale prices. As MSA population increases, secondary market sale prices also increase, as a 10% increase in logged MSA population results in a 6.34% increase in secondary market PSL and STR sale prices [1.1^.645 = 1.0634].

Also of interest are the negative coefficients of increasing magnitude on the *YEAR08* and *YEAR09* variables. The results of these variables suggest that all else equal, PSL sale prices have declined over time as compared to the 2005 baseline. Since other influential variables are controlled for, these results could suggest that a PSL decreases in value over time as a PSL owner's window to purchase season tickets from the franchise shrinks. These findings could represent the possibility of asset depreciation or simply consumer awareness of the aforementioned scenario.

Table 2.5 shows the estimation results using only the PSL data. The fixed effects model produces a R^2 value of .7356. Fifteen of the twenty-seven covariates in this model are significant at the 0.01 level with two others significant at the 0.05 level. The Baltimore team indicator was again treated as the baseline and *TYPEPSL*, *LOGSTADIUMCAPACITY*, *LOGLIST*, *NOLIST* and *DOME* were excluded from the model due to perfect collinearity with the team indicator variables. This model produces results that are similar to the previous pooled fixed effects model, which is due to 92.1% of the data being comprised of PSL observations.

The positive and significant effect of LOGINCOME implies that secondary market PSL sale prices have increased along with per capita levels of income in a franchise's MSA. Additionally, *LOGSTADIUMAGE* shows a strong negative effect, suggesting that once team effects are accounted for, fan preferences for newer venues emerge. This could also suggest that consumers are willing to pay higher PSL prices in venues where they hold an option to purchase season tickets directly from the franchise for a longer period of time. Once again, *LOGROW*, *SEATQUAL1* and *SEATQUAL3* show strong effects on the dependent variable, supporting the idea that consumer preferences for prime seating locations drive PSL sale prices on the secondary market.

The random effects model for the PSL observations produces a R^2 value of .6706. Fifteen of the twenty-one covariates in this model are significant at the 0.01 level. Because this model includes only PSL observations, *TYPEPSL* is omitted from the estimation.

LOGROW and each of the *SEATQUAL* variables, which measure the quality of the seating and entertainment experience, were again highly significant in determining secondary market sales price. The positive effect of *LOGPOPULATION* again supports population being one of the five primary drivers of demand. Consumer preferences for team quality are also strong as illustrated by the significant positive coefficient on *WIN3*.

LOGLIST produces a positive and significant effect on secondary market PSL sale prices. This suggests that that secondary market PSL sale prices increase along with the primary market length of a franchise's season ticket waiting list. *DOME* also illustrates a significant negative relationship with the dependent variable which could be interpreted as reduced consumer preferences for the indoor NFL product.

Table 2.6 shows the estimation of only STRs. Because there are fewer STR observations, these estimations are not as robust. The fixed effects model produces a R^2 value of .4399. Only four of the twenty-two variables are significant at standard levels. The results vary from the previous models as the lack of significant variables is evident. No 2005, 2006 or 2007 observations exist in this portion of the data, so 2008 secondary market STR sales are used as the baseline and only the 2009 year indicator variable is included in the estimation. The Buffalo team indicator is excluded from the estimation and is used as the baseline. Additionally, *TYPEPSL*, *LOGSTADIUMCAPACITY*, *LOGLIST, NOLIST*, *LOGINCOME* and *DOME* are excluded from the estimation due to perfect collinearity with the team indicator variables.

LOGROW is once again the variable of largest influence on secondary market sale price. Coupled with *SEATQUAL1* and *SEATQUAL3* showing a significant positive effects on the dependent variable, these results again support the notion that consumer preferences for high quality seating locations drive demand for access to NFL season tickets on the secondary market.

The inclusion of team indicator variables results in two covariates which have traditionally been positively linked to demand to show non-significance. Namely, *LOGPOPULATION* and *WIN3* fail to produce any significant effect on the dependent variable. The non-significant relationship between WIN3 and the dependent variable is particularly unique based on the strong and consistent positive relationship between team quality and demand in the sports economics literature. This result suggests that short-term team quality does not significantly impact secondary market sale prices of season ticket rights.

Also of note is that only one of the ten team indicators in this model is significant at any level (New Orleans at the 0.05 level). This is a notable difference from the first two fixed effects models, where the bulk of team indicators were significant. This suggests that the vast majority of STR franchises included in the data set do not differ significantly from Buffalo, the baseline franchise. This can be interpreted further by stating that this collection of STR franchises does not exhibit characteristics beyond what is controlled for by other variables in this model that would significantly differentiate them from each other.

Lastly, the STR random effects model produces a R^2 value of .4339 with six of the eighteen covariates showing significance at standard levels. L*OGROW* and *SEATQUAL1* again produce significant effects, supporting the notion that fan preferences for high quality seating locations drive prices on the secondary market. *WIN3* remains non-significant in the random effects model supporting the result seen in the fixed effects STR only model.

The random effects model also produces a significant negative effect on the *NOLIST* indicator variable, signifying reduced secondary market STR prices for those franchises in the data set without season ticket waiting lists. Interestingly though, there is also a significant negative effect on the *LOGLIST* variable, which would suggest that as the length of a team's season ticket waiting list increases, secondary market STR sales prices decrease. At first glance, this result appears to lack sensibility, but a plausible explanation exists. The extensive, yet questionable[13] season ticket waiting list length of 155,000 that the Washington Redskins claim to possess could be impacting the direction and significance of this variable. It would seem reasonable that if a waiting list of this size existed, secondary market sale prices for this franchise would be higher.

2.9 Comparing PSL and STR Markets

The empirical models outlined in this chapter illustrate clear differences in demand for NFL football between PSL and STR markets. Specifically, the pooled random effects model shows that PSLs sell for significantly higher prices on the secondary market as compared to STRs. As a result of this finding, a closer examination of the cost of attendance between PSL and STR markets is warranted. The following section outlines secondary market sale prices for PSLs and STRs by seating location, sale prices by franchise, and sale prices by seating location and franchise over the examination period.

Table 2.7 demonstrates average secondary market sale prices for PSLs and STRs in each of the five seating locations identified in the empirical models. This table shows

[13] Personal correspondence with a member of the Washington Redskins ticket sales staff prior to the start of the 2010 season revealed immediate purchase availability for premium seating locations at FedEx Field. Based on this information, it is questionable whether or not a season ticket waiting list of this length is accurate.

that over the examination period, the average secondary market sale price paid to acquire the rights to purchase a season ticket directly from a NFL franchise was $2,848.47. Naturally, acquisition rights for seating locations closer to the field of play and closer to midfield were substantially higher, as is evidence by the higher sale prices for seating locations one and three.

Average secondary market sale prices for PSLs only and STRs only are illustrated in Tables 2.8 and 2.9, respectively. In support of the findings from the pooled random effects estimation, these two tables confirm the disparity between PSL and STR sale prices on the secondary market. At a sale price of $3,060.38, the average PSL sells for approximately ten times the price of an average STR on the open market. This result confirms the stark difference in attendance demand existing between PSL and STR markets. Another interesting disparity between PSL and STR consumers is visible in the average sale prices for seat quality five, or end zone seating locations. While differences in sale prices for the four other seating locations are comparable between PSLs and STRs, the opposite is true with end zone seating. While end zone seating is clearly the least desirable (according to average sale price) in STR markets, it ranks as third most desirable in PSL markets. While this result could be due to simple differences in the distribution of upper level versus lower level end zone seating, it could also be an indicator of differences in consumer preference between markets. While many would be quick to dismiss this theory, the differences in demand between PSL and STR markets is a topic worthy of future examination.

Table 2.10 highlights average secondary market PSL sale prices by franchise. While the distribution of seat locations may vary by team, this table shows that Chicago

Bears and Pittsburgh Steelers PSLs sell for approximately $9,000 each – the highest amount of all of the teams in the sample. Alternatively, the franchises with the lowest secondary market sale prices are the St. Louis Rams and Cincinnati Bengals with average transaction prices in the $655-$815 range.

Each of the six empirical models illustrates both demand for high quality seating locations and significant differences in demand between seating locations. In order to further expand on these important findings, Table 2.11 demonstrates secondary market PSL sale prices by franchise for the highest price seating location – lower level seating between the 30-yard lines. In agreement with the general results in Table 2.10, the averages in Table 2.11 again show that Chicago Bears and Pittsburgh Steelers PSLs in this prime seating area sell for the highest prices on the secondary market. St. Louis and Cincinnati PSLs again trail the pack by a wide margin in secondary market sale prices.

Table 2.12 provides average secondary market STR sale prices for all STR franchises in the sample. While STRs for many of these franchises sell for an average of $115-$245 each, there are several franchises with STRs selling for significantly higher prices on the secondary market. For example, New Orleans Saints and Miami Dolphins STRs trade hands for approximately $800 apiece, while Indianapolis Colts STRs sell for approximately $550 each. Table 2.12 shows that Washington Redskins STRs sell for the highest secondary market prices while Tampa Bay Buccaneers STRs sell for the lowest. However, because of the low number of sales transactions for Redskins and Buccaneers STRs, there is less certainty regarding the true average secondary market value of these items.

STR sale prices for seating locations in the lower level between the 30-yard lines are shown in Table 2.13. Despite only three franchises having sales transactions in this seating area, this table is useful for comparison against the PSL teams shown in Table 2.11. Again, New Orleans STRs in prime seating locations sell for higher prices as compared to the rest of the STR franchises. However, their average secondary market sale price in this seating category is still lower than all of the PSL franchises.

The tables presented in this section illustrate that PSLs sell for significantly higher prices on the secondary market as compared to STRs. However, as established in the introductory and theoretical sections of this chapter, the purchase of a secondary market PSL or STR does not comprise the entire cost of attendance to the consumer. The purchase of a PSL or STR only grants the consumer the right to purchase a season ticket directly from the franchise on a yearly basis. The following section illustrates examples of the full cost of attendance for secondary market PSL and STR purchasers based on seating location.

Table 2.14 provides a breakdown of the combined cost of attendance for secondary market PSL purchasers while Figure 2.2 offers a visual presentation. Table 2.14 demonstrates the combined cost of attendance over a ten-year period by summing the average secondary market PSL purchase price for a given seat location and the average face value season ticket price plus standard yearly price increases for the given seat quality location. Because the consumer is required to purchase a ticket to ten total home games per season (eight regular season and two preseason games) the ticket price for each season is multiplied by ten. Figure 2.2 illustrates that if a consumer purchases the average secondary market PSL and maintains his or her PSL rights for ten years through the purchase of season tickets for that entire ten-year period, 21% of the total cost is paid in the form of the initial PSL purchase and the remaining 79% is paid through season ticket fees. Also of note is the substantial variation in the total cost of attendance between seating locations. Specifically, the average ten-year cost for a PSL and season ticket in the highest quality seat location is approximately 3.4 times greater than the lowest quality seating location. This illustrates the premium consumers are willing to pay for the highest quality seating locations in NFL stadiums.

Table 2.15 serves as a comparison against Table 2.14 by outlining the total cost of attendance for a secondary market STR purchaser over a ten-year period. Again, the purchase of a STR grants the consumer the right to purchase a season ticket directly from the franchise, so the initial STR purchase is only a fraction of the total cost of attendance. Analogous to the PSL example, this illustration assumes a 3% yearly increase in the face value season ticket price and also that the STR fee was paid in full and not financed at the time of purchase. Table 2.15 shows that for lower level seating between the 30-yard lines (seat quality 1) STR face value ticket prices are comparable to that of PSL franchises. Additionally, STR face value ticket prices tickets in the upper level between the 30-yard line and the end zone line are actually more expensive than PSL season ticket prices in the same location. However, for the three other seating locations, PSL face value ticket prices are significantly higher than their STR counterparts. Lastly, Figure 2.3 illustrates that secondary market STR purchases account for only 4% of the total cost of attendance in STR markets while season ticket expenses comprise the remaining 96%. This result is considerably different that the total cost breakdown in the PSL case.

This section along with the results from the empirical models clearly demonstrates differences in demand between PSL and STR markets. As a result, a closer examination of the distinctions between National Football League PSL and STR markets appears to be necessary. The background section at the beginning of the chapter outlines a few key differences between PSL and STR markets which have the ability to influence demand. For example, by nature of the NFL PSL program, franchises in PSL markets, on average, have newer venues as compared to STR franchises. Previous empirical research has shown that new stadiums have the ability to increase attendance demand (for example, Clapp & Hakes, 2005). In addition, PSL franchises likely have implemented PSL programs due to strong existing demand for their product. After all, adequate attendance demand must exist in order to conduct a successful PSL program. The case of the Oakland Raiders is a prime example of the need for sufficient market demand to support a PSL program (Dickey, 2004).

Beyond these factors, a closer look at the demographics in both PSL and STR markets has the potential to shed light on this state of affairs. Tables 2.16 and 2.17 provide population and income data for PSL and STR markets, respectively. During the 2005-2009 examination period, the average population in NFL PSL markets was 3.86 million, while the average population in STR markets was 3.10 million. The average per capita personal income in PSL and STR markets was $42,992 and $45,452, respectively. Based on these figures, it appears that PSL markets are approximately 24.5% larger in terms of population, but 5.72% less wealthy than STR markets. However, the San Francisco 49ers and Oakland Raiders share a single market and are both STR franchises. Once accounting and adjusting for this fact, the results change substantially. The far right column in Table 2.17 shows adjusted population data for STR markets. Following this adjustment, on average, PSL markets are approximately 42.4% larger than STR markets. It is inappropriate to adjust income in this case, because having two franchises in a single market does not alter the income levels of consumers in that market. Therefore, a cursory examination of demographics in PSL and STR markets illustrates that PSL markets are significantly larger on average than STR markets and slightly less wealthy. Following in line with Fort's (2003) description of demand shifters in sports, a plausible explanation for the substantial variation in demand between PSL and STR markets can be contributed to the significant population discrepancies between the two market types. This finding is supported by the positive and significant coefficients on the *LOGPOPULATION* variables in the pooled and PSL only random effects models presented earlier in the chapter.

2.10 Team Quality and Sale Prices in PSL and STR Markets

Perhaps the most important finding in the empirical modeling section of this chapter is the variation in the relationship between team quality and demand for PSL and STR franchises. Specifically, the modeling results show a strong positive relationship between short-term team quality and secondary market PSL sale prices. This positive relationship between team quality and attendance demand is a fundamental empirical finding in the sports economics literature and is discussed in great detail by Borland and Macdonald (2003) in their survey of the sports demand literature. However, in the empirical estimations completed in this chapter, the recurrent positive relationship between team quality and demand is not found in STR markets. Based on the historical strength of this association, this finding is certainly of interest. The following section delves deeper into the relationship between team quality and demand in the NFL.

Table 2.18 provides information on the distribution of three-year cumulative team winning percentages for PSL franchises along with the corresponding average secondary market PSL sale prices. Figure 2.4 provides the same information in chart format. Visible is a gradual increase in PSL sale prices as cumulative team winning percentage increases from a low of .271. Beginning with a winning percentage of .583 is a steep increase in secondary market PSL sale prices which holds throughout the remainder of the distribution.

The distribution of three-year team winning percentages for STR franchises and corresponding secondary market STR sale prices is shown in Table 2.19 and Figure 2.5. Unlike the case of PSLs where sale prices increase along with team quality, this association does not hold with STR sale prices. Similar to the PSL example, there is a slight gradual increase in sale prices in the beginning of the winning percentage distribution. However, once team winning percentage rises above .333, there is no clear relationship evident between team quality and secondary market STR sale prices. Figure 2.5 illustrates this nicely with sale price variation both above and below the mean transaction price of $370.

Despite the fact there are less STR observations as compared to PSL observations, which renders the STR results less robust, the lack of a significant relationship between team quality and secondary market STR sale prices is an important finding. These results point to the possibility of two classes of NFL franchises. The first class is made up of PSL franchises which show stronger demand for the live NFL product and also strong sensitivity to fluctuations in short-term team quality. The second class is made up of STR franchises which on average show significantly reduced demand for the live NFL product and no statistically significant relationship between team quality and attendance demand. Future research targeting a more comprehensive understanding of consumer demand for NFL football may look to investigate this relationship in greater detail.

2.11 PSLs as Depreciable Assets

Significant negative effects on the year of sale indicator variables from the pooled and PSL only random effects models suggest PSLs decrease in value over time when holding all else equal. Based on the fact that the empirical models control for factors such as income, population, market unemployment, team quality, stadium quality and seating location, it is reasonable to suggest that PSLs are depreciable assets based on the modeling results. As described in the background section, a PSL owner's window to purchase a season ticket directly from the franchise shrinks as time advances. As the club moves closer and closer to either constructing a new venue or renovating their existing venue, the opportunity arises for the NFL franchise to establish a new PSL program. With a new PSL program, even existing PSL owners would be forced to purchase a new PSL in order to maintain their rights as a season ticket holder. Under this scenario, the value of the PSL in the open market decreases as the number of remaining seasons the club will play its home games in its existing venue decreases. The math behind this outcome is clearly visible in equation (1) in the Theoretical Model section of the chapter.

Based on this empirical result, the following section will take a closer look at secondary market PSL sales over time. Tables 2.20 and 2.21 provide year-by-year sale prices for PSLs in seating locations one and three, respectively. Both tables illustrate that in general PSL prices have decreased from the beginning to the end of the examination period. For seating quality one, shown in Table 2.20, it is clear that secondary market PSL sale prices have declined steadily for the Chicago, Cincinnati, Cleveland, Houston, St. Louis and Tennessee franchises. This finding is also provided visually in Figures 2.6 and 2.7. Carolina and Philadelphia PSLs have also gradually declined in value, but not before slight increases in value in the earlier years of the data set. Overall, a gradual decline in secondary market PSL sale prices is evident.

Table 2.21 provides secondary market sale prices for seat quality three, which is seating in the lower level sidelines between the 30-yard line and end zone. Again, the general trend is a gradual decrease in the secondary market value of PSLs from the start to the end of the examination period. Table 2.21 echoes this point by tracking a steep decline in PSL sale prices for Baltimore, Carolina, Chicago, Houston and Tennessee. A visual treatment of this result is provided in Figure 2.8. Despite this general decline, Pittsburgh and Seattle PSL sale prices in this seating area have actually increased slightly from 2005 to 2009. However, there is a clear general downward trend in secondary market PSL sale prices as a whole. Along with the significant and negative year indicators found in the empirical models, substantial evidence exists to suggest that PSLs are depreciable assets. Future research should look to examine this issue further.

2.12 Directions for Future Research

In the landscape of empirical research focused on professional sport demand, contributions specific to professional football are relatively scarce when compared to the overall attention paid to the topic. This is primarily due to the venue capacity constraint which has hampered attendance demand estimations specific to the NFL (Noll, 1974).

This chapter avoids the venue capacity constraint through the utilization of a unique data set of secondary market PSL and STR sale prices to estimate demand for access to NFL attendance. The completion of this study has generated two key outcomes that are appropriate for future investigation.

The first focuses on the relationship between team quality and demand in NFL markets. Empirical modeling results in this chapter suggest two classes of NFL franchises – one with sizeable demand for the live product and a second with significantly reduced demand. Despite the relative popularity of NFL football, there is a dearth of empirical work focused on the league (Fizel, 2006). Hence, a need exists for an extension to this chapter which examines the degree to which NFL demand varies by market. The results have to ability to inform regarding pricing, large market versus small market subsidization, advertising fees, radio and new media broadcast rights, and franchise valuation, just to name a few.

A second area appropriate for future research investigates the PSL sales program as an incentive mechanism promoting the construction of new venues. The revenue generation benefits of a PSL program to the franchise are clear. However, under the NFL Collective Bargaining Agreement, franchises are only eligible to initiate a PSL sales program in conjunction with the construction of a new venue or renovation of an existing venue. Based on the revenue sharing benefits and the substantial new revenue generation opportunities associated with conducting a successful PSL program, it is reasonable to suggest that the PSL program is in place specifically to act as an incentive for the construction of new venues. This system has the ability to increase both individual franchise revenues and league-wide revenues through the NFL revenue sharing arrangement. While there is a substantial amount of literature focused on revenue sharing and its subsequent impact on investment in playing talent, there are also other important investments made by franchises. Specifically, the effect of revenue sharing on alternative investments, such as investments in stadia, has yet to be analyzed. This future project will attempt to model and illustrate through examples, the incentives and motives of policies such as the NFL PSL program. Initial results suggest that leagues should differentiate between various types of revenues when developing revenue sharing systems. More specifically, early evidence suggests that shifting revenue sharing towards media revenues and away from stadium revenues enhances the incentive for individual franchises to invest in a new venue. The NFL PSL program illustrates this finding nicely.

2.13 Summary and Conclusions

The sale of personal seat licenses as a franchise revenue generation mechanism has become a common practice in the NFL over the past two decades. Because of the substantial demand for NFL football, a strong secondary market has emerged where current NFL season ticket holders sell their PSLs and STRs over the internet. This has led to an opportunity to estimate demand for NFL attendance using secondary market PSL and STR sale prices. Deriving demand using secondary market data allows for the avoidance of the venue capacity constraint which has previously dominated NFL attendance demand estimations. The utilization of this unique data has reopened a line of investigation which has been relatively inactive since Noll (1974).

With the venue capacity constraint eliminated, fan preferences for high quality seating locations are found to be the strongest drivers of demand for access to NFL season tickets. *LOGROW* is significant in all six models at the 0.01 level. Additionally, each of the *SEATQUAL* variables is significant at the 0.01 level in the pooled and PSL only models. The STR only models also illuminate strong consumer preferences for specific seating locations[14].

More importantly, this research demonstrates clear differences in attendance demand among NFL markets. As one would expect, PSLs sell for significantly higher prices than STRs. This is evidenced by *TYPEPSL* producing the strongest influence on sale price in the random effects pooled data model. *WIN3*, or short-term team quality, is significant at the 0.01 level in both the pooled data models and the PSL only models, but not in the STR only models. This suggests that secondary market sale price is less likely to fluctuate in STR markets as short-term team quality changes. Alternatively, in PSL markets, demand for PSLs will increase along with an increase in on-field performance. A clear and persistent finding in the team sports literature illustrates the positive relationship between team quality and demand (Borland & Macdonald, 2003). Therefore, evidence that certain markets are not sensitive to changes team performance is certainly an important finding. Future research on NFL demand may look to examine this closer. Lastly, the empirical models provide evidence that the value of PSLs decline in value over time – suggesting they are depreciable assets. This result is supported by several negative coefficients on the year of sale indicators in multiple models.

In regards to industry implications, it appears that in the primary market the vast majority of NFL franchises which have the prerequisite demand needed to implement a PSL program have done so. This is supported by the significant differences in secondary

[14] Our results illuminate consumer preferences for high quality seating locations. However, stadium configurations are such that there are fewer "good" seating locations as opposed to "bad" seating locations. Therefore, these estimations illustrate consumer preferences given the relative scarcity of the given seat location.

market PSL and STR sale prices. The notable exception is seen in the case of the New England Patriots, who appear to have decided to simply increase the face value price of their season tickets instead of selling PSLs. This allows the Patriots to increase ticket prices in accordance with changes in demand, but the trade-off is that they are forced to share a portion of their increase in ticket revenues under the league's revenue sharing system.

Another intriguing question asks why NFL franchises have almost uniformly decided not to restrict the sale of PSLs and STRs on the secondary market. Similar to what the New England Patriots and Green Bay Packers do, clubs could simply require season ticket holders to return PSLs and STRs to the team if the consumer no longer wishes to purchase their season ticket package. This would allow the franchise to recapture unclaimed consumer surplus that is lost when the club allows season ticket holders to sell their PSLs and STRs on the secondary market. On the other hand, the current scenario where the majority of franchises sell consumers the unrestricted rights to seats likely increases the original PSL value. An interesting extension to this work would examine that question further.

2.13 References

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

Clapp, C.M., & Hakes, J.K. (2005). How long a honeymoon? The effect of new stadiums on attendance in Major League Baseball. *Journal of Sports Economics*, 6(3), 237-263.

Coates, D.C., & Humphreys, B.R. (2007). Ticket prices, concessions and attendance at professional sporting events. *International Journal of Sport Finance*, 2, 161-170.

Demmert, H. G. (1973). The economics of professional team sports. Lexington, MA: Lexington Books.

Dickey, G. (2004, December 15). Raiders PSLs expiring; Officials seek solution. *San Francisco Chronicle*. Retrieved October 15, 2010 from http://www.sfgate.com/cgi-bin/article.cgi?f=/c/a/2004/12/15/SPGGQAC6J41.DTL&ao=all#ixzz1HXXVMDsN

Drayer, J., & Martin, N.T. (2010). Establishing legitimacy in the secondary ticket market: A case study of an NFL market. *Sport Management Review*, 13(1), 39-49.

Fizel, J. (2006). The National Football League conundrum. In J. Fizel (Ed.) *Handbook of Sports Economics Research* (pp. 170-171). Armonk, NY: M.E. Sharpe, Inc.

Fizel, J.L., & Bennett, R.W. (1989). The impact of college football telecasts in college football attendance. *Social Science Quarterly*, 70, 980-988.

Fort, R.D. (2003). Sports economics. Upper Saddle River, NJ: Prentice Hall.

Fort, R.D. (2004). Inelastic sports pricing. *Managerial and Decision Economics*, 25, 87-94.

Howard, D. R., & Crompton, J.L. (2004). Financing Sport, 2nd ed., Morgantown, WV: Fitness Information Technology.

Krautmann, A., & Hadley, L. (2006). The product market for professional sports. In J. Fizel (Ed.) *Handbook of Sports Economics Research* (pp. 175-189). Armonk, NY: M.E. Sharpe, Inc.

Nagel, M.S., & Rascher, D.A. (2007). Washington “Redskins” – disparaging term or valuable tradition?: Legal and economic issues concerning Harjo v. Pro-football, Inc. *Fordham Intellectual Property, Media & Entertainment Law Journal*, 17, 789-804.

Noll, R. G. (1974). Attendance and price setting. In R.G. Noll (Ed.), *Government and the Sport Business* (pp. 115-157). Washington DC: Brookings Institution Press.

Noll, R.G., & Zimbalist, A.S. (1997). Build the stadium – create the jobs! In R.G. Noll & A.S. Zimbalist (Eds.), *Sports, Jobs & Taxes* (pp. 1-53). Washington DC: Brookings Institution Press.

NFLPA. (2006). Collective bargaining agreement between the NFL management council and the NFL players association. Retrieved October 12, 2009, from www.nflpa.org/CBA/CBA_Complete.aspx

Ostfield, A. (1995). Seat license revenue in the National Football League: Shareable or not? *Seton Hall Journal of Sport Law*, 5, 599-610.

Putsis Jr., W.P., & Sen, S.K. (2000). Should NFL blackouts be banned? *Applied Economics*, 32(12), 1495-1507.

Reese, J.T., Nagel, M.S., & Southall, R.M. (2004). National Football League ticket transfer policies: Legal and policy issues. *Journal of Legal Aspects of Sport*, 14(2), 163-190.

Sandomir, R. (2009, July 17). A Texas-size stadium. *New York Times.* Retrieved October 12, 2010, from http://www.nytimes.com/2009/07/17/sports/football/17cowboys.html

Schofield, J.A. (1983). Performance and attendance at professional team sports. *Journal of Sport Behavior*, 6, 196-206.

Sherman, L. (2007, September 7). Toughest NFL waiting lists. *Forbes.com*. Retrieved October 5, 2009, from http://www.forbes.com/2007/09/07/nfl-football-tickets-forbeslife-cx_ls_0907tickets.html

Tainsky, S. (2010). Television broadcast demand for National Football League contests. *Journal of Sports Economics*, 11(6), 629-640.

Team Marketing Report. (2009). NFL fan cost index. *Team Marketing Report*. Retrieved February 3, 2010, from http://www.teammarketing.com/blog/index.html?article_id=96.

Vrooman, J. (2010). *Dallas Cowboys Stadium Analysis*. Vanderbilt University, Nashville, TN.

Welki, A. M., & Zlatoper, T.J. (1994). US professional football: The demand for game-day attendance. *Managerial and Decision Economics*, 15(5), 489-495.

Welki, A.M. & Zlatoper, T.J. (1999). US professional football and game day attendance. *Atlantic Economic Journal*, 27, 285-298.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4), 817-838.

Figure 2.1: Secondary Market Supply and Demand of PSLs and STRs

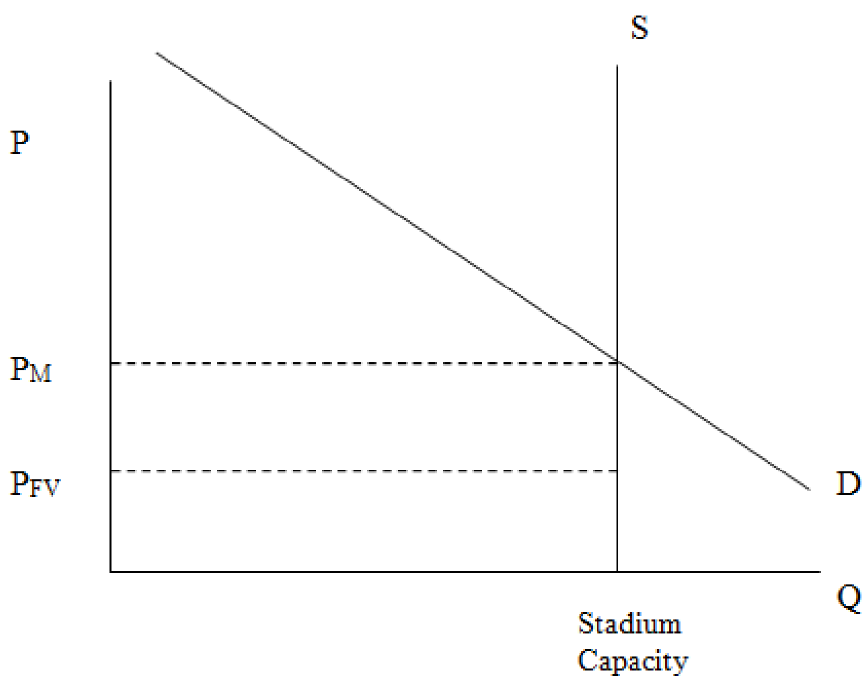


Figure 2.2: Average Consumer Cost of Purchasing a PSL and Season Ticket over Ten Years

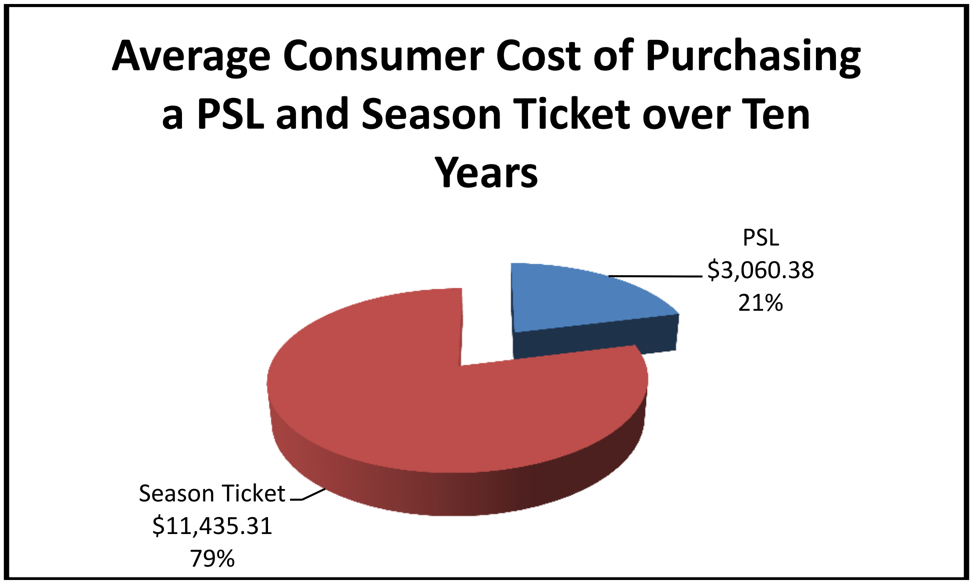


Figure 2.3: Average Consumer Cost of Purchasing a STR and Season Ticket over Ten Years

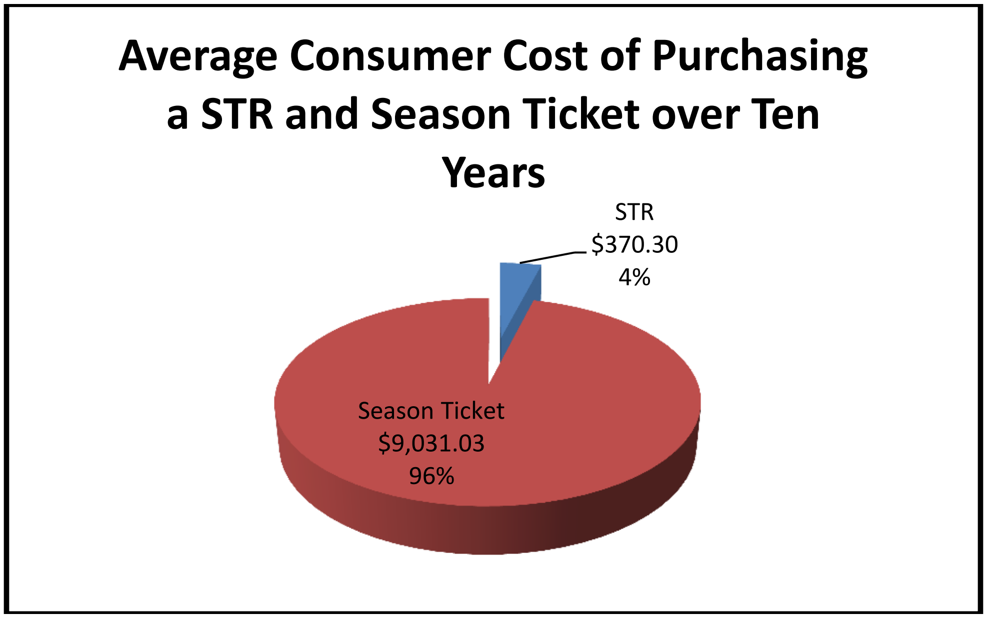


Figure 2.4: Secondary Market PSL Sale Prices and 3-Year Team Winning Percentage

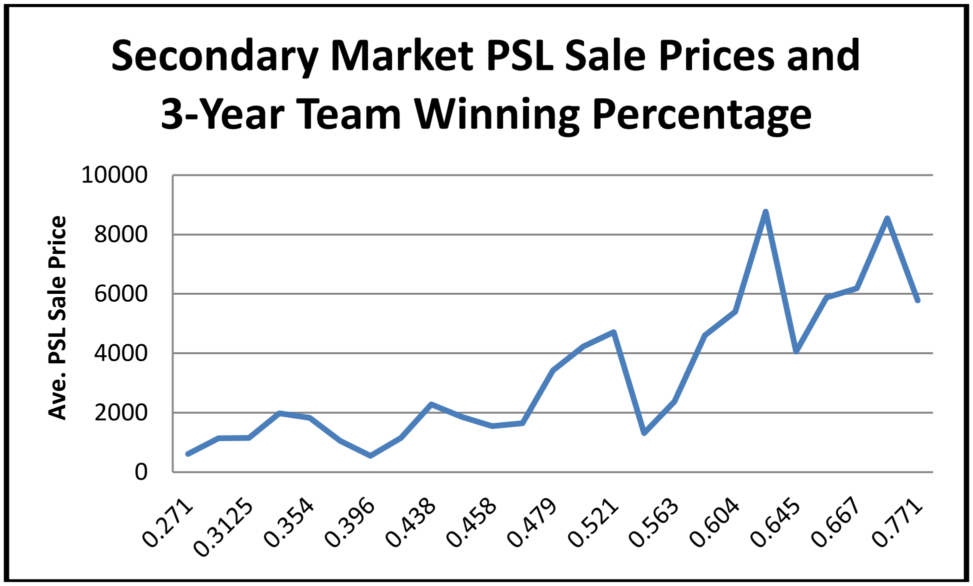


Figure 2.5: Secondary Market STR Sale Prices and 3-Year Team Winning Percentage

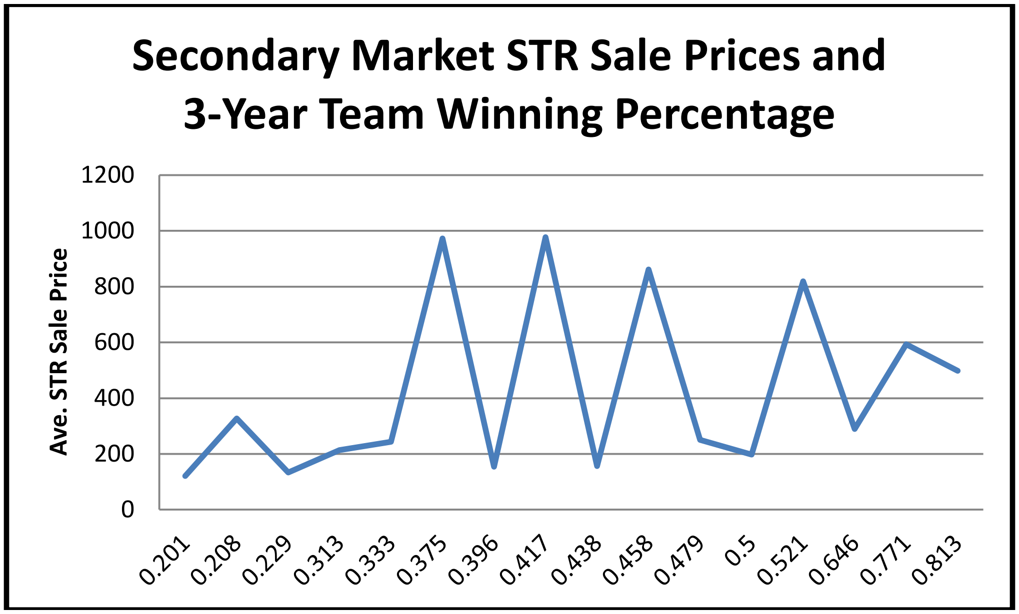


Figure 2.6: Secondary Market Sale Prices by Team and Year – Seat Quality 1

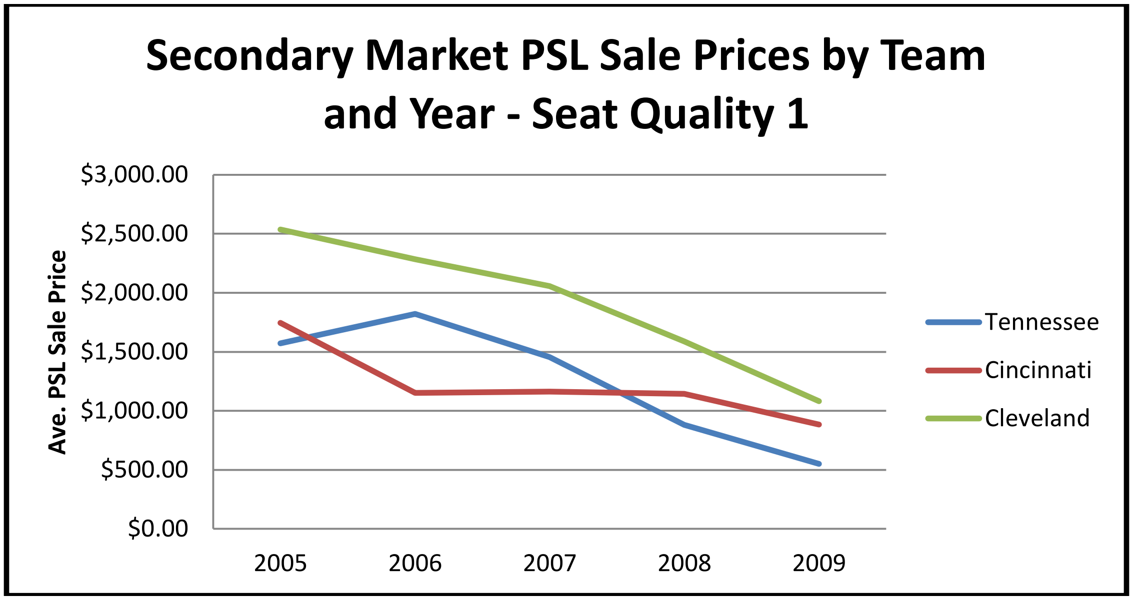


Figure 2.7: Secondary Market Sale Prices by Team and Year – Seat Quality 1

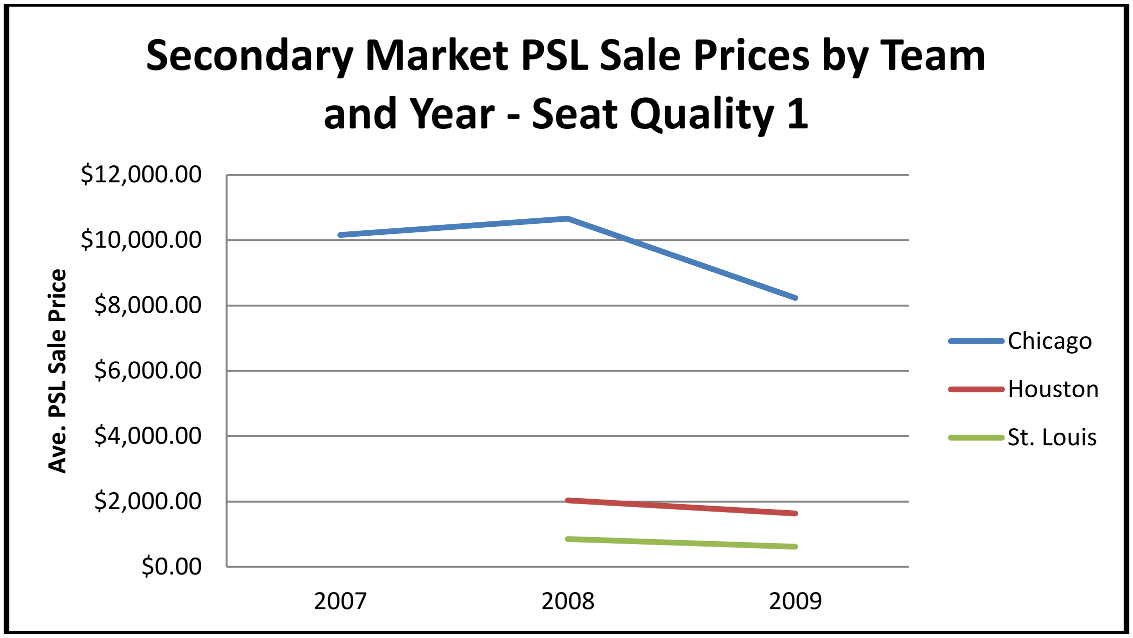


Figure 2.8: Secondary Market Sale Prices by Team and Year – Seat Quality 3

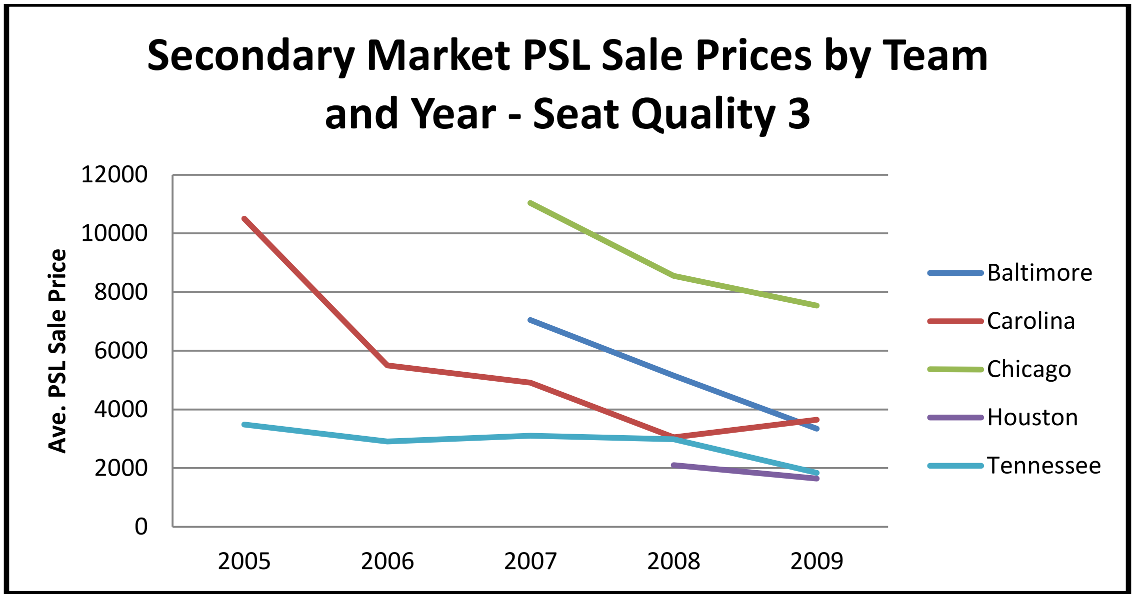


Table 2.1: Timeline of NFL PSL Programs

Year Venue Opened	Team	Name of Program
1995	St. Louis Rams	Personal Seat License
1996	Carolina Panthers	Personal Seat License
1999	Tennessee Titans	Personal Seat License
1998	Baltimore Ravens	Personal Seat License
1999	Cleveland Browns	Personal Seat License
2001	Pittsburgh Steelers	Seat License
2002	Houston Texans	Personal Seat License
2002	Seattle Seahawks	Charter Seat License
2003	Chicago Bears	Personal Seat License
2003	Green Bay Packers	Personal Seat License
2003	Philadelphia Eagles	Stadium Builder License
2004	Cincinnati Bengals	Charter Ownership Agreement
2009	Dallas Cowboys	Seat License
2010	New York Giants	Personal Seat License
2010	New York Jets	Personal Seat License

Table 2.2: Variable Descriptions

Variable	Description
LOGSEATPRICE	logged per seat sale price of a PSL or STR (dependent variable)
YEAR05	indicator variable; 1 = PSL or STR sold in 2005, 0 = not sold in 2005 (baseline category)
YEAR06	indicator variable; 1 = PSL or STR sold in 2006, 0 = not sold in 2006
YEAR07	indicator variable; 1 = PSL or STR sold in 2007, 0 = not sold in 2007
YEAR08	indicator variable; 1 = PSL or STR sold in 2008, 0 = not sold in 2008
YEAR09	indicator variable; 1 = PSL or STR sold in 2009, 0 = not sold in 2009
LOGROW	logged row of seating location associated with PSL or STR sale
SEATQUAL1	seat location: lower level between 30-yard lines
SEATQUAL2	seat location: upper level between 30-yard lines
SEATQUAL3	seat location: lower level between 30-yard lines and rear end zone lines
SEATQUAL4	seat location: upper level between 30-yard lines and rear end zone lines
SEATQUAL5	seat location: lower or upper level end zone seating (baseline category)
LOGSTADIUMAGE	logged age of team's stadium during year of PSL or STR sale
LOGSTADIUMCAPACITY	logged capacity of team's stadium during year of PSL or STR sale
TYPEPSL	indicator variable; 1 = PSL sale, 0 = STR sale
LOGTICKETPRICE	logged single game ticket price for the seating location associated with the PSL or STR sale
AISLE	indicator variable; 1 = seating location is on aisle, 0 = seating location is not on aisle
WIN3	cumulative team win percentage over the three seasons prior to the PSL or STR sale
LOCALUNEMPLOYMENT	local MSA unemployment rate during the month of the PSL or STR sale
LOGLIST	logged length of team's season ticket waiting list
NOLIST	indicator variable; 1 = team has no season ticket waiting list, 0 = team has waiting list
LOGPOPULATION	logged MSA population during the year of the year of the PSL or STR sale
LOGINCOME	logged MSA per capita income during the year of the PSL or STR sale
DOME	indicator variable; 1 = PSL or STR is located in a domed or retractable roof venue, 0 = otherwise
TEAM CITY/STATE	team indicator variables (included only in fixed effects models)

Table 2.3: Summary Statistics

Variable	N	Mean	SD	Minimum	Maximum
LOGSEATPRICE	3821	7.134	1.518	0	10.742
YEAR05	3821	0.026	0.158	0	1
YEAR06	3821	0.061	0.239	0	1
YEAR07	3821	0.121	0.326	0	1
YEAR08	3821	0.287	0.453	0	1
YEAR09	3821	0.505	0.500	0	1
LOGROW	3821	2.562	0.875	0	4.159
SEATQUAL1	3821	0.109	0.312	0	1
SEATQUAL2	3821	0.080	0.271	0	1
SEATQUAL3	3821	0.305	0.460	0	1
SEATQUAL4	3821	0.195	0.396	0	1
SEATQUAL5	3821	0.311	0.463	0	1
LOGSTADIUMAGE	3821	2.170	0.570	0	3.912
LOGSTADIUMCAPACITY	3821	11.144	0.058	11.027	11.426
TYPEPSL	3821	0.921	0.269	0	1
LOGTICKETPRICE	3821	4.426	0.523	2.890	6.310
AISLE	3821	0.027	0.163	0	1
WIN3	3821	0.485	0.119	0.201	0.813
LOCALUNEMPLOYMENT	3821	6.732	2.045	3.4	17
LOGLIST	3821	4.486	4.622	0	12.206
NOLIST	3821	0.507	0.500	0	1
LOGPOPULATION	3821	14.911	0.580	13.932	16.764
LOGINCOME	3821	10.668	0.114	10.472	11.031
DOME	3821	0.179	0.384	0	1

Table 2.4: Random and Fixed Effects Models (Pooled PSL and STR Observations)

	Random Effects			Fixed Effects		
	Coefficient	Robust S.E.	t-statistic	Coefficient	Robust S.E.	t-statistic
YEAR06	-0.054	0.102	-0.53	-0.117	0.146	-0.80
YEAR07	-0.047	0.097	-0.48	-0.059	0.218	-0.27
YEAR08	-0.325***	0.096	-3.39	-0.309	0.272	-1.14
YEAR09	-0.819***	0.111	-7.38	-0.339	0.303	-1.12
LOGROW	-0.234***	0.017	-13.95	-0.224***	0.015	-14.84
SEATQUAL1	1.069***	0.056	19.21	1.204***	0.051	23.55
SEATQUAL2	-0.150***	0.055	-2.71	-0.133***	0.050	-2.67
SEATQUAL3	0.518***	0.040	12.87	0.578***	0.037	15.74
SEATQUAL4	-0.718***	0.042	-16.90	-0.709***	0.039	-18.37
LOGSTADIUMAGE	-0.012	0.036	-0.34	-1.738***	0.229	-7.58
LOGSTADIUMCAPACITY	4.420***	0.303	14.59	x	x	x
TYPEPSL	2.112***	0.066	32.13	x	x	x
LOGTICKETPRICE	-0.254***	0.038	-6.68	-0.314***	0.035	-9.01
AISLE	-0.129	0.085	-1.51	-0.144*	0.077	-1.87
WIN3	3.162***	0.176	17.96	1.189***	0.219	5.42
LOCALUNEMPLOYMENT	-0.004	0.015	-0.24	-0.062***	0.018	-3.36
LOGLIST	-0.057***	0.021	-2.70	x	x	x
NOLIST	-1.369***	0.206	-6.65	x	x	x
LOGPOPULATION	0.645***	0.037	17.32	-0.169	0.127	-1.33
LOGINCOME	0.143	0.201	0.71	8.55***	1.987	4.30
DOME	-0.246***	0.052	-4.72	x	x	x
CONSTANT	-53.588***	4.014	-13.35	-75.346***	21.450	-3.51
PSL TEAMS						
BALTIMORE	x	x	x	baseline	baseline	baseline
CAROLINA	x	x	x	1.426***	0.365	3.91
CHICAGO	x	x	x	0.089	0.224	0.40
CINCINNATI	x	x	x	-1.537***	0.441	-3.48
CLEVELAND	x	x	x	0.154	0.354	0.43
DALLAS	x	x	x	-3.067***	0.615	-4.98
HOUSTON	x	x	x	-1.724***	0.162	-10.62
PHILADELPHIA	x	x	x	-0.809***	0.193	-4.18
PITTSBURGH	x	x	x	1.211***	0.237	5.11
SEATTLE	x	x	x	-2.047***	0.196	-10.42
ST. LOUIS	x	x	x	-0.528*	0.290	-1.82
TENNESSEE	x	x	x	0.261	0.368	0.71
STR TEAMS						
BUFFALO	x	x	x	0.692	0.611	1.13
DETROIT	x	x	x	-2.537***	0.407	-6.24
INDIANAPOLIS	x	x	x	-4.365***	0.616	-7.09
JACKSONVILLE	x	x	x	-2.171***	0.437	-4.97
KANSAS CITY	x	x	x	0.541	0.438	1.23
MIAMI	x	x	x	-0.248	0.303	-0.82
NEW ORLEANS	x	x	x	0.829**	0.366	2.26
OAKLAND	x	x	x	-5.72***	0.554	-10.32
SAN FRANCISCO	x	x	x	-2.990***	0.636	-4.70
TAMPA BAY	x	x	x	-2.189***	0.653	-3.35
WASHINGTON	x	x	x	-3.910***	0.480	-8.15
N = 3821	*** $p < .01$			N = 3821	*** $p < .01$	
$R^2 = .6936$	** $p < .05$			$R^2 = .7561$	** $p < .05$	
	* $p < .10$				* $p < .10$	

Table 2.5: Random and Fixed Effects Models (PSL Observations Only)

	Random Effects			Fixed Effects		
	Coefficient	Robust S.E.	t-statistic	Coefficient	Robust S.E.	t-statistic
YEAR06	-0.114	0.096	-1.18	-0.145	0.157	-0.92
YEAR07	-0.054	0.094	-0.57	-0.152	0.258	-0.59
YEAR08	-0.298***	0.097	-3.06	-0.427	0.327	-1.31
YEAR09	-0.533***	0.121	-4.39	-0.403	0.357	-1.13
LOGROW	-0.212***	0.016	-12.87	-0.208***	0.015	-14.06
SEATQUAL1	1.107***	0.054	20.57	1.231***	0.049	24.88
SEATQUAL2	-0.143***	0.053	-2.68	-0.149***	0.048	-3.10
SEATQUAL3	0.499***	0.040	12.45	0.592***	0.037	16.2
SEATQUAL4	-0.706***	0.041	-17.35	-0.706***	0.037	-19.27
LOGSTADIUMAGE	0.274***	0.048	5.70	-2.027***	0.234	-8.68
LOGSTADIUMCAPACITY	5.559***	0.336	16.57	x	x	x
TYPEPSL	x	x	x	x	x	x
LOGTICKETPRICE	-0.233***	0.037	-6.30	-0.315***	0.034	-9.30
AISLE	-0.122	0.092	-1.34	-0.171**	0.082	-2.08
WIN3	3.406***	0.184	18.51	1.102***	0.208	5.29
LOCALUNEMPLOYMENT	-0.095***	0.018	-5.17	-0.070***	0.018	-3.92
LOGLIST	0.070***	0.024	2.94	x	x	x
NOLIST	-0.022	0.247	-0.09	x	x	x
LOGPOPULATION	1.073***	0.059	18.10	1.887	1.467	1.29
LOGINCOME	-0.474	0.424	-1.12	10.591***	2.424	4.37
DOME	-0.621***	0.066	-9.46	x	x	x
CONSTANT	-65.536***	4.698	-13.95	-126.887***	43.122	-2.94
PSL TEAMS						
BALTIMORE	x	x	x	baseline	baseline	baseline
CAROLINA	x	x	x	2.833***	1.027	2.76
CHICAGO	x	x	x	-2.592	1.806	-1.43
CINCINNATI	x	x	x	-0.919	0.749	-1.23
CLEVELAND	x	x	x	0.969	0.702	1.38
DALLAS	x	x	x	-5.343***	1.266	-4.22
HOUSTON	x	x	x	-3.510***	1.179	-2.98
PHILADELPHIA	x	x	x	-2.585**	1.151	-2.25
PITTSBURGH	x	x	x	1.605***	0.417	3.85
SEATTLE	x	x	x	-2.757***	0.460	-6.00
ST. LOUIS	x	x	x	-0.317	0.304	-1.05
TENNESSEE	x	x	x	1.721	1.131	1.52
STR TEAMS						
BUFFALO	x	x	x	x	x	x
DETROIT	x	x	x	x	x	x
INDIANAPOLIS	x	x	x	x	x	x
JACKSONVILLE	x	x	x	x	x	x
KANSAS CITY	x	x	x	x	x	x
MIAMI	x	x	x	x	x	x
NEW ORLEANS	x	x	x	x	x	x
OAKLAND	x	x	x	x	x	x
SAN FRANCISCO	x	x	x	x	x	x
TAMPA BAY	x	x	x	x	x	x
WASHINGTON	x	x	x	x	x	x
N = 3520	*** p < .01			N = 3520	*** p < .01	
R^2 = .6706	** p < .05			R^2 = .7356	** p < .05	
	* p < .10				* p < .10	

Table 2.6: Random and Fixed Effects Models (STR Observations Only)

	Random Effects			Fixed Effects		
	Coefficient	Robust S.E.	t-statistic	Coefficient	Robust S.E.	t-statistic
YEAR06	x	x	x	x	x	x
YEAR07	x	x	x	x	x	x
YEAR08	x	x	x	x	x	x
YEAR09	-0.487	0.347	-1.4	-0.485	0.389	-1.25
LOGROW	-0.386***	0.078	-4.93	-0.373***	0.080	-4.67
SEATQUAL1	0.926***	0.334	2.77	0.898***	0.338	2.66
SEATQUAL2	0.187	0.322	0.58	0.168	0.339	0.5
SEATQUAL3	0.484	0.170	2.85	0.478***	0.171	2.8
SEATQUAL4	-0.652*	0.362	-1.8	-0.643	0.437	-1.47
LOGSTADIUMAGE	0.088	0.168	0.52	-0.375	1.014	-0.37
LOGSTADIUMCAPACITY	6.132***	1.991	3.08	x	x	x
TYPEPSL	x	x	x	x	x	x
LOGTICKETPRICE	-0.263	0.214	-1.23	-0.243	0.224	-1.09
AISLE	-0.077	0.249	-0.31	-0.079	0.251	-0.31
WIN3	-1.061	1.127	-0.94	0.276	2.803	0.1
LOCALUNEMPLOYMENT	-0.101	0.125	-0.81	-0.042	0.144	-0.29
LOGLIST	-1.052***	0.357	-2.95	x	x	x
NOLIST	-12.179***	4.308	-2.83	x	x	x
LOGPOPULATION	0.195	0.197	0.99	-0.004	0.253	-0.02
LOGINCOME	-0.284	0.760	-0.37	x	x	x
DOME	0.347	0.734	0.47	x	x	x
CONSTANT	-48.419*	26.431	-1.83	8.339	5.475	1.52
<u>PSL TEAMS</u>						
BALTIMORE	x	x	x	x	x	x
CAROLINA	x	x	x	x	x	x
CHICAGO	x	x	x	x	x	x
CINCINNATI	x	x	x	x	x	x
CLEVELAND	x	x	x	x	x	x
DALLAS	x	x	x	x	x	x
HOUSTON	x	x	x	x	x	x
PHILADELPHIA	x	x	x	x	x	x
PITTSBURGH	x	x	x	x	x	x
SEATTLE	x	x	x	x	x	x
ST. LOUIS	x	x	x	x	x	x
TENNESSEE	x	x	x	x	x	x
<u>STR TEAMS</u>						
BUFFALO	x	x	x	baseline	baseline	baseline
DETROIT	x	x	x	-0.843	1.981	-0.43
INDIANAPOLIS	x	x	x	0.060	3.035	0.02
JACKSONVILLE	x	x	x	-1.002	0.902	-1.11
KANSAS CITY	x	x	x	0.469	0.565	0.83
MIAMI	x	x	x	0.759	0.821	0.92
NEW ORLEANS	x	x	x	1.667**	0.697	2.39
OAKLAND	x	x	x	-0.939	1.401	-0.67
SAN FRANCISCO	x	x	x	0.077	0.582	0.13
TAMPA BAY	x	x	x	-1.389	1.288	-1.08
WASHINGTON	x	x	x	0.419	1.253	0.33
N = 301	*** p < .01			N = 301	*** p < .01	
R^2 = .4339	** p < .05			R^2 = .4399	** p < .05	
	* p < .10				* p < .10	

Table 2.7: Average Secondary Market PSL and STR Sale Prices (‘05-‘09)

Seat Quality	Average PSL or STR Sale Price	Standard Deviation	# of Observations
1	$6,504.91	$7,481.97	417
2	$1,642.45	$1,758.83	306
3	$3,644.72	$3,793.73	1165
4	$1,171.57	$1,433.80	746
5	$2,147.23	$2,109.23	1187
All Sales	$2,848.47	$3,870.69	3821

Note: Seat quality identifiers are those specified in Table 2

Table 2.8: Average Secondary Market PSL Sale Prices (‘05-‘09)

Seat Quality	Average PSL Sale Price	Standard Deviation	# of Observations
1	$6,898.46	$7,576.33	390
2	$1,724.55	$1,779.66	288
3	$3,973.96	$3,832.51	1055
4	$1,192.15	$1,442.36	730
5	$2,386.63	$2,112.02	1057
All Sales	$3,060.38	$3,957.48	3520

Note: Seat quality identifiers are those specified in Table 2

Table 2.9: Average Secondary Market STR Sale Prices ('05-'09)

Seat Quality	Average STR Sale Price	Standard Deviation	# of Observations
1	$850.91	$1,004.01	26
2	$312.94	$299.49	19
3	$487.06	$744.11	110
4	$232.62	$189.81	16
5	$200.70	$314.37	130
All Sales	$370.30	$611.20	301

Note: Seat quality identifiers are those specified in Table 2

Table 2.10: Secondary Market PSL Sale Prices by Franchise (‘05-‘09)

Team	Average PSL Sale Price	Standard Deviation	# of Observations
Baltimore	$3,818.16	$2,588.58	766
Carolina	$2,729.48	$2,391.89	93
Chicago	$9,428.45	$7,287.37	305
Cincinnati	$815.18	$1,035.12	529
Cleveland	$1,081.25	$1,101.36	523
Dallas	$4,059.37	$1,948.99	71
Houston	$1,737.32	$1,428.38	369
Philadelphia	$5,563.14	$2,690.69	92
Pittsburgh	$9,065.07	$5,181.42	135
Seattle	$4,058.03	$1,760.15	55
St. Louis	$655.91	$714.26	137
Tennessee	$1,671.14	$1,649.54	445
All Sales	$3,060.38	$3,957.48	3520

Note: Includes all seat locations and quality of seat locations sold may vary by team

Table 2.11: Secondary Market PSL Sale Prices by Franchise - Seat Quality 1 (‘05-‘09)

Team	Average PSL Sale Price	Standard Deviation	# of Observations
Baltimore	$8,925.78	$4,102.39	69
Carolina	$6,152.69	$5,353.55	6
Chicago	$19,484.88	$10,478.50	50
Cincinnati	$2,084.46	$1,901.49	51
Cleveland	$3,222.36	$2,225.83	18
Dallas	-	-	0
Houston	$3,202.51	$2,207.66	71
Philadelphia	$9,250.07	$4,400.35	7
Pittsburgh	$18,885.42	$5,149.87	14
Seattle	$4,332.93	$1,921.06	28
St. Louis	$1,526.63	$836.87	33
Tennessee	$3,972.70	$2,734.32	43
All Sales	$6,898.46	$7,576.33	390

Note: Includes sales for Seat Quality 1 only; lower level between the 30-yard lines

Table 2.12: Secondary Market STR Sale Prices by Franchise ('05-'09)

Team	Average STR Sale Price	Standard Deviation	# of Observations
Buffalo	$151.02	$138.86	22
Detroit	$117.56	$155.67	29
Indianapolis	$550.73	$402.86	20
Jacksonville	$134.63	$212.74	13
Kansas City	$244.99	$245.06	22
Miami	$818.77	$1,321.43	15
New Orleans	$832.79	$816.24	59
Oakland	$153.01	$247.56	40
San Francisco	$167.82	$195.84	72
Tampa Bay	$15.42	$10.13	3
Washington	$1,151.13	$1,323.12	6
All Sales	$370.30	$611.20	301

Note: Includes all seat locations and quality of seat locations sold may vary by team

Table 2.13: Secondary Market STR Sale Prices by Franchise - Seat Quality 1 (‘05-‘09)

Team	Average STR Sale Price	Standard Deviation	# of Observations
Buffalo	-	-	0
Detroit	-	-	0
Indianapolis	-	-	0
Jacksonville	-	-	0
Kansas City	-	-	0
Miami	-	-	0
New Orleans	$1,021.50	$1,137.52	18
Oakland	$391.66	$444.04	7
San Francisco	$995.00	$0.00	1
Tampa Bay	-	-	0
Washington	-	-	0
All Sales	$850.91	$1,004.01	26

Note: Includes sales for Seat Quality 1 only; lower level between the 30-yard lines

Table 2.14: Example of Consumer PSL and Season Ticket Expenditure over Ten Years

Seat Quality	Average PSL Sale Price	Year 1 Average Ticket Price	PSL & Season Ticket Cost over 10 Years
1	$6,898.46	$165.65	$25,887.84
2	$1,724.55	$71.79	$9,953.91
3	$3,973.96	$136.27	$19,595.46
4	$1,192.15	$55.82	$7,591.54
5	$2,386.63	$76.95	$11,207.97
Average	$3,060.38	$99.75	$14,495.69

Note 1: Seat quality identifiers are those specified in Table 2
Note 2: Assumes PSL paid in full at time of purchase
Note 3: Assumes 3% yearly increase in face value ticket price
Note 4: Assumes PSL and season ticket prices from 2005-2009

Table 2.15: Example of Consumer STR and Season Ticket Expenditure over Ten Years

Seat Quality	Average STR Sale Price	Year 1 Average Ticket Price	STR & Season Ticket Cost over 10 Years
1	$820.34	158.13	$18,948.38
2	$328.91	50.17	$6,079.95
3	$487.06	84.57	$10,181.74
4	$232.62	62.00	$7,340.22
5	$200.70	63.43	$7,471.67
Average	$370.30	78.78	$9,401.32

Note 1: Seat quality identifiers are those specified in Table 2
Note 2: Assumes STR paid in full at time of purchase
Note 3: Assumes 3% yearly increase in face value ticket price
Note 4: Assumes STR and season ticket prices from 2005-2009

Table 2.16: Population and Income Data for PSL Markets (‘05-‘09)

Team	Average Population	Average Income
Baltimore	2675093	$47,008
Carolina	1661729	$38,792
Chicago	9570747	$45,270
Cincinnati	2160335	$38,424
Cleveland	2095028	$38,969
Dallas	6418509	$42,392
Houston	5832369	$48,259
Philadelphia	5868017	$45,702
Pittsburgh	2357703	$41,530
Seattle	3342002	$49,376
St. Louis	2826928	$41,323
Tennessee	1540685	$38,855
PSL Average	3862429	$42,992

Note: Average population and income figures over the examination period

Table 2.17: Population and Income Data for STR Markets (‘05-‘09)

Team	Average Population	Average Income	Adjusted Population
Buffalo	1123850	$36,408	1123850
Detroit	4407921	$39,806	4407921
Indianapolis	1730969	$39,318	1730969
Jacksonville	1323555	$39,304	1323555
Kansas City	2064606	$40,367	2064606
Miami	5511763	$43,709	5511763
New Orleans	1134029	$44,136	1134029
Oakland	4313521	$61,747	2156760
San Francisco	4313521	$61,747	2156760
Tampa Bay	2742768	$36,918	2742768
Washington	5436871	$56,510	5436871
STR Average	3100307	$45,452	2708168

Note: Average population and income figures over the examination period

Table 2.18: Secondary Market PSL Sale Prices and 3-Year Team Winning Percentage

3-Year Win %	Average PSL Sale Price	Standard Deviation
0.271	616.80	710.46
0.292	1140.65	987.71
0.313	1151.30	1153.74
0.333	1985.40	1707.84
0.354	1832.14	1781.06
0.375	1051.57	1036.34
0.396	549.51	907.94
0.417	1155.80	1271.68
0.438	2283.62	1900.20
0.446	1866.80	1886.01
0.458	1549.31	1503.39
0.459	1645.10	1301.01
0.479	3418.56	1471.32
0.500	4225.10	2745.78
0.521	4717.50	4013.63
0.542	1316.41	1360.99
0.563	2376.38	2222.06
0.583	4607.98	2817.11
0.604	5397.49	5216.08
0.625	8766.53	4981.58
0.645	4059.37	1948.99
0.646	5876.35	7138.65
0.667	6185.55	3576.66
0.708	8547.11	4419.14
0.771	5779.69	1571.20
Average	3060.38	3957.48

Table 2.19: Secondary Market STR Sale Prices and 3-Year Team Winning Percentage

3-Year Win %	Average PSL Sale Price	Standard Deviation
0.201	121.26	170.16
0.208	326.81	355.22
0.229	133.70	231.60
0.313	213.33	231.58
0.333	243.14	184.63
0.375	973.30	1523.21
0.396	153.58	206.53
0.417	977.15	716.04
0.438	156.06	144.09
0.458	861.88	1355.33
0.479	250.00	0.00
0.500	197.86	494.77
0.521	819.42	829.61
0.646	289.38	338.80
0.771	593.74	407.63
0.813	498.17	414.85
Average	370.30	611.20

Table 2.20: Average PSL Sale Price by Team and Year - Seat Quality 1

Team	2005	2006	2007	2008	2009
Baltimore	-	-	$4,740.48	$4,083.29	$2,958.28
Carolina	$4,717.50	$4,269.14	$2,821.80	$1,924.48	$2,233.31
Chicago	-	-	$10,161.81	$10,658.95	$8,232.76
Cincinnati	$1,572.63	$1,819.74	$1,454.20	$882.28	$549.50
Cleveland	$1,745.02	$1,151.30	$1,161.47	$1,142.58	$884.29
Houston	-	-	-	$2,038.60	$1,635.80
Philadelphia	$5,779.69	$5,265.41	$7,268.18	$5,704.59	$4,854.07
Pittsburgh	$8,453.24	$8,789.05	$8,547.11	$10,026.33	$8,766.53
Seattle	$2,810.00	$4,316.50	$6,041.67	$4,278.62	$3,418.56
St. Louis	-	-	-	$849.78	$616.80
Tennessee	$2,536.14	$2,283.62	$2,055.85	$1,588.67	$1,081.94
Average	$3,058.38	$2,718.82	$3,580.60	$3,907.32	$2,440.48

Note: Includes only seat quality 1; lower level seating between 30-yard lines

Table 2.21: Average PSL Sale Price by Team and Year - Seat Quality 3

Team	2005	2006	2007	2008	2009
Baltimore	-	-	$7,053.14	$5,158.38	$3,340.79
Carolina	$10,500.00	$5,500.50	$4,916.67	$3,050.00	$3,650.00
Chicago	-	-	$11,040.18	$8,550.90	$7,537.28
Cincinnati	$1,670.00	$2,125.63	$3,304.17	$1,430.51	$820.57
Cleveland	$2,083.21	$2,313.95	$2,067.50	$2,239.94	$1,246.62
Houston	-	-	-	$2,100.36	$1,644.27
Philadelphia	$5,667.50	$5,224.16	$5,512.50	$6,063.31	$4,171.28
Pittsburgh	$8,502.43	$12,703.58	$11,500.00	$13,761.06	$12,229.50
Seattle	$2,500.00	-	$6,500.00	$3,903.36	$3,409.06
St. Louis	-	-	-	$572.90	$463.05
Tennessee	$3,487.50	$2,906.39	$3,102.29	$2,987.74	$1,836.74
Average	$3,567.41	$3,672.15	$5,493.53	$5,144.74	$3,062.89

Note: Includes only seat quality 3; lower level seating between 30-yard line and end zone

CHAPTER 3

Training and the Major League Baseball Draft

3.1 Introduction

Empirical research focused on the effect of training on employment outcomes has received a great deal of attention in the labor economics literature (for example, Mincer, 1962; Card & Sullivan, 1988; Gritz, 1993). However, the professional sports landscape in North America also allows for an opportunity to examine the relationship between training and employment outcomes. This research will examine labor market outcomes in North American professional baseball for players with varying training backgrounds. In Major League Baseball (MLB) the rights to amateur players are typically acquired by clubs via the June First-Year Player Draft. Selections are made directly from one of three sources; high school players, junior or senior players from four-year institutions, or players from junior colleges or community colleges (Winfree & Molitor, 2006).

More importantly, the fact that players are drafted by clubs at various stages in their development represents variation in both the type and quality of training acquired prior to assignment into the minor league system of the MLB parent club. With the top clubs spending over $10 million each year just on draftee signing bonuses (Castrovince, 2010), it is of practical significance to examine whether relationships exist between training and employment outcomes in the professional baseball labor market.

Certain clubs utilize drafting strategies focused on acquiring players from specific training backgrounds. Previous accounts suggest that some clubs prefer to draft players from four-year institutions with the thought that they are more polished and have a shorter turnaround time for a potential contribution at the big league level (Lewis, 2003). The opposing line of thought supports the strategy of drafting younger high school players who will be subjected to an extended period of instruction provided and monitored by the parent club. Naturally, there are also proponents supporting the selection of players from both backgrounds. The training groups will be classified as follows:

1 – Players drafted directly out of high school
2 – Players drafted directly out of a four-year institution
3 – Players drafted directly out of a junior college or community college

There are two primary employment outcomes of interest in this study. The first is determining the rate at which individuals drafted in the MLB June First-Year Player Draft reach the major league level from each of the three training categories. The second is determining the Major League career duration of players drafted in the MLB June First-Year Player Draft from each of the three training groups.

On the firm side, this study will document the behavior of the collective market through an examination of historical selection decisions in regards to drafting players. In order to determine how the market has operated and whether or not adjustments have occurred over time, documentation of its behavior is necessary. This will shed light on whether there has been a propensity for the market to prefer players from specific training groups. The results should allow for an assessment of whether selection into the labor market matches with employment outcomes.

This study is warranted not only based on the objectives outlined above, but because of the previous empirical work examining the professional baseball players' labor market. Past studies have focused on the ability to locate talent and the rates of return to clubs in the MLB Draft (Spurr, 2000; Burger & Walters, 2009). While these pieces have enhanced our understanding of the Draft, a potential shortcoming is their decision to use very limited data samples – both in number of years and number of draft rounds utilized. Through the use of a more comprehensive data set, this study will be able to provide more robust information to the parent discipline in economics regarding the effect of training differences on employment outcomes in a specialized labor market.

Other work in the area has addressed potential market inefficiencies associated with the Draft. Specifically, these contributions have highlighted the debate between selecting high school versus college talent. To date, numerous popular media outlets, novelists (Lewis, 2003), and academicians (Bradbury, 2011) have addressed, but only loosely analyzed the issue. These debates suggest that a detailed analysis regarding this topic may be a missing link in the sports labor literature.

Subsequently, this chapter contributes to the literature in the field of sports economics by thoroughly examining the relationship between training and employment outcomes in this specialized labor market. Through the use of historical MLB Draft data, the analysis examines historical selection into the labor market and labor market outcomes as measured by probability of reaching MLB and MLB career duration. Specific to selection into the market, the results highlight a period of market adjustment in the mid-1970s which is illustrated by a drastic increase in the selection of four-year college players. This shift in selection into the labor market has been maintained ever since.

In terms of labor market outcomes, players drafted out of four-year institutions have shown significantly higher probabilities of reaching MLB over the history of the Draft as illustrated by five separate logistic regression models. This result can likely be contributed to the fact that college players have larger quantities of accumulated training when drafted. Accordingly, these players are more developed and competed against a higher level of competition, which has allowed MLB clubs to make more accurate predictions regarding the probability of reaching the majors. This result supports economic theory highlighting the positive relationship between accumulated training and the value of the worker to the firm as measured by the probability of employment success.

Four Cox proportional hazard models are used the estimate major league career duration of drafted players and the results oppose traditional economic theory. In all four hazard models, high school players are significantly more likely to remain employed in the MLB labor market at all points in time as compared to four-year college players. This finding suggests that while college players may be less risky draft selections due to the fact that they are much more likely to reach MLB, high school players appear to have a higher payoff to clubs as they remain in the labor market for a longer period of time.

The chapter proceeds with a section outlining the MLB Draft, a theoretical treatment of the relationship between training and labor market outcomes, and a literature review. Following these introductory sections, separate sections outlining the data utilized, econometric methods and estimation results will be presented for each of the two labor market outcomes identified. The chapter concludes with discussion and conclusion sections.

3.2 The MLB Draft

The MLB Amateur Draft was first established in 1965 and was renamed the MLB First-Year Player Draft in 1998 (Rausch, 2002). In its original conception, clubs selected amateur players in reverse-order-of-finish from the previous season's standings with teams from the American and National Leagues alternating picks (Koppett, 1965). The 1966 version of the Amateur Draft was the first in which teams selected in a true reverse-order fashion, and this formatting alteration has remained since. The reverse-order-of-finish mechanism was promoted by franchise owners as a competitive balance mechanism needed to provide poor performing franchises the opportunity to acquire the most talented amateur players. Despite several economists (for example, Fort and Quirk, 1995) illustrating how league institutional configuration is arranged so that large revenue clubs can simply purchase playing talent from lower revenue generating clubs, the reverse-order of finish draft is still operational.

It is also appropriate to distinguish between the June First-Year Player Draft and other smaller versions of the MLB Draft. Over time, MLB has held a January Regular Draft, a January Secondary Draft and a June Secondary Draft. Each of these alternative versions of the June First-Year Player Draft existed from 1966 to 1986. The American Legion Draft also existed until 1966 (Madden, 2001). These alternative drafts were much smaller than the traditional Draft as they consisted of only a limited number of rounds and selections. Alternatively, the First-Year Player Draft has served as the primary mechanism in which MLB clubs have selected amateur talent into the North American professional baseball labor market. In order to create a data set where players were selected into the professional baseball labor market under the same set of rules, only player data from the June First-Year Player Draft will be used to analyze the objectives identified in this chapter. Moving forward, the MLB June First-Year Player Draft will be referred to simply as the “MLB Draft” or the “Draft.”

3.3 Theory

This chapter utilizes theory on human capital and employment training and applies it to the unique process of selection into the labor market found in the MLB Draft. Research on human capital is concerned with determining how abilities and skills are accumulated in people and establishing the value of these collected assets. The majority of human capital is acquired through education and informal training (Becker, 1961). In fact, previous empirical research estimates that training accounts for at least half of a worker’s human capital (Mincer, 1962). Following in line with Mincer’s (1962) definition, we will assume that “training” refers to investment in the acquisition of skill or improvement in worker productivity. The author also claims that when training is identified as a process of human capital formation in workers, key empirical questions are appropriate for economic analysis. Of specific importance is 1) determining the rate of return of this training and 2) assessing how training informs on features of labor force behavior (Mincer, 1962).

Theory regarding general training specifies that the value of marginal product (VMP) for a worker is dependent on the amount of training acquired by that worker when holding other variables such as innate ability constant (Borjas, 2005). Therefore, in a case where there are two workers, with worker A possessing a larger amount of general training than worker B, represented by $T_A > T_B$, when holding all else constant, $VMP_A > VMP_B$ and $W_A > W_B$, where W represents the wage paid to the worker.

The MLB Draft is an appropriate mechanism to evaluate this relationship as players drafted by MLB clubs are not only selected from different backgrounds but they are also drafted at different chronological ages. In the MLB Draft, three different "treatments" of players are clearly identifiable. For classification purposes, group A will signify players drafted directly out of high school, group B will signify players drafted out of four-year institutions, and group C will signify players drafted out of a junior or community college. At the time of the draft, players in group A have accumulated four years of general training, players in group B have accumulated either seven or eight years of general training, and players in group C have accumulated either five or six years of general training[15]. This scenario represents differences in accumulated human capital based on variation in the amount of general training accumulated by players at the time they are drafted. This training is considered to be general training, because the skills acquired can be utilized with all MLB clubs. Clearly, variation exists in both the quantity and quality of training accrued at the time of the Draft. But since this training is not controlled by the drafting MLB club and virtually every player spends significant time in the minor leagues, which represents an additional period of training controlled by the drafting club, it is not clear that $T_B > T_C > T_A$ and $VMP_B > VMP_C > VMP_A$. It is important to note that this chapter is focused on the relationship between training and employment outcomes from the perspective of the firm at the time the firm must make

[15] MLB First-Year Player Draft rules specify that players from four-year institutions are eligible to be drafted only after completing their junior or senior seasons. Players drafted from junior or community colleges are eligible to be drafted following either their first or second seasons.

the decision on selecting the player. Therefore it is appropriate to evaluate this relationship at the time of selection into the labor market – namely, at the MLB Draft.

More specifically, the training to value production relationship linking the worker and the firm can be illustrated through theory outlined by Borjas (2005), but originally developed by Becker (1961 & 1962) and Mincer (1962). Specifically, assume that the employment relationship between the drafted player and the MLB club lasts two periods. Assuming the player reaches MLB, period 1 is the phase between the time the player is drafted and the time the player reaches MLB – also known as the player's minor league career. Period 2 is the phase between the time the player reaches MLB and the time the player exits the MLB labor market, either voluntarily or involuntarily. Total labor costs assumed by the MLB club in period 1 and period 2 are denoted by TC_1 and TC_2, respectively. The player's values of marginal product in each of the two periods are specified by VMP_1 and VMP_2 with *r* representing the discount rate. Directly from the work of Borjas (2005), the profit-maximizing condition for the optimal level of employment for the firm for the two periods is given by:

$$TC_1 + \frac{TC_2}{1+r} = VMP_1 + \frac{VMP_2}{1+r}$$

The left-hand side of the equation represents the present value of the firm's costs of drafting a player, training that player following their high school and/or college baseball training and promoting that player to MLB. The right-hand side of the equation represents the present value of the player's contribution to the MLB parent club in both the minor league and major league periods. In this equation, the wage paid to the player by the MLB club is equal to the player's value of marginal product.

However, in the current scenario there is no guarantee that the drafted player will reach period 2. Specifically, the player will never reach period 2 if the firm estimates that the value generated by the player in period 2 and to a lesser degree period 1 will not match or exceed the combined total labor costs in periods 1 and 2. The player will reach period 2 if the club estimates that the player's value will meet or exceed those combined total labor costs. Moreover, the length of a player's career duration in period 2 is also a function of the value created by the player in relationship to total labor costs incurred by the club. If the player's VMP is equal to or exceeds the wage that they are paid by the club, that player will continue to be employed in the MLB labor market. At the point where the player's production falls below either the league minimum salary or the lowest wage the player is willing to accept in exchange for playing, the player will exit the MLB labor market.

Following in line with Mincer's (1962) suggestions for empirical investigation outlined above, this study will evaluate the relationship between training and the value of the worker to the firm through the measurement of labor market outcomes of draftees. This will be achieved by 1) determining the rate at which players from the three training categories reach MLB and 2) measuring variation in MLB career duration, for players in each of the three training groups. More specifically, this work will examine these labor market outcomes through the use of econometric techniques. Specific to the first outcome, the probability of a drafted player reaching MLB is specified based on the following general function:

$$P = f\,(\textit{training}, \textit{draft}, \textit{demographic}),$$

where P is the probability that the player will reach MLB following selection in the MLB Draft, *training* is a vector of training specific variables, *draft* is a vector of draft specific variables and *demographic* is a vector of demographic variables specific to the selected player. It is a reasonable assumption that each of these three categories of variables influences the probability that a given player will reach MLB.

Explicit to the second labor market outcome, the MLB career duration of a drafted player, given that they have reached MLB is specified based on the following general function:

$$D = f\,(\textit{training}, \textit{draft}, \textit{demographic}),$$

where D is the MLB career duration of the specified player, *training* is a vector of training specific variables, *draft* is a vector of draft specific variables and *demographic* is a vector of demographic variables specific to the selected player. While the exact vector of variables differs slightly based on the labor market outcome specified, it is reasonable to assume that the general categories of variables used to estimate both labor market outcomes is equivalent. The exact methodology and empirical specifications for both labor market outcomes are described in detail in their respective "data and methods" sections below.

3.4 Literature Review

Empirical research focused on the effect of training on employment outcomes has been a mainstay in the labor economics literature. This line of investigation was born out of the original contributions exploring human capital formation (most notably, Becker, 1961 & 1962 and Schultz, 1961 & 1962). Mincer (1962) was the first to empirically explore the topic and provided a survey on the relationship between education, on-the-job training and returns to these investments. His work sparked further investigation on the relationship between training and employment outcomes.

As scholarship in the area expanded, Kaitz (1979), Ridder (1986) and Card and Sullivan (1988) all examined the effects of government subsidized training programs on employment. The work by Card and Sullivan (1988) is likely the most well-known contribution in this area. Following controversy regarding the effectiveness of government training programs, the authors focused on a simple employment outcome – the probability of gaining employment following training. Their contributions illustrated significant effects for both on-the-job training and classroom training programs. Not only were those included in training programs more likely to become employed, but they were also more likely to remain in the labor market, providing empirical support for theory regarding the relationship between training and employment outcomes.

Gritz (1993) filled a void in the literature by examining private training programs by analyzing the frequency and duration of employment spells. By utilizing continuous time duration modeling, the analysis uncovered positive effects for women, men and minorities following completion of a private training program. Though the effect for men was smaller than that of women, the results illustrated that training programs were associated with longer employment spells and shorter non-employment episodes.

In the context of professional sports, Rottenberg (1956) produced the first exposition on the unique nature of the labor market for professional baseball players. Since that time, empirical investigation on professional sports labor markets has been one of the most thoroughly investigated topics in the sports economics literature (Fort, 2006).

As noted by Kahn (2000), professional sports are fertile ground for the investigation of labor markets due to the immense availability of performance statistics. Naturally, baseball has been the most analyzed of the North American professional sports labor markets. Henceforth, this review of literature will be constrained to contributions focused on the Major League Baseball Draft, measurement of high school versus college players and those specific to the methodology used in this investigation.

Shughart and Goff (1992) were the first to produce empirical work comparing employment outcomes for high school and college trained minor league baseball players. The analysis uncovered that four-year college players spent less time in the minor leagues prior to being promoted to the major league level. Spurr and Barber (1994) followed up by examining promotion, demotion and turnover for minor league pitchers for the fourteen year period ranging from 1975 to 1988. This contribution was unique in that the analysis used performance data to estimate a worker's career path – an issue that previously was examined from solely a theoretical perspective. Probit modeling illustrated that players were both promoted and demoted quicker within the minor league system when performance deviated further away from the mean. Duration modeling also established that MLB clubs were able to make quicker determinations on promotion and demotion for college players as compared to high school players.

Spurr (2000) added to the literature by providing an analysis of the ability of MLB clubs to identify talent through the Draft. The author utilizes a probit model to estimate the probability of a drafted player reaching the major leagues and finds no statistically significant differences in the ability of clubs to locate talent. Instead, the factors influencing the probability of success were draft position and the background of the selected player. Namely, players from four-year institutions and those from elite college baseball programs had increased probabilities of reaching MLB. Spurr identifies that for a period of time, the market failed to identify the abilities of players from the collegiate ranks. The idea of this market inefficiency was later popularized in the mainstream through the commercial best-seller *Moneyball* (Lewis, 2003).

While Spurr's (2000) work is undoubtedly a central contribution to the literature, there are two key factors necessitating an expanded analysis of the Draft. First, the author utilizes draft data for four years only; 1966-1968 and 1983. This is an understandable choice based on the substantial amount of time associated with collecting, cleaning and organizing data of this nature. However, a short-run empirical examination of employment outcomes stemming from the Draft has the ability to produce spurious results. Specifically, the current analysis illustrates that over time, significant effects have the ability to both appear and disappear depending on the period of examination. Secondly, the primary goal of Spurr's (2000) paper was to estimate whether there is a difference in ability of clubs to locate talent through the draft mechanism. Alternatively, the current contribution seeks to analyze the relationship between training and employment outcomes through the utilization of historical MLB Draft data – a missing link in the sports labor literature.

Winfree and Molitor (2006) approach the Draft from a different perspective by evaluating the financial decision of high school athletes to turn professional or attend college based on draft position. The results confirm that players drafted in earlier rounds of the draft are better off entering professional baseball, while those selected in the twelfth round or later have higher estimated career earnings by attending college. The authors also find that players selected in the later rounds have a higher probability of reaching the major leagues if they attend college. This result hints at a potential relationship between the amount of training a player accumulates prior to being selected into the labor market and employment success.

Burger and Walters (2009) analyze the draft in a manner which estimates the financial returns to clubs based on the decision to select high school versus college players. Using 1990-1997 data over the first ten rounds of the draft, the authors find that clubs generate higher returns from college as opposed to high school selections. Additionally, findings point to the market overvaluing pitchers and overcompensating high school players in the early rounds of the Draft. Similar to that of Spurr (2000), the authors add to the growing support for the existence of inefficiencies in the market.

Most recently, in his book *Hot Stove Economics*, Bradbury (2011) also addresses the debate comparing the selection of high school versus college players. Bradbury acknowledges that the data support the notion that college players may be undervalued, but he cautions against claims of market inefficiency. He instead suggests that movement toward the selection of a larger number of college players is not warranted. Namely, he opines that high school players have greater upside and an increase in the selection of college players would simply increase the number of “duds” selected from that pool of talent.

Furthermore, and as it directly relates to the second labor market outcome of interest, the use of duration modeling has shown to be an effective tool in estimating labor market survival times in professional sports. Atkinson and Tschirhart (1986) were the first to use this approach in the context of professional sport in their examination of

National Football League (NFL) career length. Staw and Hoang (1995) examined the relationship between sunk costs and survival in the National Basketball Association (NBA) using duration modeling. Hoang and Rascher (1999) and Groothuis and Hill (2004) used the same techniques to investigate exit discrimination in the NBA. Frick, Pietzner, and Prinz (2007) examined the labor market for players in the German Bundesliga soccer league and discovered positional effects on career duration. Most recently, Volz (2009) used survival modeling to estimate managerial tenure in MLB. Despite the increased use of duration modeling in the general labor economics and sports economics literature, it does not appear that any published work examining the effect of training on employment outcomes has been completed in the context of professional sports. Subsequently, this chapter addresses this missing link in the context of the Major League Baseball.

3.5 Labor Market Outcome 1: Entry into Major League Baseball (Data/Methods)

Historical player data from the MLB Draft is used to examine the objectives outlined in this chapter. Because the two objectives regarding labor market outcomes are dissimilar, two separate samples of data are used for each. Hence, the data samples, empirical specifications and results for each outcome are presented separately in order to aid readability. This section will describe the data and empirical specification utilized to evaluate the first labor market outcome – probability of entry into Major League Baseball following selection in the MLB Draft.

The first objective aims to determine the rate at which drafted players from each of the three training groups reach Major League Baseball. The data were collected from the MLB Draft pages at www.baseball-reference.com. This sample includes players drafted in the MLB Draft from 1966 to 2005 and follows the playing career of players through the completion of the 2010 season. In total, 40,678 observations are included. In order to avoid a biased sample, players drafted in the 2005 draft class are the final year of draftees included in the sample. Since the average minor league career is 4.67 years, including players drafted in 2006 or later would be inappropriate, as a significant portion of those players would still make the major leagues following the 2010 season.

It is also important to point out the exclusion of certain observations. Specifically, the data does not include observations for players who were drafted but did not sign a professional contract and returned to school. Holding out this subsample of players is appropriate because even though a player may be drafted up to five times in the June First-Year Draft, they only are able to sign a professional contract and enter the professional baseball labor market once. In order to capture only the final time a player is drafted, previous player observations are withheld.

Logistic regression is the empirical modeling approach utilized to examine the probability of a player selected in the MLB Draft reaching the major league level of professional baseball. This scenario represents a binary outcome as there are only two possible results – either the player does or does not reach the major leagues during his career in professional baseball. Therefore, it is appropriate to model this behavior using a binary response model. Accordingly, logistic regression, a type of binary response model is utilized to estimate the probability of a drafted player reaching MLB based on the values of the set of covariates included in the model.

Logistic regression offers several advantages over traditional linear regression or more specifically a linear probability model, which is simply ordinary least squares (OLS) estimated on a binary variable. For example, a logistic regression produces estimated probabilities that are constrained between zero and one. In some cases, a linear probability model would produce predicted probabilities less than zero and greater than one, resulting in values which are theoretically inappropriate (Horowitz & Savin, 2001). OLS also assumes constant variance of the dependent variable across values of the covariates. With a binary dependent variable, this is assumption is violated as the variance will approach zero as the probability approaches both the zero and one boundaries. Logistic regression relaxes the assumption of homoscedasticity, or constant variance, which is a hallmark of the traditional linear model. Lastly, logistic regression does not impose the restriction of normal error distribution, an assumption which would be violated if using OLS on a binary dependent variable (Long, 1997).

Instead of imposing a linear restriction as is the case with a linear probability model, a logistic regression, or logit model, uses the cumulative logistic distribution function to force probabilities between zero and one. Logit models are also traditionally estimated by maximum likelihood, which performs well in large samples (Horowitz & Savin, 2001). Because of the numerous advantages as compared to the linear probability model, logistic regression will be utilized here.

The general specification of a binary probability model states that the probability of a player reaching the major leagues is conditional on the set of variables specified in the model. This general form applied to the current scenario is illustrated below:

$$P(Y = 1|tr, dr, de) = F(\beta_0 + \beta_1 tr + \beta_2 dr + \beta_3 de) + \varepsilon,$$

where Y = 1 if the drafted player played in MLB during his career, F is a function of the specified model, *tr* is a vector of training background variables, *dr* is a vector of draft specific variables, *de* is a vector of demographic variables, and ε is the error term. β_0 is a constant and β_1 to β_3 are the coefficients to be estimated. In order for a logistic regression model to be properly specified, the function of the specified model must be fit to the cumulative logistic distribution function (Horowitz & Savin, 2001). The functional form of the specified logistic regression model then becomes:

$$P(Y = 1|tr, dr, de) = F_L(\beta_0 + \beta_1 tr + \beta_2 dr + \beta_3 de) + \varepsilon,$$

where F_L is the cumulative logistic distribution function, $F_L(z) = 1 / (1 + e^{-z})$, and

$$z = \beta_0 + \beta_1 tr + \beta_2 dr + \beta_3 de.$$

To identify the factors which influence the probability of a drafted player reaching MLB, the following equation is specified which models the relationship between the dependent variable and the covariates derived from the functional form outlined above.

LOGIT (PLAYEDINMLB = 1) = β_0 + β_1 NOCLASSIFICATION + β_2 JUCO + β_3 4YEARCOLLEGE + β_4 ROUND+ β_5 ROUNDPICK + β_6 LHP + β_7 C + β_8 1B + β_9 3B + β_{10} 2B + β_{11} SS + β_{12} OF + β_{13} CANADA + β_{14} INTERNATIONAL+ β_{15} MIDATLANTIC + β_{16} SOUTHATLANTIC + β_{17} EASTSOUTHCENTRAL + β_{18} EASTNORTHCENTRAL + β_{19} WESTSOUTHCENTRAL + β_{20} WESTNORTHCENTRAL + β_{21} MOUNTAIN + β_{22} PACIFIC + β_{23} COMPORSUPP + β_{24} REDRAFTED + β_{25} THROWSLEFT+ β_{26} BATSLEFT + β_{27} BATSSWITCH + β_{28} HEIGHT + β_{29} WEIGHT+ ε

While Table 3.1 provides a list and description of all variables, they will also be presented here. The dependent variable for the first objective is PLAYEDINMLB. This variable is an indicator equal to one if the player played at least one game at the major league level and is coded equal to zero otherwise.

As highlighted in the description of logistic regression above, the covariates included in the empirical estimations can be divided into three distinct categories; training background, draft specific and demographic variables. The training background variables are included to capture whether the player was drafted out of high school, a four-year institution or from a junior or community college. HIGHSCHOOL is an indicator variable equal to one if the player was drafted out of high school. Likewise, 4YEARCOLLEGE and JUCO are indicator variables identifying whether the player was drafted out of a four-year institution or a junior or community college, respectively. Despite the fact that the vast majority of players are drafted from one of these three categories, there are some players that do not fit into a specific classification. These are often players that are home schooled or are international players without a high school or collegiate affiliation. Subsequently, NOCLASSIFICATION is an indicator variable equal to one, identifying the non-specific training background of these drafted players.

Draft specific variables are included to account for differences between drafted players that are a function of the mechanisms associated with the reverse-order-of-finish nature of the MLB Draft. ROUND is variable representing the round the player was drafted in a specific year of the MLB Draft. From 1966 to 1997, draft rules allowed clubs to continue selecting players until every team exhausted their willingness to make a selection. This format often resulted in drafts lasting up to 100 rounds. Beginning in 1998, MLB imposed a fifty round limit to the Draft and continued with clubs making one selection per round in reverse-order-of finish from the previous season’s standings.

ROUNDPICK is an integer variable representing a player's draft position in a specific round.

COMPORSUPP is an indicator variable equal to one if a player was selected as a compensatory or supplementary draft selection. Compensatory picks are awarded to clubs that lost one or more prominent free agents during the previous off-season. Supplementary picks are awarded to clubs as compensation above and beyond a standard compensatory pick based on the loss of a Type A free agent, as specified by the Elias Sport Bureau. Since compensatory and supplemental picks are awarded to clubs in addition to their standard single selection per round, clubs may be willing to take on additional risk when making compensatory or supplemental selections. The inclusion of COMPORSUPP accounts for this possibility. REDRAFTED is an indicator variable equal to one if a player was previously drafted in the MLB Draft and did not sign a professional contract. In this scenario, players are able to retain their amateur eligibility and accrue additional training at the college level while still maintaining the ability to be redrafted at a later time[16]. This variable is included to measure whether or not redrafted players who have larger quantities of training prior to entering the professional baseball labor market have enhanced labor market outcomes as compared to players who are not redrafted.

Demographic variables are included to account for between-player differences at time of selection into the labor market. HEIGHT and WEIGHT are continuous variables included to capture the physical attributes of drafted players. Height is measured in

[16] Players are able to be drafted following the exhaustion of their high school eligibility, at any time during their Junior or Community College playing career or at the conclusion of their junior or senior seasons at a four-year institution. All players with the exception of athletes who have exhausted their collegiate eligibility are able to return to an amateur playing career at the collegiate level following being drafted, assuming they do not sign a contract (MLB.com, 2011).

inches and weight in pounds. BATSRIGHT, BATSLEFT and BATSSWITCH are indicator variables accounting for whether the player bats from the right, left or both sides of the plate. THROWSRIGHT and THROWSLEFT are indicators specifying whether the player throws right-handed or left-handed.

Position indicators are included to determine whether or not labor market outcomes vary based on position played at the time of the Draft. If a player is a pitcher, LHP and RHP account for whether the player is a right-handed or left-handed pitcher. Likewise, C, 1B, 2B, 3B, SS, and OF are variables representing the primary position played by a drafted player at the time of the draft. The number of players selected each year varies by position and due to this, it is appropriate to account for the possibility that there is also variation in labor market outcomes based on position played. The inclusion of indicator variables for position played controls for this possibility.

There is also variation in the geographical background of players selected in the MLB Draft. A common notion exists which hypothesizes that players residing in warmer climates have a developmental advantage over players residing in colder climates as there is an enhanced opportunity for training due to more favorable weather conditions over a longer portion of the calendar year. Indicator variables accounting for the location of drafted players are included to account for this hypothesis. In order to generate unbiased geographical classifications, the US Census Bureau Census Regions and Divisions are utilized for the determination of geographical area territories. NEWENGLAND is indicator set equal to one if a player was drafted from a high school, four-year institutional or junior or community college in Maine, New Hampshire, Vermont, Massachusetts, Rhode Island or Connecticut. MIDATLANTIC is an indicator for players drafted out of New Jersey, New York or Pennsylvania. SOUTHATLANTIC is a dummy variable equal to one if a player was drafted from a school in Delaware, West Virginia, Maryland, Washington DC, Virginia, North Carolina, South Carolina, Georgia or Florida. EASTSOUTHCENTRAL is set equal to one for players drafted out of Kentucky, Tennessee, Mississippi or Alabama. Players drafted from schools in Ohio, Michigan, Indiana, Illinois or Wisconsin are classified to the EAST NORTHCENTRAL region. The WESTSOUTHCENTRAL region is comprised of Louisiana, Arkansas, Oklahoma and Texas. WESTNORTHCENTRAL is an indicator for players drafted from schools in Minnesota, Iowa, Missouri, Kansas, Nebraska, North Dakota and South Dakota. MOUNTAIN represents players from Arizona, Colorado, New Mexico, Utah, Nevada, Wyoming, Idaho and Montana. Lastly, PACIFIC is an indicator for players drafted out of California, Oregon, Washington, Hawaii and Alaska.

In addition to variables controlling for geographical regions within the United States, indicators are also included to account for players drafted from outside of the US. Over time, MLB has relaxed the regulations regarding the drafting of international players, resulting in a larger number of international players being selected in the Draft. CANADA is an indicator set equal to unity if a player is drafted from a Canadian school. Players have also been drafted out of International locations such as Puerto Rico, Dominican Republic, Cuba, Guam, the US Virgin Islands, Holland and Australia. In order to measure whether or not there is variation in labor market outcomes between international and American players, the indicator variable INTERNATIONAL is included. It is reasonable to assume that geographical variation between players' impacts training availability and quality, which subsequently has the capacity to affect the probability of reaching MLB.

Summary statistics for the variables included in this sample are illustrated in Table 3.3.

3.6 Labor Market Outcome 1: Entry into Major League Baseball (Results)

Before evaluating the results of the logistic regression model to estimate the probability of drafted players reaching MLB, it is useful to examine the historical summary statistics. Tables 3.5, 3.6 and 3.7 illustrate the historical probabilities of a drafted player reaching MLB by position, by position and training classification, and by round and training classification, respectively. Table 3.5 demonstrates that independent of other factors, 10.95% of all players drafted in the MLB Draft from 1966 to 2005 have reached the major leagues. Players drafted as pitchers have achieved MLB status at an 11.94% clip, while position players have been 1.84% less likely to reach the majors. Left-handed pitchers have been 1.73% more likely to reach MLB as compared to right-handed pitchers. Specific to position players, individuals drafted as shortstops are the most likely to reach MLB, while those drafted as catchers are the least likely.

Table 3.6 illustrates the historical probabilities of drafted players reaching MLB by position and training classification. The most important finding is the realization that at every position, four-year college draftees have reached MLB at a greater clip than both high school and JUCO draftees. This homogeneous outcome illustrates that four-year college players, with a greater amount of accumulated training prior to selection in the

MLB Draft, have enhanced employment outcomes as measured by the probability of reaching the majors.

The summary statistics in Table 3.7 are also telling as they illustrate the historical probability of drafted players reaching MLB when controlling for the round of selection. Again, the results are powerful as four-year college players have been more likely to reach the majors as compared to high school players in forty-seven of fifty rounds. In fact, in the first seven rounds of the draft, four-year college players have reached MLB at an 11.77% greater rate when holding all else equal. A visual treatment of the differences in the historical probabilities of college versus high school players reaching MLB by round is shown in Figure 3.6.

Summary statistics are beneficial in illustrating that four-year college players have been more likely to reach MLB as compared to high school or JUCO draftees. However, in order to examine the effects of the identified covariates on the predicted probability of a drafted player reaching MLB, the use of a binomial response model is necessary. Table 3.8 illustrates logistic regression estimation results for the probability of a drafted player reaching MLB from 1966 to 2005. Likewise, Tables 3.9 through 3.12 display estimation results in ten year periods over the same forty years of draft data. For ease of interpretation, logistic regression results are reported as odds ratios instead of traditional coefficient values. In a logistic regression, traditional coefficients are transformed into odds ratios by way of the formula:

$$\text{odds ratio} = \exp(b),$$

where b is the value of the coefficient from the estimated logistic regression model. Odds ratios offer the advantage of easier interpretation specific to the impact of covariate values on changes in the odds of a drafted player reaching MLB. However, a byproduct of the use of odds ratios is the elimination of the constant term in the results output, as the constant term is no longer interpretable.

Table 3.8 shows logistic regression results over the entirety of the data sample. A total of 29 covariates are included in the model with seventeen showing significance at the 0.01 level and one more each at both the 0.05 and 0.10 levels. Overall, the model produces a likelihood ratio chi-squared (29) = 7587.41, $p < .0001$, which suggests that the model fits significantly better than the null model. HIGHSCHOOL is withheld from the estimation in order to serve as the comparison group for the training classification categories. RHP and NEWENGLAND are omitted in order to serve as the baseline categories for the position and geographical variables. Likewise, BATSRIGHT and THROWSRIGHT are omitted to serve as the baseline for the batting and throwing variables. These five variables are withheld in each of the five logistic regression models.

The model results in Table 3.8 illustrate that both 4YEARCOLLEGE and JUCO draftees have a higher probability of reaching the major leagues as compared to HIGHSCHOOL players, the reference category. The odds ratio of 1.1976 on 4YEARCOLLEGE is interpreted such that players drafted out of four-year institutions are 19.76% [(1.1976-1) x 100] more likely to reach MLB as compared to high school players, when holding all other variables constant. JUCO players are 15.62% more likely to reach MLB as compared to HIGHSCHOOL. Together, these two variables illustrate that when holding all else equal, players with larger amounts of accumulated training prior to selection in the MLB Draft have enhanced labor market outcomes as measured by the probability of reaching the major leagues.

Beyond the identified training classifications, the draft specific covariates included in the model significantly affect the probability of a drafted player reaching the majors. ROUND is highly significant and produces an odds ratio of .9194. Comparing a first round draftee versus a second round draftee can be interpreted that all else equal, a second round draftee has 8.06% [(1 - .9194) x 100 = 8.06] lower odds of reaching MLB. Examining a first versus third round draftee shows that third round selections have 15.47% [$(.9194)^2$ = .84530 => (1 - .84530) x 100 = 15.47)] lower odds of reaching the major league level. Likewise, ROUNDPICK is also significant at the 0.01 level. A player selected with the seventh pick in a particular round as opposed to the sixth pick in that same round has 0.76% [(1 - .9924) x 100 = 0.76] lower odds of reaching MLB.

REDRAFTED shows strong statistical significance with an odds ratio of 5.2264, meaning that players whom eventually sign a professional contract following being previously drafted have a 422.64% increase in odds of reaching MLB. At first glance, this seems like an exceptional change in odds, but a closer look reveals additional information. Draftees coded equal to one for the REDRAFTED variable have two clearly identifiable qualities. First, these players have been selected at least twice in the MLB Draft – suggesting that the individual is viewed by those selecting into the labor market as a player with substantial potential to succeed. Second, these players, inherent with being redrafted, all accumulate additional training prior to finally entering into the professional baseball labor market. So when holding other factors constant, on average, players that are redrafted are likely to have a greater amount of accumulated training as compared to the remainder of the sample of drafted players.

Players drafted as a compensatory or supplementary pick also show an increased probability of reaching the major leagues as is evidence by the COMPORSUPP variable. This odds ratio can be interpreted as players drafted in a compensatory or supplementary draft position have a 93.93% increase in odds of reaching the majors. Again, this result is not surprising given that compensatory and supplemental picks have historically been awarded within the first few rounds of the draft.

Position indicator variables were included to measure the effects of player position at the time of the draft on the probability of reaching MLB. As compared to the reference category, RHP, every position indicator variable was statistically significant at the 0.01 level. A substantial difference was found between the baseline and the LHP indicator, as drafted left-handed pitchers have 64.08% lower odds of reaching MLB as compared to drafted right-handed pitchers. Overall, every position indicator has a statistically significant odds ratio below one, suggesting that drafted right-handed pitchers have a higher probability of reaching MLB as compared to all of the other positions included in the model.

Other demographic variables included in the model were also highly significant. As evidenced by THROWSLEFT, players of any position who throw left-handed have a 168.91% increase in the odds of reaching MLB as compared to players who throw right-handed. Likewise BATSLEFT and BATSSWITCH were also significant at 0.01. As such, left-handed hitters have a 132.36% increase in odds and switch hitters have a 217.74% increase in the odds of reaching the majors as compared to players who bat right-handed only. HEIGHT and WEIGHT are also significantly associated with the probability of a drafted player reaching MLB. HEIGHT is negatively associated with the probability of reaching MLB as a one standard deviation increase in HEIGHT (SD = 1.335 inches) above the mean of 73.54 inches results in a 7.80% decrease in the odds of reaching MLB. Alternatively, WEIGHT is positively associated with the probability of reaching MLB. A one standard deviation increase in WEIGHT (SD = 11.09 pounds) above the mean of 195.17 pounds equates to a 12.50% increase in the odds of reaching the majors.

Geographical indicator variables were included in the estimation to control for the possibility that players drafted from specific regions of the country or internationally may reach the majors at varying rates. As compared to NEWENGLAND, the reference category, none of the geographical variables produced a significant effect. This can be interpreted by stating that players from any specific geographical location do not exhibit characteristics which would significantly alter the odds of them reaching MLB as compared to a player from another location.

Tables 3.9 through 3.12 illustrate logistic regression results from the original data sample divided into ten year periods. The benefit of this approach lies in the ability to examine whether or not the effects of the included covariates have changed over time in relationship to the dependent variable. In lieu of readdressing each of the significant covariates in these four models, the focus will be placed on highlighting specific covariates that have changed markedly over the examination period.

Results from players drafted from 1966-1975 are shown in Table 3.9. CANADA and INTERNATIONAL are excluded from the model as the Draft was restricted to United States amateurs only during this time period. Among the training classification variables, there was little change from the previous model, with the exception that the indicator classifying JUCO draftees was no longer significant. This suggests that there was no statistically significant difference between 1966 to 1975 HIGHSCHOOL and JUCO draftees in the probability of reaching MLB.

In reference to the demographic variables, an interesting change was illuminated regarding the HEIGHT and WEIGHT variables. In the full model, HEIGHT was negatively associated and WEIGHT was positively associated with the probability of reaching MLB. In data including only 1966-1975 draftees, these relationships are not only reversed, but are both statistically significant at standard levels.

Table 3.10 provides logistic regression results for players selected in the MLB Draft from 1976 to 1985. Once again, due to eligibility restrictions regarding players from outside of the United States, CANADA and INTERNATIONAL are excluded from the model. The most pertinent finding is illustrated in the 4YEARCOLLEGE odds ratio. While athletes drafted out of four-year institutions were still more likely to reach the majors as compared to those selected from high school, there is no statistically significant difference between the training classifications at standard levels of significance.

Estimation results for players drafted from 1986 to 1995 are shown in Table 3.11. Similar to the model using the entire data sample, 4YEARCOLLEGE is again positively and significantly associated with the probability of reaching MLB. The odds ratio of 1.3030 demonstrates that players drafted out of four-year colleges were 30.30% more likely to reach the majors as compared to players drafted out of high school. The estimation from the 1986-1995 sample also produced the first significant geographical indictor as players from the MOUNTAIN region had 39.67% lower odds of reaching MLB as compared to the baseline region. Additionally, both HEIGHT and WEIGHT reverted back to having significant negative and positive respective effects on the probability of a player reaching MLB.

Output from the final logistic regression model which is generated from players drafted from 1996 to 2005, is highlighted in Table 3.12. The most crucial result is again found in the 4YEARCOLLEGE indicator, which is non-significant. Along with the other training classification indicators showing non-significance, this can be interpreted that for players drafted between 1996 and 2005, differences in training background are not significantly associated with the probability of a player reaching the majors. WEIGHT was also highly significant and positively linked with the probability of a drafted player reaching MLB. Based on WEIGHT having a mean value of 197.91 over the examination period, a one standard deviation (SD = 15.54) increase in weight increases a player's odds of reaching MLB by 70.50%. HEIGHT was negatively associated with probability of entry into MLB for players drafted between 1996 and 2005. Namely, a one standard deviation (SD = 1.86) increase in a player's height decreases the probability of reaching MLB by 24.20%.

In addition to traditional estimation results, one advantage of the use of the specified methodology lies in the ability to create artificial player profiles in order to predict a future player's probability of reaching MLB based on training specific, draft specific and demographic characteristics. Table 3.21 illustrates several artificial player profiles and the corresponding predicted probabilities of entry into MLB based on the estimation results from the logistic regression model estimated on the 1966 to 2005 data sample.

3.7 Labor Market Outcome 2: Major League Baseball Career Duration (Data/ Methods)

The second objective aims to estimate the major league career duration of drafted players from each of the three training groups. Historical player data is again used as this data set includes players selected in the MLB Draft from 1966 to 1997. The data were collected from the MLB Draft pages at www.baseball-reference.com. Only players that played at least one game in the major leagues are included in the sample. In total, 3,557 observations are included. Again, in order to avoid creating a biased sample, players chosen in the 1997 draft class are the final year of draftees included as the average major league career lasts 6.5 years. In order to allow for a player to have both an average length minor league and major league career, players drafted in 1998 and beyond are not incorporated into the sample. Including players drafted in 1998 and beyond would force a large number of active players into the sample, which is a concern when estimating a duration model. This choice will be discussed in more detail later in the methods section.

Similar to the first objective, this portion of the analysis again withholds observations for players who were drafted but did not sign a professional contract and returned to school. Holding out this subsample of players is appropriate because even though a player may be drafted up to five times in the June First-Year Draft, they only are able to sign a professional contract and enter the professional baseball labor market once. In order to capture only the final time a player is drafted, previous player observations are withheld.

Duration modeling is the empirical approach utilized to analyze the major league career duration of players selected in the MLB Draft. Because the variable of interest is the length of a drafted player's major league playing career, duration modeling, which is also known as hazard modeling or survival modeling, offers significant advantages over other statistical techniques. For example, the use of a logistic regression where the dependent variable is an indicator representing whether or not the player exited MLB is a potential option, but is also problematic. Specifically, this technique cannot incorporate the effect of duration, or the amount of time spent in the league prior to the occurrence of exiting the league. Additionally, standard regression techniques prove challenging because of the case of right-censored observations (Hoang & Rascher, 1999; Frick, et al., 2007). As is the case with many studies utilizing duration modeling, not all players will have exited the league within the time period of the study. An advantage of duration modeling allows for these observations to be identified.

In duration modeling, the hazard rate is the dependent variable. The hazard rate measures the probability of a player exiting MLB, either voluntarily or involuntarily at a point in time based on the fact that the player has reached that point in time. The model used in this study will be specified in which the hazard rate, or the probability of a player exiting the major leagues, is a function of time. This allows the risk of a player exiting the labor market to either increase or decrease over time. The utilization of duration modeling also assumes that the hazard rate is a function of the covariates included in the model, allowing for an interpretation of the effects of the included variables on the response (Cox, 1972).

Among a variety of hazard model options, the Cox (1972) proportional hazard model was selected. The Cox model is the preferred option because it does not impose any functional form to the shape of the hazard function over time – meaning that the hazard could increase, decrease or shift. Therefore, the baseline hazard function is not given a specific parameterization and the model covariates have the ability to shift the baseline hazard function. Based on these computational advantages, the semiparametric Cox (1972) model is the most popular of all of the hazard models (Cleves, et. al, 2010). Alternatives to the Cox model are parametric models which impose an estimate of the baseline hazard model. Selecting a parametric hazard model, such as the Weibull, exponential or Gompertz has the ability to produce very efficient estimates, but the accuracy of the estimates are contingent on making correct assumptions regarding the underlying shape of the hazard function. If incorrect, choosing a parametric model over the Cox (1972) model can produce less than efficient estimates (Cleves, et. al, 2010). Due in part to these considerations, the Cox model was selected.

The conditional hazard function in the Cox model provides the conditional probability of a player exiting MLB at time *t* based on the values of the covariates included in the model and is expressed as follows:

$$\lambda\ (t \mid \mathrm{b, d, p}) = \lambda_o\,(t)\ \exp(\beta_1\ X_b + \beta_2\ X_d + \beta_3\ X_p) + \varepsilon,$$

where λ is the hazard rate, t is the time variable, $\lambda_o(t)$ is the baseline hazard function, X_b is a vector of training background variables. X_d is a vector of draft specific variables and X_p is a vector of demographic variables. β_1, β_2 and β_3 are the coefficients to be estimated and ε is the error term. The hazard rate is exponentiated in order to keep the hazard rate greater than zero. The baseline hazard can be interpreted as the probability of a drafted player exiting MLB when each of the covariates in the model is equivalent to zero. The exact training, draft and demographic variables included in the model are outlined below.

As stated previously, in order to properly estimate MLB career duration, the use of Cox proportional hazard modeling is appropriate. To identify the factors which impact career duration, the following equation is specified which models the relationship between the dependent variable and the covariates derived from the functional form outlined above.

HAZARD RATE (*TIME* | TRAINING, DRAFT AND DEMOGRAPHIC VARIABLES) = λ_o (*TIME*) [EXP (β_1 NOCLASSIFICATION + β_2 JUCO + β_3 4YEARCOLLEGE + β_4 ROUND+ β_5 ROUNDPICK + β_6 LHP + β_7 C + β_8 1B + β_9 3B + β_{10} 2B + β_{11} SS + β_{12} OF + β_{13} COMPORSUPP + β_{14} REDRAFTED + β_{15} THROWSLEFT + β_{16} BATSLEFT + β_{17} BATSSWITCH + β_{18} HEIGHT + β_{19} WEIGHT + β_{20} 1STSTAGE)] + ε

Table 3.2 provides a list and description of all variables included in the Cox models, but they will also be presented here. The hazard rate is the dependent variable for the Cox hazard models and YEARSPLAYEDMLB is the time variable used to calculate the hazard rate. This is an integer equal to the number of seasons the individual played for any team at the major league level. Players are coded have played one season at the major league level if they played at least one game in the majors during the course of a season.

Similar to the first objective, the covariates included in the empirical estimations for the second labor market objective can be divided into three distinct categories; training background, draft specific and demographic variables. The training background variables are included to capture whether the player was drafted out of high school, a four-year institution or from a junior or community college. HIGHSCHOOL is an indicator variable equal to one if the player was drafted out of high school. Likewise, 4YEARCOLLEGE and JUCO are indicator variables identifying whether the player was drafted out of a four-year institution or a junior or community college. Despite the fact that the vast majority of players are drafted from one of these three categories, there are some players that do not fit into a specific classification. These are often players that are home schooled or are international players without a high school or collegiate affiliation.

Subsequently, NOCLASSIFICATION is an indicator variable equal to one, identifying the non-specific training background of these drafted players.

Draft specific variables are included to account for differences between drafted players that are a function of the mechanisms associated with the reverse-order-of-finish nature of the MLB Draft. ROUND is variable representing the round the player was drafted in a specific year of the MLB Draft. From 1966 to 1997, draft rules allowed clubs to continue selecting players until every team exhausted their willingness to make a selection. This format often resulted in drafts lasting up to 100 rounds. Beginning in 1998, MLB imposed a fifty round limit to the Draft and continued with clubs making one selection per round in reverse-order-of finish from the previous season's standings. ROUNDPICK is an integer variable representing a player's draft position in a specific round.

COMPORSUPP is an indicator variable equal to one if a player was selected as a compensatory or supplementary draft selection. Compensatory picks are awarded to clubs that lost one or more prominent free agents during the previous off-season. Supplementary picks are awarded to clubs as compensation above and beyond a standard compensatory pick based on the loss of a Type A free agent, as specified by the Elias Sport Bureau. Since compensatory and supplemental picks are awarded to clubs in addition to their standard single selection per round, clubs may be willing to take on additional risk when making compensatory or supplemental selections. The inclusion of COMPORSUPP accounts for this possibility. REDRAFTED is an indicator variable equal to one if a player was previously drafted in the MLB Draft and did not sign a professional contract. In this scenario, players are able to retain their amateur eligibility and accrue additional training at the college level while still maintaining the ability to be redrafted at a later time[17]. This variable is included to measure whether or not redrafted players who have larger quantities of training prior to entering the professional baseball labor market have enhanced labor market outcomes than players who are not redrafted.

Demographic variables are included to account for between-player differences at time of selection into the labor market. HEIGHT and WEIGHT are continuous variables included to capture the physical attributes of drafted players. Height is measured in inches and weight in pounds. BATSRIGHT, BATSLEFT and BATSSWITCH are indicator variables accounting for whether the player bats from the right, left or both sides of the plate. THROWSRIGHT and THROWSLEFT are indicators specifying whether the player throws right-handed or left-handed.

Position indicators are included to determine whether or not labor market outcomes vary based on position played at the time of the Draft. If a player is a pitcher, LHP and RHP account for whether the player is a right-handed or left-handed pitcher. Likewise, C, 1B, 2B, 3B, SS, and OF are variables representing the primary position played by a drafted player at the time of the draft. The number of players selected each year varies by position and due to this, it is appropriate to account for the possibility that there is variation in career duration based on position played. The inclusion of indicator variables for position played controls for this possibility.

Lastly, the case can be made that the two labor market outcomes of interest are not completely independent. Specifically, factors which impact the ability of a player to

[17] Players are able to be drafted following the exhaustion of their high school eligibility, at any time during their Junior or Community College playing career or at the conclusion of their junior or senior seasons at a four-year institution. All players with the exception of athletes who have exhausted their collegiate eligibility are able to return to an amateur playing career at the collegiate level following being drafted, assuming they do not sign a contract (MLB.com, 2011).

reach the major leagues may also play a role in career duration. In order to control for this possibility, the variable 1STSTAGE is included in the Cox hazard model. This variable is the fitted value, or the predicted probability of a player reaching MLB estimated from the first stage logistic regression model. The inclusion of this variable allows for an investigation of the covariates which influence career duration while also accounting for the factors influencing the player reaching the major leagues in the first place.

Table 3.4 shows summary statistics for the variables included in this sample.

3.8 Labor Market Outcome 2: Major League Baseball Career Duration (Results)

The second labor market objective is focused on analyzing MLB career duration for players selected in the MLB Draft. Tables 3.13, 3.14 and 3.15 illustrate the average career duration of drafted players who have reached MLB by position, by position and training classification, and by round and training classification, respectively. Table 3.13 displays that for players drafted between 1966 and 1997 the average MLB career length is 6.48 seasons. Position players have had longer MLB careers than pitchers by an average of 1.17 seasons and left-handed pitchers have slightly longer careers than right-handed pitchers.

Table 3.14 illustrates career duration by position and training classification. For players of any position, those drafted out of high school have longer careers than those from four-year institutions by an average of 1.24 seasons. This finding holds at every position, but significant differences between high school and four-year college players are evident in the middle infield positions. Players drafted as second basemen out of high school have longer careers by 2.11 years on average as compared to college players. Likewise, high school shortstops stay in the majors for an additional 1.65 seasons on average. Results for non-classified players are included as a reference, but because of low sample sizes, the results are less robust.

Summary statistics for career duration by round of selection and training classification are illustrated in Table 3.15, while a visual treatment is shown in Figure 3.7. This figure shows that for each training classification the mean declines slowly and the variance increases along with the round of selection. In all but one of the first thirty-five rounds of the draft, high school players have had longer career durations than their college counterparts. JUCO players exhibit more variation as compared to the round average and this is not surprising considering markedly less JUCO players have been selected over the history of the Draft.

Tables 3.16, 3.17, 3.18 and 3.19 illustrate results from four separate Cox proportional hazard models. Each model is significant at 0.01, indicating a better fit than the null model. The first table demonstrates results from the entire sample while the following three tables provide estimation results divided into periods similar to that of the logistic regression models. Because the sample for the duration models stops with 1997 Draft data, the fourth and final model utilizes 1986-1995 data in order for comparison with the logistic regression model spanning the same time period. Again, the benefit of multiple models lies in the ability to examine whether or not the effects of the included covariates have changed over time in relationship to the dependent variable. As opposed to readdressing each of the significant covariates each model, the focus will be placed on highlighting specific covariates that have changed markedly over each examination period.

Table 3.16 highlights Cox regression results for the entirety of the 32-year sample. The most pertinent finding is illuminated in the 4YEARCOLLEGE variable. Because the hazard ratio is greater than one at 1.3557, it can be interpreted that players drafted from four-year colleges have a 35.57% increase in the hazard ratio of exiting MLB as compared to high school players, the baseline category. Additionally, JUCO draftees that reach MLB also have a higher probability of exiting MLB as compared to high school players as illustrated by a hazard ratio of 1.1389. These two results do not support the identified theoretical relationship between training and employment outcomes.

The inclusion of 1STSTAGE produces several interesting findings. First, its hazard ratio of 0.2571 is highly significant and suggests that players with higher predicted probabilities of reaching MLB also are significantly less likely to exit MLB given that they have reached MLB. Once accounting for variation in the probability of reaching MLB through 1STSTAGE, the interpretation of REDRAFTED becomes of specific interest. Namely, the hazard ratio on REDRAFTED suggests that players previously selected in the Draft have a 44.64% increase in the likelihood of exiting MLB, when holding all other variables constant. While this may appear to be a counterintuitive result, a closer examination explains this finding. Specifically, the logistic regression models show that redrafted players have significantly higher probabilities of reaching MLB, all else equal. However, redrafted players have slightly shorter MLB careers than non-redrafted players. So after controlling for the probability of reaching MLB, through the inclusion of 1STSTAGE, redrafted players are more likely to exit MLB as compared to non-redrafted players.

Similar to the previous result, the hazard ratio on COMPORSUPP illustrates that players selected with compensatory or supplementary draft picks have a 26.32% increase in the likelihood of exiting MLB, all else equal. Surprisingly, variables that capture when a player is selected in a particular draft are non-significant in determining career duration when holding all else constant. ROUND was non-significant at standard levels and ROUNDPICK just missed significance at the 0.10 level. These results can be explained in a comparable manner to the REDRAFTED variable, as players selected with compensatory selections, supplementary selections and those selected in higher draft slots reach MLB at a significantly higher clip, but have career durations that are similar to that of other MLB players.

Player demographics exhibit significant effects on career duration. WEIGHT is significantly and positively associated with the length of a player's MLB career as a one pound increase in weight above the average is associated with a 0.0035% decrease in the hazard rate. Alternatively, switch hitters show a 21.86% increase in the hazard rate of exiting MLB as compared to players that bat from the right side only. This result can be explained by the large number of switch-hitters that reach MLB coupled with career lengths that are similar to non-switch hitters.

The position played at the time of selection into the labor market also plays a factor in career duration. Players drafted as shortstops, second basemen, third basemen and catchers all have reduced probabilities of exiting MLB as compared to the RHP baseline. Players at shortstop are the least likely to exit MLB once they have entered and are 27.20% less likely to be cut from MLB as compared to right-handed pitchers, the reference category. Similarly, third basemen have a 25.44% reduced probability of exit and catchers are 21.68% less likely to be cut from the league as compared to the baseline. However, no statistically significant differences in career duration exist on the mound as the LHP variable is not significant as compared to the RHP baseline.

Cox estimation results for 1966 to 1975 draftees are illustrated in Table 3.17. Of note is the lack of significant variables in this reduced model as compared to the estimation utilizing the entire sample shown in Table 3.16. Only six variables are significant at conventional levels with 4YEARCOLLEGE, C, SS, OF, REDRAFTED and 1STSTAGE showing statistical significance. Similar to the previous model, players drafted from four-year institutions show a 34.26% increase in the probability of exit from MLB as compared to high school players. During this period, shortstops, catchers and outfielders were the only positions which showed a statistical difference from the RHP baseline. Parallel to the previous model, 1STSTAGE is significant and illuminates that players with higher predicted probabilities of reaching MLB also stay in the MLB labor market longer.

Table 3.18 displays Cox proportional hazard results for players drafted from 1976 to 1985. Only four covariates in the model show significance at standard levels. While four-year college players still show shorter career durations as compared to high school players, there are also several new findings of note in this model. Specifically, 1STSTAGE is not significant even though the hazard ratio coefficient suggests increased career length for players with higher probabilities of entry into the league. Additionally, SS is the only position indicator which is statistically significant as compared to the RHP baseline. The result is in contrast to previous Cox models where the majority of position indicators where significant. Lastly, THROWSLEFT is significant with a hazard ratio of 0.7087, signifying that players which throw left-handed are 29.13% less likely to exit MLB as compared to players which throw right-handed.

Estimation results for players drafted from 1986 to 1995 are illustrated in Table 3.19. Comparable to the results in the full model, players from four-year colleges were 30.06% more likely to exit MLB as compared to high school draftees. Likewise, hazard ratio coefficients on 1STSTAGE, REDRAFTED and WEIGHT all exhibited similar effects to what was seen in the Cox model utilizing the entire 32-year sample. In general, position players are significantly less likely to be removed from the MLB labor market as compared to pitchers and right-handed and left-handed pitchers illustrate no statistically significant differences in regards to career duration.

Beyond traditional estimation results, one advantage of the use of the specified methodology lies in the ability to create artificial player profiles in order to predict a future player's hazard rate at a given point in time. Table 3.21 illustrates several artificial player profiles and the corresponding hazard rates based on the estimation results from the Cox proportional hazard model estimated on the 1966 to 1997 sample.

3.9 Selection into the Labor Market – Data

Lastly, the sample used to measure selection into the labor market is similar to the one used to estimate the first labor market objective. Accordingly, this sample includes players selected in the MLB Draft from 1966 to 2005. In order to correspond to the samples utilized to examine the two labor market objectives, players selected in the 2005 draft class are the final year of draftees included. The primary difference here is that players that do not sign professional contracts following the draft are included in this sample. Because the objective is to understand the behavior of the overall market in terms of its operation and any adjustments over time, there is no need to exclude certain players from the sample.

3.10 Selection into the Labor Market – Results

This section briefly evaluates historical market behavior by way of selection of players into the labor market from the MLB Draft. Figure 3.1 illustrates player selection over the entire sample by training classification and round. The graph clearly shows that over the first four rounds of the draft, MLB clubs have selected significantly more high school players as compared to four-year college players. From rounds five to twenty-eight, the opposite has been true with teams showing a preference for college players. Following the thirty-third round, clubs have returned to favoring high school talent.

Figures 3.2 through 3.5 demonstrate historical selection divided into ten-year periods. The graphs show that the 1966-1975 period differs significantly as compared to the following thirty years of draft history. From 1966 to 1975, the market preferred high school players by a wide margin over four-year college players with this trend staying consistent over each round of the draft. The 1976 to 1985 period begins a significant change in drafting behavior as the market continues to prefer high school players in the first two draft rounds, but moves swiftly toward four-year college players until round 29 where the two training classifications are selected at comparable rates for the remainder of the first 50 rounds. From 1986 to 1995 there is a noticeable shift in selection behavior at the top of the draft as four-year college players are selected more frequently in the first round. Four-year players are also preferred in the middle draft rounds until round 34 where high school players were chosen with more frequency for the remainder of the draft. The selection of JUCO players became relevant for the first time during this period as JUCOs were selected at approximately double the frequency as compared to the previous ten-year period. Lastly, the 1996-2005 period produces market behavior that is comparable to the previous ten-year period. The primary difference is evident in the top end of the draft where college and high school players are drafted at roughly the same frequency over the first four rounds. From rounds five to thirty-one, four-year college players are selected at a significantly higher rate with ratios exceeding 2.5 to 1 in the middle draft rounds.

Overall, the graphs in Figures 3.2 through 3.5 are beneficial in illustrating historical behavior of selection into the labor market. The 1966 to 1975 period appears to be the only sample significantly differing from the others. From 1976 forward, selection into the labor market has been relatively stable with the exception of fluctuation in preference for high school versus four-year college players in the first few draft rounds. Otherwise, four-year college players have been preferred over high school players by a wide margin for the majority of the first half of the draft. In the final third of the draft, the opposite has been true.

3.11 Discussion

The current analysis adds to the sports economics literature through a comprehensive investigation of the relationship between training and employment outcomes in a specialized labor market. Using historical MLB Draft data, the analysis uncovers many relevant general and industry specific findings. Specifically, the theorized relationship between training and the value of the worker to the firm holds in only one of the two labor market outcomes examined.

The logistic regression models estimated demonstrate telling relationships regarding the training accumulated by a worker at the time of selection into the labor market and the probability of reaching the highest level of this specialized market. Over the history of the Draft, players selected from four-year institutions were 19.76% more likely to reach MLB as compared to high school draftees. JUCO players were also significantly more likely to reach the majors with a 15.62% increased probability as compared to the baseline. These results support economic theory regarding the positive relationship between training and employment outcomes.

However, the results of the four Cox proportional hazard models illuminate findings that oppose the theorized relationship between training and employment outcomes. First, as compared to high school players, those selected out of four-year institutions and junior and community colleges face an increased probability of exit from the major leagues at all points in time. In addition to the estimation results, Figure 3.8 illustrates the estimated hazard rates for the three training classifications of players. The figure demonstrates this result as the “fouryear” and “JUCO” hazard rates trend above the “highschool” hazard rate over the entirety of the career. Additionally, the difference in hazard rates between four-year college players and high school players expands as a player’s career lengthens. More generally, it is also evident that the slope of the hazard increases along with time, suggesting that all players face an increased risk of exiting the

MLB labor market as their careers progress. This result can be explained in part through the historical rules of the MLB collective bargaining agreement as players earn less during the first three years of service time and prior to arbitration eligibility. Once arbitration eligibility is attained, salaries tend to increase and players with lower performance to salary ratios are more likely to exit the labor market. This would explain the increasing slope of the hazard following year three.

As an alternative to the hazard function, Figure 3.9 provides an illustration of the survivor function for players from each of the three training classifications. The figure shows that high school players are more likely to stay employed in the labor market at all career points as compared to either four-year college or JUCO players. These results suggest that in this specialized labor market, the theorized positive relationship between training and labor market outcomes does not hold when employment outcomes are measured by career duration. A possible explanation for this result could be tied to the age at which players are drafted. Specifically, four-year college players are on average 3.25 years older than their high school counterparts when drafted (please see Table 3.20). Because older players naturally have less "prime" years available in the MLB labor market before they are replaced by younger and likely cheaper players, age could play a part in explaining the relationship between training and career duration.

Somewhat surprisingly, being selected earlier in a specific draft produces mixed labor market outcomes as specified by the two objectives identified. As expected, ROUND is significant in all five logistic regression models and ROUNDPICK is significant in three of the five models. However, in the Cox hazard models, ROUND is insignificant in each model while ROUNDPICK is significant in only the 1986-1995 sample. These results suggest that in general, players selected earlier in a given draft have enhanced probabilities of labor market success as measured by reaching MLB. Alternatively, after controlling for the probability that the player reaches MLB, ROUND and ROUNDPICK are insignificant in determining MLB career duration.

In terms of demographics, players who throw left-handed and bat left-handed are more likely to reach MLB as compared to those who bat or throw right-handed. However, neither throwing nor batting left-handed produces a significant effect in influencing career duration. Two possible explanations exist for these outcomes. First, it is feasible that players who bat and/or throw left-handed are simply more productive on average in the minor leagues, resulting in lefties reaching the major leagues at higher rates. A second possible interpretation is that MLB clubs prefer to have left-handed pitchers and/or batters on their rosters for strategic purposes – hence why left-handers would reach the major leagues at a higher clip. Since studies have estimated that only 10% of the population is left-handed, it is reasonable to assume that less left-handed baseball players are available in the pool of potential workers (Hardyck & Petrinovich, 1977). Assuming this is accurate; in order to keep the desired number of left-handed pitchers and position players on the active roster, clubs may be forced to promote left-handed players to the major league level. However, once at the major league level, clubs likely retain only the most productive workers, regardless of hand dominance. If this scenario is accurate, it could explain the positive relationships between the left-handed variables in the logistic regression models and the non-significant results in the Cox hazard models.

The effects of HEIGHT and WEIGHT on employment outcomes are also of interest due to the physical nature of the skills associated with playing professional baseball. Due to the demands of the sport, as one would expect, WEIGHT is positively associated with both the probability of reaching MLB and career duration. However, HEIGHT is negatively linked to the probability of reaching the majors and is not a statistically significant factor in determining career duration. If MLB clubs have a propensity to overvalue and hence overdraft taller players, then this negative effect could be visible if an appropriate percentage of these "tall" drafted players are not reaching MLB.

Position indicators also exhibit interesting information regarding the North American professional baseball labor market. Over the entirety of the data set, position players had statistically significant reductions in the probability of reaching MLB as compared to pitchers. But interestingly, position players face a much lower risk of exit once they have reached the majors. Of specific interest are players drafted as C, 2B, SS and 3B. Summary statistics in Table 3.5 show that MLB clubs draft a relatively large amount of middle infielders and catchers. These positions are often viewed as premium positions with many players unable to adequately handle the defensive and strategic responsibilities. Therefore, it is possible that once a shortstop, second basemen or catcher has reached the major league level, the probability of exiting the labor market is reduced because the player has exhibited proficiency in the elements of the position that make it difficult to reach MLB in the first place. If this is indeed true, it would explain the low rates of reaching the major leagues and also the lower risk of being removed from the labor market for players at these premium defensive positions.

Lastly, when comparing selection into the labor market with employment outcomes, there are three specific items that stand out. Specifically, there appears to be some evidence of possible market inefficiency in the 1966 to 1975 draft era. During this time period, high school players were much more likely to be selected into the labor market, yet four-year college players were 56% more likely to reach MLB. On the other hand, high school players did have significantly longer MLB careers once they reached the majors, so claims of full-scale inefficiency should be approached with caution. Secondly, it appears that the market adjusted in the 1976 to 1985 period with the increased selection of four-year college players. Once clubs adjusted, no statistically significant difference in the probability of reaching MLB was evident between high school and four-year college players. Lastly, despite fluctuation at the top of the draft over time in terms of preference for high school versus college players, ROUND is highly significant in the logistic regression models but not so in the Cox hazard models. This suggests that regardless of the training background of the player, MLB clubs have shown a propensity for promoting players to the major league level who are viewed as top end prospects. However, once reaching MLB, clubs appear to retain the most productive players – as is evidenced by the non-significant hazard ratio on the ROUND variable.

3.12 Conclusions

This study contributes to the literature in the field of sports economics by examining the relationship between training and employment outcomes in a specialized labor market. Through the use of historical MLB Draft data, the analysis examines historical selection into the labor market and labor market outcomes as measured by the probability of reaching MLB and MLB career duration. The results highlight a period of market adjustment in the mid-1970s which is illustrated by a drastic increase in the selection of four-year college players. This shift in selection into the labor market has been maintained ever since.

In terms of labor market outcomes, players drafted out of four-year institutions have shown significantly higher probabilities of reaching MLB over the history of the Draft. This result can likely be contributed to the fact that college players have larger quantities of accumulated training when drafted. Accordingly, these players are more developed and have competed against a higher level of competition prior to being drafted, which has allowed MLB clubs to make more accurate predictions regarding the probability of reaching the majors. This result supports economic theory concerning the positive relationship between accumulated training and the value of the worker to the firm as measured by the probability of employment success.

The Cox proportional hazard models used the estimate major league career duration uncover results which oppose traditional economic theory. In all four hazard models, high school players were significantly more likely to remain employed in the MLB labor market at all points in time as compared to four-year college players. This finding suggests that while college players may be less risky draft selections due to the fact that they are much more likely to reach MLB, high school players appear to have a higher payoff to clubs as they remain in the labor market for a longer period of time. This result is supported by the previous work of Spurr and Barber (1994) which demonstrated that players will be removed from the professional baseball labor market if production deteriorates.

In summation, this study empirically tests a long-standing theory regarding the accumulation of training and the value of a worker to the firm, but in the context of a specialized labor market. However, the theorized relationship holds in only one of the two labor market outcomes measured. This suggests a need for future research in the area, specifically in regards to the specialized labor markets in professional sports.

3.13 References

Atkinson, S.E., & Tschirhart, J. (1986). Flexible modeling of time to failure in risky careers. *The Review of Economics and Statistics*, 68(4), 558-566.

Becker, G.S. (1961). Investment in people. *National Bureau of Economic Research*. Unpublished Manuscript.

Becker, G.S. (1962). Investment in human capital: An empirical analysis. *Journal of Political Economy*, 70(5), 9-49.

Borjas, G.J. (2005). Labor economics, 3rd Edition. New York: McGraw-Hill.

Bradbury, J.C. (2011). *Hot stove economics: Understanding baseball's second season*. New York: Copernicus Books.

Brown, M. (2010, August 25). Will leaked MLB financials alter revenue sharing? *Fangraphs.com*. Retrieved October 3, 2011, from: http://www.fangraphs.com /blogs/index. php/will- leaked-mlb-financials-kill-revenue-sharing/.

Burger, J.D., & Walters, S.J. (2009). Uncertain prospects: Rates of return in the baseball draft. *Journal of Sports Economics*, 10(5), 485-501.

Card, D., & Sullivan, D. (1988). Measuring the effect of subsidized training programs on movements in and out of employment. *Econometrica*, 56(3), 497-530.

Castrovince, A. (2010, August 23). Indians opt for Draft signees over free agents. *MLB.com*. Retrieved October 8, 2010, from: http://cleveland.indians.mlb .com/news/article.jsp?ymd=20100823&content_id=13821810&vkey=news_cle&f ext=.jsp&c_id=cle

Cleves, M., et. al. (2010) An introduction to survival analysis using Stata. 3rd Edition. College Station, TX: Stata Corp.

Cox, D.R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society, Series B*, 34(2), 187-220.

Dierker, L. (2005, May 2). Big teams spend more on scouts. *MLB.com*. Retrieved October 3, 2011, from: http://mlb.mlb.com/news/print.jsp?ymd=20050502&conte nt_id=1035748&vkey=perspectives&fext=.jsp&c_id=mlb&affiliateId=Comment Widget.

Fort, R. (2006). Competitive balance in North American professional sports. In J. Fizel (Ed.) *Handbook of Sports Economics Research* (pp. 190-206). Armonk, NY: M.E. Sharpe, Inc

Fort, R., & Quirk, J. (1995). Cross-subsidization, incentives, and outcomes in professional team sports leagues." *Journal of Economic Literature*, 23, 1265-1299.

Gritz, R.M. (1993). The impact of training on the frequency and duration of employment. *Journal of Econometrics*, 57, 21-51.

Groothuis, P.A., & Hill, J.R. (2004). Exit discrimination in the NBA: A duration analysis of career length. *Economic Inquiry*, 42(2), 341-349.

Hardyck, C., & Petrinovich, L.F. (1977). Left-handedness. *Psychological Bulletin*. 87(3), 385-404.

Hoang, H., & Rascher, D. (1999). The NBA, exit discrimination, and career earnings. *Industrial Relations*, 38(1), 69-91.

Horowitz, J.L., & Savin, N.E. (2001). Binary response models: Logits, probits and semiparametrics. *Journal of Economic Perspectives*, 15(4), 43-56.

Kahn, L.M. (2000). The sports business as a labor market laboratory. *Journal of Economic Perspectives*, 14(3), 75-94.

Kaitz, H.B. (1979). Potential use of Markov process models to determine program impact. In F.E. Bloch, (Ed.). *Research in labor economics*, (pp. 259-283). Supplement I: Evaluating manpower training programs. Greenwich, CT: JAI Press.

Koppett, L. (1965, February 28). Baseball's New Draft. *New York Times*: pp. S2

Lewis, M. (2003). *Moneyball: The art of winning an unfair game*. New York: W.W. Norton.

Long, J. S. (1997). *Regression models for categorical and limited dependent variables*. Thousand Oaks, CA: Sage Publications.

Madden, W.C. (2001). *Baseball's first-year player draft, team by team through 1999*. Jefferson, NC: McFarland & Company, Inc.

Mincer, J. (1962). On-the-job training: Costs, returns and some implications. *Journal of Political Economy*, 70(5, Part 2), 50-79.

MLB.com. (2011, September 30) First-year player draft FAQ. *MLB.com*. Retrieved September 30, 2011, from: http://mlb.mlb.com/mlb/draftday/faq.jsp

Rausch, G. (2002, May 16). Evolution of the draft. *MLB.com*. Retrieved September 29, 2011, from: http://mlb.mlb.com/news/article.jsp?ymd=20020516&content_id=26646&vkey=news_mlb&fext=.jsp&c_id=null

Ridder, G. (1986). An event history approach to the evaluation of training, recruitment and employment programmes. *Journal of Applied Econometrics*, 1, 109-126.

Schultz, T.W. (1961). Investment in human capital. *The American Economic Review*, 51(1), 1-17.

Shughart, W.F. II, & Goff, B.L. (1992). Field of dreams: On the construction of professional baseball talent in the minor leagues. In G.W. Scully (Ed.), *Advances in the economics of sport*, (pp. 91-114). Volume 1. Greenwich, CT: JAI Press.

Schultz, T.W. (1962). Reflections on investment in man. *Journal of Political Economy*, 70(5, Part 2), 1-8).

Spurr, S.J. (2000). The baseball draft: A study in the ability to find talent. *Journal of Sports Economics*, 1, 66-85.

Spurr, S.J., & Barber, W. (1994). The effect of performance on a worker's career: Evidence from minor league baseball. *Industrial and Labor Relations Review*, 47(4), 692-708.

Staw, B.M., & Hoang, H. (1995). Sunk costs in the NBA: Why Draft order affects playing time and survival in professional basketball. *Administrative Science Quarterly*, 40(3), 474-494.

Volz, B. (2009). Minority status and managerial survival in Major League Baseball. *Journal of Sports Economics*, 10(5), 522-542.

Winfree, J.A., & Molitor, C. (2006). The value of college: Drafted high school baseball players. *Journal of Sports Economics* 8(4), 378-393.

Table 3.1: Variable Descriptions: Logistic Regression Models

Variable	Exp Sign	Description
PLAYEDINMLB	N/A	indicator variable; 1 = player played in MLB, 0 = otherwise (dependent variable)
HIGHSCHOOL	-	indicator variable; 1 = player is drafted from a high school program, 0 = otherwise (reference category)
JUCO	+/-	indicator variable; 1 = player is drafted from a junior college program, 0 = otherwise
4YEARCOLLEGE	+	indicator variable; 1 = player is drafted from a 4-year college program, 0 = otherwise
NOCLASSIFICATION	+/-	indicator variable; 1 = player is not drafted from a high school, 4-year college or junior college, 0 = otherwise
ROUND	-	draft round in which player is selected
ROUNDPICK	-	pick number in a draft round in which player is selected
LHP	+/-	indicator variable; 1 = player is drafted as a left-handed pitcher, 0 = otherwise
RHP	+/-	indicator variable; 1 = player is drafted as a right-handed pitcher, 0 = otherwise (reference category)
C	+/-	indicator variable; 1 = player is drafted as a catcher, 0 = otherwise
1B	+/-	indicator variable; 1 = player is drafted as a first baseman, 0 = otherwise
3B	+/-	indicator variable; 1 = player is drafted as a third baseman, 0 = otherwise
2B	+/-	indicator variable; 1 = player is drafted as a second baseman, 0 = otherwise
SS	+/-	indicator variable; 1 = player is drafted as a shortstop, 0 = otherwise
OF	+/-	indicator variable; 1 = player is drafted as an outfielder, 0 = otherwise
CANADA	-	indicator variable; 1 = player is drafted out of Canada, 0 = otherwise
INTERNATIONAL	+/-	indicator variable; 1 = player is drafted out of an International location (outside of US and Canada), 0 = otherwise
NEWENGLAND	-	indicator variable; 1 = player is drafted out of Maine, New Hampshire, Vermont, Massachusetts, Rhode Island or Connecticut, 0 = otherwise
MIDATLANTIC	-	indicator variable; 1 = player is drafted out of New Jersey, New York or Pennsylvania, 0 = otherwise
SOUTHATLANTIC	+	indicator variable; 1 = player is drafted out of Delaware, West Virginia, Maryland, Washington DC, Virginia, North Carolina, South Carolina, Georgia or Florida, 0 = otherwise
EASTSOUTHCENTRAL	+	indicator variable; 1 = player is drafted out of Kentucky, Tennessee, Mississippi or Alabama, 0 = otherwise
EASTNORTHCENTRAL	-	indicator variable; 1 = player is drafted out of Ohio, Michigan, Indiana, Illinois or Wisconsin, 0 = otherwise
WESTSOUTHCENTRAL	+	indicator variable; 1 = player is drafted out of Louisiana, Arkansas, Oklahoma and Texas, 0 = otherwise
WESTNORTHCENTRAL	-	indicator variable; 1 = player is drafted out of Minnesota, Iowa, Missouri, Kansas, Nebraska, North Dakota and South Dakota, 0 = otherwise
MOUNTAIN	+/-	indicator variable; 1 = player is drafted out of Arizona, Colorado, New Mexico, Utah, Nevada, Wyoming, Idaho and Montana, 0 = otherwise
PACIFIC	+/-	indicator variable; 1 = player is drafted out of California, Oregon, Washington, Hawaii and Alaska, 0 = otherwise
COMPORSUPP	+/-	indicator variable; 1 = player is drafted with a compensatory or supplementary draft selection, 0 = otherwise
REDRAFTED	+	indicator variable; 1 = player was previously drafted and did not sign a professional contract , 0 = otherwise
THROWSLEFT	+/-	indicator variable; 1 = player throws left-handed, 0 = otherwise
THROWSRIGHT	+/-	indicator variable; 1 = player throws handed-handed, 0 = otherwise (reference category)
BATSLEFT	+/-	indicator variable; 1 = player bats left-handed, 0 = otherwise
BATSRIGHT	+/-	indicator variable; 1 = player bats right-handed, 0 = otherwise (reference category)
BATSSWITCH	+	indicator variable; 1 = player is a switch hitter, 0 = otherwise
HEIGHT	+	height of player in inches
WEIGHT	+	weight of player in pounds

Note: Expected signs are the hypothesized relationship between the identified covariate and the employment outcome specified

Table 3.2: Variable Descriptions: Cox Proportional Hazard Models

Variable	Exp Sign	Description
YEARSPLAYEDMLB	N/A	number of seasons player played in MLB (used to calculate the hazard rate (dependent variable))
HIGHSCHOOL	-	indicator variable; 1 = player is drafted from a high school program, 0 = otherwise (reference category)
JUCO	+/-	indicator variable; 1 = player is drafted from a junior college program, 0 = otherwise
4YEARCOLLEGE	+	indicator variable; 1 = player is drafted from a 4-year college program, 0 = otherwise
NOCLASSIFICATION	+/-	indicator variable; 1 = player is not drafted from a high school, 4-year college or junior college, 0 = otherwise
ROUND	-	draft round in which player is selected
ROUNDPICK	-	pick number in a draft round in which player is selected
LHP	+/-	indicator variable; 1 = player is drafted as a left-handed pitcher, 0 = otherwise
RHP	+/-	indicator variable; 1 = player is drafted as a right-handed pitcher, 0 = otherwise (reference category)
C	+/-	indicator variable; 1 = player is drafted as a catcher, 0 = otherwise
1B	+/-	indicator variable; 1 = player is drafted as a first baseman, 0 = otherwise
3B	+/-	indicator variable; 1 = player is drafted as a third baseman, 0 = otherwise
2B	+/-	indicator variable; 1 = player is drafted as a second baseman, 0 = otherwise
SS	+/-	indicator variable; 1 = player is drafted as a shortstop, 0 = otherwise
OF	+/-	indicator variable; 1 = player is drafted as an outfielder, 0 = otherwise
COMPORSUPP	+/-	indicator variable; 1 = player is drafted with a compensatory or supplementary draft selection, 0 = otherwise
REDRAFTED	+	indicator variable; 1 = player was previously drafted and did not sign a professional contract , 0 = otherwise
THROWSLEFT	+/-	indicator variable; 1 = player throws left-handed, 0 = otherwise
THROWSRIGHT	+/-	indicator variable; 1 = player throws handed-handed, 0 = otherwise (reference category)
BATSLEFT	+/-	indicator variable; 1 = player bats left-handed, 0 = otherwise
BATSRIGHT	+/-	indicator variable; 1 = player bats right-handed, 0 = otherwise (reference category)
BATSSWITCH	+	indicator variable; 1 = player is a switch hitter, 0 = otherwise
HEIGHT	+	height of player in inches
WEIGHT	+	weight of player in pounds
1STSTAGE	+	predicted probability of player reaching MLB (generated from 1st stage logistic regression)

Note: Expected signs are the hypothesized relationship between the identified covariate and the employment outcome specified

Table 3.3: Summary Statistics: 1966-2005 Data for Logistic Regression Models

Variable	N	Mean	SD	Minimum	Maximum
PLAYEDINMLB	40678	0.1157	0.3199	0	1
NOCLASSIFICATION	40678	0.0106	0.1025	0	1
HIGHSCHOOL	40678	0.4291	0.4950	0	1
JUCO	40678	0.1342	0.3409	0	1
4YEARCOLLEGE	40678	0.4261	0.4945	0	1
ROUND	40678	23.8962	16.3779	1	100
ROUNDPICK	40678	12.7422	8.0803	1	52
LHP	40678	0.1301	0.3364	0	1
RHP	40678	0.3230	0.4676	0	1
C	40678	0.1083	0.3107	0	1
1B	40678	0.0535	0.2251	0	1
3B	40678	0.0502	0.2183	0	1
2B	40678	0.0363	0.1871	0	1
SS	40678	0.1010	0.3013	0	1
OF	40678	0.1886	0.3912	0	1
CANADA	40678	0.0058	0.0761	0	1
INTERNATIONAL	40678	0.0105	0.1019	0	1
NEWENGLAND	40678	0.0275	0.1636	0	1
MIDATLANTIC	40678	0.0747	0.2630	0	1
SOUTHATLANTIC	40678	0.2077	0.4057	0	1
EASTSOUTHCENTRAL	40678	0.0674	0.2507	0	1
EASTNORTHCENTRAL	40678	0.1047	0.3062	0	1
WESTSOUTHCENTRAL	40678	0.1282	0.3343	0	1
WESTNORTHCENTRAL	40678	0.0527	0.2234	0	1
MOUNTAIN	40678	0.0641	0.2450	0	1
PACIFIC	40678	0.2550	0.4359	0	1
COMPORSUPP	40678	0.0133	0.1147	0	1
REDRAFTED	40678	0.0522	0.2225	0	1
THROWSLEFT	40678	0.0757	0.2645	0	1
THROWSRIGHT	40678	0.2899	0.4537	0	1
BATSLEFT	40678	0.0998	0.2997	0	1
BATSRIGHT	40678	0.2407	0.4275	0	1
BATSSWITCH	40678	0.0250	0.1562	0	1
HEIGHT	40678	73.6111	1.3347	60	82
WEIGHT	40678	195.1715	11.0906	90	255

Table 3.4: Summary Statistics: 1966-1997 Data for Cox Proportional Hazard Models

Variable	N	Mean	SD	Minimum	Maximum
MLBCAREERDURATION	3557	6.4782	5.0279	1	25
NOCLASSIFICATION	3557	0.0045	0.0669	0	1
HIGHSCHOOL	3557	0.3731	0.4837	0	1
JUCO	3557	0.0779	0.2680	0	1
4YEARCOLLEGE	3557	0.5446	0.4981	0	1
ROUND	3557	10.4557	11.2109	1	89
ROUNDPICK	3557	13.4864	8.1669	1	51
LHP	3557	0.1555	0.3624	0	1
RHP	3557	0.3433	0.4749	0	1
C	3557	0.0905	0.2870	0	1
1B	3557	0.0506	0.2192	0	1
3B	3557	0.0503	0.2186	0	1
2B	3557	0.0298	0.1701	0	1
SS	3557	0.1161	0.3204	0	1
OF	3557	0.1566	0.3635	0	1
COMPORSUPP	3557	0.0484	0.2145	0	1
REDRAFTED	3557	0.2148	0.4107	0	1
THROWSLEFT	3557	0.2311	0.4216	0	1
THROWSRIGHT	3557	0.7689	0.4216	0	1
BATSLEFT	3557	0.2856	0.4518	0	1
BATSRIGHT	3557	0.6297	0.4829	0	1
BATSSWITCH	3557	0.0846	0.2784	0	1
HEIGHT	3557	73.5448	2.2203	66	82
WEIGHT	3557	192.5687	17.2728	150	255
1STSTAGE	3557	0.3055	0.2183	0.0002	0.9393

Table 3.5: Prob. of a Drafted Player Reaching MLB by Position (‘66-‘05)

Player Type	Probability	Standard Deviation	# of Players
All Players	.1095	.3123	42981
All Pitchers	.1194	.3243	19954
All Position Players	.1010	.3013	23027
Left Handed Pitchers	.1321	.3386	5685
Right-Handed Pitchers	.1148	.3188	14204
Catchers	.0899	.2860	4585
First Basemen	.1019	.3026	2266
Second Basemen	.0964	.2952	1515
Third Basemen	.1144	.3184	2115
Shortstops	.1234	.3289	4222
Outfielders	.0932	.2907	8015

Note: Player type is determined by how player was classified at the time of the draft.

Table 3.6: Prob. of a Drafted Player Reaching MLB by Position and Class. (‘66-‘05)

Player Type	Classification	Probability	Standard Deviation	# of Players
All Players	High School	.1252	.2404	20839
	4-Year College	.1585	.3310	18510
	JUCO	.0960	.3652	6555
	No Classification	.0524	.2946	488
All Pitchers	High School	.0956	.2941	8410
	4-Year College	.1621	.3686	8240
	JUCO	.0746	.2628	3095
	No Classification	.0574	.2332	209
All Position Players	High School	.0835	.2767	10309
	4-Year College	.1326	.3391	9609
	JUCO	.0619	.2410	2861
	No Classification	.0524	.2233	248
Left-Handed Pitchers	High School	.1024	.3032	2423
	4-Year College	.1806	.3848	2359
	JUCO	.0854	.2797	843
	No Classification	.0833	.2787	60
Right-Handed Pitchers	High School	.0934	.2911	5951
	4-Year College	.1551	.3620	5861
	JUCO	.0708	.2565	2246
	No Classification	.0479	.2144	146
Catchers	High School	.0735	.2611	2203
	4-Year College	.1185	.3233	1789
	JUCO	.0693	.2543	548
	No Classification	.0000	.0000	45
First Basemen	High School	.0814	.2736	909
	4-Year College	.1329	.3396	1046
	JUCO	.0572	.2327	297
	No Classification	.0714	.2673	14
Second Basemen	High School	.0729	.2603	384
	4-Year College	.1102	.3132	926
	JUCO	.0825	.2758	194
	No Classification	.0000	.0000	11
Third Basemen	High School	.0975	.2967	903
	4-Year College	.1519	.3591	935
	JUCO	.0460	.2098	261
	No Classification	.0000	.0000	16
Shortstops	High School	.1028	.3038	2179
	4-Year College	.1603	.3670	1535
	JUCO	.0998	.3000	451
	No Classification	.1053	.3096	57
Outfielders	High School	.0768	.2664	3579
	4-Year College	.1281	.3343	3254
	JUCO	.0448	.2070	1093
	No Classification	.0674	.2522	89

Note: Player type is determined by how player was classified at the time of the draft.

Table 3.7: Prob. of a Drafted Player Reaching MLB by Round and Class. ('66-'05)

Round	High School	4-Year College	JUCO	Round Average
1	.5809	.7607	.4500	.6570
2	.3524	.6156	.3871	.4525
3	.2963	.4345	.3000	.3698
4	.2471	.3656	.3143	.3006
5	.2073	.3250	.2632	.2667
6	.1663	.2869	.3200	.2311
7	.1422	.2744	.2295	.2130
8	.1184	.1964	.1447	.1619
9	.1046	.1934	.1867	.1541
10	.0993	.1801	.1974	.1465
11	.1078	.1731	.0843	.1387
12	.0636	.1492	.1333	.1179
13	.0761	.1209	.1183	.1040
14	.0702	.0969	.0636	.0834
15	.0540	.0904	.1415	.0823
16	.0724	.0963	.0667	.0850
17	.0499	.1000	.1485	.0849
18	.0508	.0859	.0877	.0719
19	.0466	.0840	.1300	.0739
20	.0562	.0707	.1429	.0756
21	.0369	.0688	.0857	.0592
22	.0408	.0530	.0424	.0477
23	.0382	.0589	.1016	.0571
24	.0248	.0711	.0726	.0561
25	.0404	.0535	.0579	.0499
26	.0265	.0690	.0625	.0511
27	.0245	.0459	.0226	.0341
28	.0093	.0431	.0584	.0325
29	.0324	.0465	.0496	.0410
30	.0240	.0562	.0719	.0453
31	.0132	.0374	.0694	.0335
32	.0134	.0447	.0308	.0305
33	.0059	.0654	.0813	.0393
34	.0156	.0498	.0226	.0286
35	.0125	.0094	.0385	.0163
36	.0169	.0308	.0541	.0277
37	.0164	.0365	.0154	.0220
38	.0108	.0549	.0511	.0331
39	.0108	.0452	.0076	.0191
40	.0074	.0359	.0087	.0161
41	.0038	.0329	.0088	.0129
42	.0111	.0511	.0165	.0222
43	.0000	.0331	.0407	.0183
44	.0082	.0160	.0603	.0224
45	.0130	.0286	.0080	.0150
46	.0000	.0000	.0179	.0045
47	.0220	.0316	.0174	.0226
48	.0172	.0120	.0010	.0143
49	.0145	.0120	.0000	.0104
50	.0047	.0563	.0000	.0131

Table 3.8: Determinants of the Prob. of a Drafted Player Reaching MLB (‘66-‘05)

Variable	Odds Ratio	Standard Error	z-statistic	p-value
NOCLASSIFICATION	0.6561	0.1506	-1.84	0.066
HIGHSCHOOL	baseline	baseline	baseline	baseline
JUCO	1.1562	0.0787	2.13	0.033
4YEARCOLLEGE	1.1976	0.0486	4.44	0.000
ROUND	0.9194	0.0018	-43.02	0.000
ROUNDPICK	0.9924	0.0022	-3.48	0.000
RHP	baseline	baseline	baseline	baseline
LHP	0.3592	0.0285	-12.89	0.000
C	0.6376	0.0424	-6.77	0.000
1B	0.4108	0.0385	-9.49	0.000
3B	0.7826	0.0655	-2.93	0.003
2B	0.6764	0.0714	-3.70	0.000
SS	0.8491	0.0541	-2.57	0.010
OF	0.4445	0.0262	-13.77	0.000
CANADA	1.2079	0.3474	0.66	0.511
INTERNATIONAL	1.2457	0.2605	1.05	0.294
NEWENGLAND	baseline	baseline	baseline	baseline
MIDATLANTIC	0.8493	0.1066	-1.30	0.193
SOUTHATLANTIC	0.8996	0.1009	-0.94	0.346
EASTSOUTHCENTRAL	0.8839	0.1101	-0.99	0.322
EASTNORTHCENTRAL	0.9978	0.1177	-0.02	0.985
WESTSOUTHCENTRAL	0.9429	0.1087	-0.51	0.610
WESTNORTHCENTRAL	0.9107	0.1198	-0.71	0.477
MOUNTAIN	0.9874	0.1247	-0.10	0.920
PACIFIC	0.9925	0.1101	-0.07	0.946
COMPORSUPP	1.9393	0.1937	6.63	0.000
REDRAFTED	5.2264	0.2944	29.36	0.000
THROWSRIGHT	baseline	baseline	baseline	baseline
THROWSLEFT	2.6891	0.2206	12.06	0.000
BATSRIGHT	baseline	baseline	baseline	baseline
BATSLEFT	2.3236	0.1389	14.11	0.000
BATSSWITCH	3.1774	0.2487	14.77	0.000
HEIGHT	0.9412	0.0119	-4.79	0.000
WEIGHT	1.0107	0.0015	7.15	0.000

Note 1: LR Test = 7587.41; N = 40678

Note 2: Logistic regression

Table 3.9: Determinants of the Prob. of a Drafted Player Reaching MLB (‘66-‘75)

Variable	Odds Ratio	Standard Error	z-statistic	p-value
NOCLASSIFICATION	0.1411	0.1467	-1.88	0.060
HIGHSCHOOL	baseline	baseline	baseline	baseline
JUCO	1.3516	0.2876	1.42	0.157
4YEARCOLLEGE	1.5684	0.1592	4.43	0.000
ROUND	0.9231	0.0050	-14.63	0.000
ROUNDPICK	0.9856	0.0064	-2.22	0.027
RHP	baseline	baseline	baseline	baseline
LHP	0.3390	0.0639	-5.74	0.000
C	0.7630	0.1169	-1.77	0.077
1B	0.4168	0.0953	-3.83	0.000
3B	0.8095	0.1628	-1.05	0.293
2B	0.5977	0.1643	-1.87	0.061
SS	0.8493	0.1232	-1.13	0.260
OF	0.4407	0.0656	-5.51	0.000
CANADA	x	x	x	x
INTERNATIONAL	x	x	x	x
NEWENGLAND	baseline	baseline	baseline	baseline
MIDATLANTIC	0.9683	0.2599	-0.12	0.905
SOUTHATLANTIC	0.8834	0.2313	-0.47	0.636
EASTSOUTHCENTRAL	0.8594	0.2529	-0.51	0.607
EASTNORTHCENTRAL	1.1780	0.3074	0.63	0.530
WESTSOUTHCENTRAL	1.0996	0.2944	0.35	0.723
WESTNORTHCENTRAL	0.9534	0.2907	-0.16	0.876
MOUNTAIN	1.1111	0.3485	0.34	0.737
PACIFIC	1.4151	0.3463	1.42	0.156
COMPORSUPP	x	x	x	x
REDRAFTED	19.2151	6.0591	9.37	0.000
THROWSRIGHT	baseline	baseline	baseline	baseline
THROWSLEFT	2.8734	0.5628	5.39	0.000
BATSRIGHT	baseline	baseline	baseline	baseline
BATSLEFT	2.8193	0.4181	6.99	0.000
BATSSWITCH	4.2088	0.8103	7.46	0.000
HEIGHT	1.0889	0.0397	2.34	0.019
WEIGHT	0.9809	0.0049	-3.88	0.000

Note 1: LR Test = 1198.31; N = 8049

Note 2: Logistic regression

Table 3.10: Determinants of the Prob. of a Drafted Player Reaching MLB (‘76-‘85)

Variable	Odds Ratio	Standard Error	z-statistic	p-value
NOCLASSIFICATION	0.9260	0.4514	-0.16	0.875
HIGHSCHOOL	baseline	baseline	baseline	baseline
JUCO	0.7114	0.1974	-1.23	0.220
4YEARCOLLEGE	1.1383	0.1077	1.37	0.171
ROUND	0.9080	0.0054	-16.24	0.000
ROUNDPICK	0.9801	0.0058	-3.40	0.001
RHP	baseline	baseline	baseline	baseline
LHP	0.3095	0.0599	-6.06	0.000
C	0.6881	0.1007	-2.55	0.011
1B	0.3512	0.0764	-4.81	0.000
3B	0.4922	0.1034	-3.37	0.001
2B	0.4688	0.1122	-3.17	0.002
SS	0.5025	0.0782	-4.42	0.000
OF	0.3274	0.0463	-7.90	0.000
CANADA	x	x	x	x
INTERNATIONAL	x	x	x	x
NEWENGLAND	baseline	baseline	baseline	baseline
MIDATLANTIC	0.7745	0.2126	-0.93	0.352
SOUTHATLANTIC	0.7633	0.1915	-1.08	0.282
EASTSOUTHCENTRAL	0.6666	0.1927	-1.40	0.161
EASTNORTHCENTRAL	0.8842	0.2288	-0.48	0.634
WESTSOUTHCENTRAL	0.6799	0.1776	-1.48	0.140
WESTNORTHCENTRAL	0.7693	0.2376	-0.85	0.396
MOUNTAIN	0.8423	0.2407	-0.60	0.548
PACIFIC	0.7507	0.1841	-1.17	0.242
COMPORSUPP	2.4968	0.6960	3.28	0.001
REDRAFTED	26.1410	4.8129	17.73	0.000
THROWSRIGHT	baseline	baseline	baseline	baseline
THROWSLEFT	3.6606	0.7429	6.39	0.000
BATSRIGHT	baseline	baseline	baseline	baseline
BATSLEFT	3.4742	0.5468	7.91	0.000
BATSSWITCH	5.9305	1.0969	9.62	0.000
HEIGHT	1.1180	0.0401	3.11	0.002
WEIGHT	0.9879	0.0049	-2.46	0.014

Note 1: LR Test = 2066.90; N = 7359

Note 2: Logistic regression

Table 3.11: Determinants of the Prob. of a Drafted Player Reaching MLB (‘86-‘95)

Variable	Odds Ratio	Standard Error	z-statistic	p-value
NOCLASSIFICATION	0.6756	0.2890	-0.92	0.359
HIGHSCHOOL	baseline	baseline	baseline	baseline
JUCO	1.3086	0.1414	2.49	0.013
4YEARCOLLEGE	1.3030	0.1003	3.44	0.001
ROUND	0.9387	0.0028	-21.57	0.000
ROUNDPICK	1.0022	0.0042	0.53	0.595
RHP	baseline	baseline	baseline	baseline
LHP	0.3176	0.0464	-7.85	0.000
C	0.6023	0.0744	-4.10	0.000
1B	0.2995	0.0562	-6.43	0.000
3B	0.7174	0.1134	-2.10	0.036
2B	0.5508	0.1165	-2.82	0.005
SS	0.8964	0.1026	-0.95	0.340
OF	0.3612	0.0387	-9.52	0.000
CANADA	2.0326	0.9087	1.59	0.113
INTERNATIONAL	1.3792	0.4435	1.00	0.317
NEWENGLAND	baseline	baseline	baseline	baseline
MIDATLANTIC	0.9300	0.2280	-0.30	0.767
SOUTHATLANTIC	0.8682	0.1888	-0.65	0.516
EASTSOUTHCENTRAL	0.8334	0.1986	-0.76	0.445
EASTNORTHCENTRAL	0.8716	0.1987	-0.60	0.547
WESTSOUTHCENTRAL	0.8457	0.1895	-0.75	0.454
WESTNORTHCENTRAL	0.6630	0.1700	-1.60	0.109
MOUNTAIN	0.6033	0.1515	-2.01	0.044
PACIFIC	0.7645	0.1665	-1.23	0.218
COMPORSUPP	1.6832	0.2987	2.93	0.003
REDRAFTED	21.7193	3.3351	20.05	0.000
THROWSRIGHT	baseline	baseline	baseline	baseline
THROWSLEFT	4.4234	0.7277	9.04	0.000
BATSRIGHT	baseline	baseline	baseline	baseline
BATSLEFT	3.0795	0.3823	9.06	0.000
BATSSWITCH	4.9194	0.7659	10.23	0.000
HEIGHT	0.9302	0.0254	-2.65	0.008
WEIGHT	1.0073	0.0033	2.24	0.025

Note 1: LR Test = 3213.92; N = 14134

Note 2: Logistic regression

Table 3.12: Determinants of the Prob. of a Drafted Player Reaching MLB (‘96-‘05)

Variable	Odds Ratio	Standard Error	z-statistic	p-value
NOCLASSIFICATION	0.9544	0.3606	-0.12	0.902
HIGHSCHOOL	baseline	baseline	baseline	baseline
JUCO	1.1100	0.1292	0.90	0.370
4YEARCOLLEGE	0.9205	0.0727	-1.05	0.294
ROUND	0.9172	0.0032	-24.41	0.000
ROUNDPICK	0.9954	0.0035	-1.30	0.194
RHP	baseline	baseline	baseline	baseline
LHP	0.5910	0.0908	-3.42	0.001
C	0.5755	0.0751	-4.23	0.000
1B	0.4285	0.0742	-4.90	0.000
3B	0.9775	0.1464	-0.15	0.879
2B	1.1221	0.2119	0.61	0.542
SS	1.2150	0.1521	1.56	0.120
OF	0.6225	0.0680	-4.34	0.000
CANADA	0.8245	0.3391	-0.47	0.639
INTERNATIONAL	0.8980	0.3202	-0.30	0.763
NEWENGLAND	baseline	baseline	baseline	baseline
MIDATLANTIC	0.7396	0.1941	-1.15	0.250
SOUTHATLANTIC	0.9984	0.2198	-0.01	0.994
EASTSOUTHCENTRAL	1.0214	0.2436	0.09	0.929
EASTNORTHCENTRAL	1.0022	0.2394	0.01	0.993
WESTSOUTHCENTRAL	1.1130	0.2493	0.48	0.633
WESTNORTHCENTRAL	1.2849	0.3201	1.01	0.314
MOUNTAIN	1.2490	0.2988	0.93	0.353
PACIFIC	1.1048	0.2441	0.45	0.652
COMPORSUPP	1.7425	0.2553	3.79	0.000
REDRAFTED	2.4861	0.2073	10.92	0.000
THROWSRIGHT	baseline	baseline	baseline	baseline
THROWSLEFT	1.5274	0.2257	2.87	0.004
BATSRIGHT	baseline	baseline	baseline	baseline
BATSLEFT	1.7442	0.1716	5.65	0.000
BATSSWITCH	1.6191	0.2314	3.37	0.001
HEIGHT	0.8618	0.0167	-7.68	0.000
WEIGHT	1.0349	0.0023	15.23	0.000

Note 1: LR Test = 2062.89; N = 11136

Note 2: Logistic regression

Table 3.13: MLB Career Duration of Drafted Players by Position (‘66-‘97)

Player Type	Career Duration	Standard Deviation	# of Players
All Players	6.4782	5.0279	3557
All Pitchers	5.8890	4.7278	1774
All Position Players	7.0639	5.2458	1783
Left Handed Pitchers	6.1284	4.9890	553
Right-Handed Pitchers	5.7813	4.6024	1221
Catchers	7.0373	5.1499	322
First Basemen	7.1722	5.6882	180
Second Basemen	6.6415	4.3429	106
Third Basemen	7.5587	5.0415	179
Shortstops	7.4213	5.4150	413
Outfielders	6.7415	5.1532	557
Middle Infielders	7.2620	5.2196	519
Corner Infielders	7.3650	5.3715	359

Table 3.14: MLB Career Duration of Drafted Players by Position and Class. ('66-'97)

Player Type	Classification	Career Duration	Standard Deviation	# of Players
All Players	High School	7.2638	5.4051	1327
	4-Year College	6.0211	4.7439	1937
	JUCO	5.7870	4.5327	227
	No Classification	8.6250	5.7023	16
All Pitchers	High School	6.4855	5.0915	620
	4-Year College	5.5935	4.5292	984
	JUCO	5.3540	4.2682	161
	No Classification	6.7778	4.2361	9
All Position Players	High School	7.9463	5.5807	707
	4-Year College	6.4627	4.9191	953
	JUCO	6.3879	4.8306	116
	No Classification	11.0000	6.7578	7
Left-Handed Pitchers	High School	6.4607	5.4025	191
	4-Year College	6.0064	4.7715	314
	JUCO	5.5000	4.7324	44
	No Classification	6.7500	4.5735	4
Right-Handed Pitchers	High School	6.4965	4.9533	429
	4-Year College	5.4000	4.4015	670
	JUCO	5.2991	4.1006	117
	No Classification	6.8000	4.4944	5
Catchers	High School	7.8309	5.6484	136
	4-Year College	6.5714	4.7430	161
	JUCO	5.7200	4.2965	25
	No Classification	0.0000	0.0000	0
First Basemen	High School	7.9831	5.6705	59
	4-Year College	6.7727	5.6514	110
	JUCO	5.6000	4.9933	10
	No Classification	19.000	0.0000	1
Second Basemen	High School	8.2381	5.1760	21
	4-Year College	6.1282	4.0879	78
	JUCO	7.5714	3.5989	7
	No Classification	0.0000	0.0000	0
Third Basemen	High School	8.4789	5.3502	71
	4-Year College	6.8333	4.8545	102
	JUCO	9.0000	1.6733	6
	No Classification	0.0000	0.0000	0
Shortstops	High School	8.2660	5.6921	188
	4-Year College	6.5714	4.9499	189
	JUCO	7.0882	5.6481	34
	No Classification	14.0000	1.4142	2
Outfielders	High School	7.4279	5.3672	222
	4-Year College	6.3367	4.9642	297
	JUCO	5.7059	4.8338	34
	No Classification	7.5000	6.8557	4

Table 3.15: MLB Career Duration of Drafted Players by Round and Class. (‘66-‘97)

Round	High School	4-Year College	JUCO	Round Average
1	8.7974	8.6386	8.5000	8.7148
2	6.8793	6.8772	3.1429	6.8040
3	7.8319	6.5940	8.0000	7.2176
4	6.5900	5.6019	5.6667	6.0654
5	6.4615	6.5304	4.6667	6.4300
6	6.9683	5.8485	6.2727	6.2832
7	6.0200	5.3429	5.1818	5.5361
8	8.9174	5.4265	4.5714	6.6250
9	6.8333	5.4400	5.3333	5.8500
10	7.7059	5.0000	6.2222	5.9364
11	6.4571	6.2188	5.5714	6.2547
12	3.7059	5.3051	4.6250	5.0000
13	6.1538	5.1628	9.7500	6.0380
14	7.0000	4.0556	5.0000	4.9836
15	7.1250	4.3125	7.1333	5.8125
16	5.5455	5.0256	2.6667	5.1159
17	7.3125	6.5714	6.4444	6.7500
18	4.2667	6.2514	8.2000	5.8909
19	5.7333	4.5714	4.7778	4.9423
20	6.7500	4.9032	3.9091	5.3500
21	5.6250	4.5000	6.6667	4.9400
22	7.8333	4.8261	2.5000	5.5128
23	7.1111	5.4545	8.2000	6.4878
24	4.8750	4.3600	4.0000	4.3250
25	7.3000	5.8636	6.2500	6.3056
26	7.4286	4.2273	7.6667	5.2500
27	4.8750	3.3571	4.0000	3.9200
28	10.5000	4.2857	4.5714	4.9130
29	9.2500	2.7273	3.7500	5.1739
30	5.3333	4.2353	7.1667	5.0690
31	6.6667	4.3000	6.2000	5.2222
32	6.2500	2.6364	3.6667	3.6111
33	3.5000	3.9231	3.5000	3.7619
34	6.6667	4.0000	12.0000	5.4545
35	6.6667	3.0000	4.0000	4.7500
36	6.4000	4.4000	9.4000	6.7333
37	7.0000	5.2857	2.5000	5.3846
38	2.5000	5.0000	5.0000	4.6154
39	7.3333	6.3333	0.0000	6.6667
40	1.5000	3.6000	2.0000	2.8750
41	0.0000	2.5000	4.0000	2.8000
42	14.5000	5.0000	4.0000	7.2500
43	0.0000	9.3333	8.5000	8.8571
44	7.0000	2.0000	6.0000	5.7778
45	2.0000	8.5000	2.0000	4.6000
46	0.0000	0.0000	7.0000	7.0000
47	7.7500	2.6667	0.0000	5.5714
48	8.2500	0.0000	0.0000	8.2500
49	7.5000	0.0000	0.0000	7.5000
50	3.0000	3.7500	0.0000	3.6000

Table 3.16: Determinants of MLB Career Duration (‘66-‘97)

Variable	Hazard Ratio	Standard Error	z-statistic	p-value
NOCLASSIFICATION	0.6623	0.1794	-1.52	0.128
HIGHSCHOOL	baseline	baseline	baseline	baseline
JUCO	1.1389	0.0843	1.76	0.079
4YEARCOLLEGE	1.3557	0.0563	7.33	0.000
ROUND	0.9972	0.0031	-0.89	0.376
ROUNDPICK	1.0036	0.0022	1.63	0.104
RHP	baseline	baseline	baseline	baseline
LHP	0.9465	0.0960	-0.54	0.588
C	0.7832	0.0544	-3.52	0.000
1B	0.7688	0.0762	-2.65	0.008
3B	0.7456	0.0635	-3.45	0.001
2B	0.7497	0.0833	-2.59	0.010
SS	0.7280	0.0469	-4.93	0.000
OF	0.7934	0.0553	-3.32	0.001
COMPORSUPP	1.2632	0.1217	2.43	0.015
REDRAFTED	1.4464	0.1471	3.63	0.000
THROWSRIGHT	baseline	baseline	baseline	baseline
THROWSLEFT	1.0401	0.0978	0.42	0.676
BATSRIGHT	baseline	baseline	baseline	baseline
BATSLEFT	1.0815	0.0698	1.21	0.224
BATSSWITCH	1.2186	0.1054	2.29	0.022
HEIGHT	1.0025	0.0113	0.22	0.828
WEIGHT	0.9965	0.0014	-2.46	0.014
1STSTAGE	0.2571	0.0716	-4.88	0.000

Observations: 3557

Number of Failures: 3389

Likelihood Ratio $\chi 2$: 199.61

Table 3.17: Determinants of MLB Career Duration (‘66-‘75)

Variable	Hazard Ratio	Standard Error	z-statistic	p-value
NOCLASSIFICATION	0.3294	0.3356	-1.09	0.276
HIGHSCHOOL	baseline	baseline	baseline	baseline
JUCO	1.3341	0.2618	1.47	0.142
4YEARCOLLEGE	1.3426	0.1241	3.19	0.001
ROUND	0.9979	0.0063	-0.33	0.738
ROUNDPICK	0.9992	0.0058	-0.14	0.891
RHP	baseline	baseline	baseline	baseline
LHP	0.9799	0.2103	-0.09	0.925
C	0.7510	0.1075	-2.00	0.045
1B	0.8844	0.1926	-0.56	0.573
3B	0.7612	0.1361	-1.53	0.127
2B	0.7829	0.1988	-0.96	0.335
SS	0.7804	0.1043	-1.86	0.063
OF	0.6970	0.1049	-2.40	0.017
COMPORSUPP	x	x	x	x
REDRAFTED	1.9655	0.7000	1.90	0.058
THROWSRIGHT	baseline	baseline	baseline	baseline
THROWSLEFT	1.0887	0.2251	0.41	0.681
BATSRIGHT	baseline	baseline	baseline	baseline
BATSLEFT	1.0006	0.1481	0.00	0.997
BATSSWITCH	1.3359	0.2762	1.40	0.161
HEIGHT	0.9737	0.0270	-0.96	0.337
WEIGHT	0.9946	0.0041	-1.33	0.182
1STSTAGE	0.2855	0.1748	-2.05	0.041

Observations: 744
Number of Failures: 744
Likelihood Ratio $\chi 2$: 40.88

Table 3.18: Determinants of MLB Career Duration (‘76-‘85)

Variable	Hazard Ratio	Standard Error	z-statistic	p-value
NOCLASSIFICATION	1.0994	0.4579	0.23	0.820
HIGHSCHOOL	baseline	baseline	baseline	baseline
JUCO	2.0578	0.5270	2.82	0.005
4YEARCOLLEGE	1.4020	0.1122	4.22	0.000
ROUND	1.0114	0.0075	1.54	0.124
ROUNDPICK	1.0031	0.0044	0.72	0.473
RHP	baseline	baseline	baseline	baseline
LHP	1.1311	0.2129	0.65	0.513
C	1.0158	0.1267	0.13	0.900
1B	0.9078	0.1552	-0.57	0.572
3B	0.7658	0.1334	-1.53	0.126
2B	0.8801	0.1712	-0.66	0.511
SS	0.8022	0.1072	-1.65	0.099
OF	0.9793	0.1380	-0.15	0.882
COMPORSUPP	1.1170	0.1764	0.70	0.483
REDRAFTED	1.0125	0.2710	0.05	0.963
THROWSRIGHT	baseline	baseline	baseline	baseline
THROWSLEFT	0.7087	0.1257	-1.94	0.052
BATSRIGHT	baseline	baseline	baseline	baseline
BATSLEFT	1.0124	0.1336	0.09	0.925
BATSSWITCH	1.0243	0.1805	0.14	0.892
HEIGHT	0.9756	0.0230	-1.05	0.295
WEIGHT	1.0012	0.0032	0.37	0.713
1STSTAGE	0.7425	0.3529	-0.63	0.531

Observations: 995

Number of Failures: 994

Likelihood Ratio $\chi 2$: 65.76

Table 3.19: Determinants of MLB Career Duration ('86-'95)

Variable	Hazard Ratio	Standard Error	z-statistic	p-value
NOCLASSIFICATION	0.8094	0.3375	-0.51	0.612
HIGHSCHOOL	baseline	baseline	baseline	baseline
JUCO	1.0584	0.1027	0.58	0.559
4YEARCOLLEGE	1.3006	0.0882	3.87	0.000
ROUND	1.0020	0.0038	0.54	0.589
ROUNDPICK	1.0082	0.0034	2.44	0.015
RHP	baseline	baseline	baseline	baseline
LHP	0.9711	0.1536	-0.19	0.853
C	0.6637	0.0740	-3.68	0.000
1B	0.6811	0.1109	-2.36	0.018
3B	0.7002	0.0912	-2.74	0.006
2B	0.6450	0.1166	-2.43	0.015
SS	0.6898	0.0678	-3.78	0.000
OF	0.7494	0.0826	-2.62	0.009
COMPORSUPP	1.0220	0.1294	0.17	0.864
REDRAFTED	1.8357	0.4403	2.53	0.011
THROWSRIGHT	baseline	baseline	baseline	baseline
THROWSLEFT	1.2159	0.1958	1.21	0.225
BATSRIGHT	baseline	baseline	baseline	baseline
BATSLEFT	1.0709	0.1152	0.64	0.524
BATSSWITCH	1.2607	0.1850	1.58	0.114
HEIGHT	1.0221	0.0176	1.27	0.204
WEIGHT	0.9929	0.0020	-3.55	0.000
1STSTAGE	0.3199	0.1351	-2.70	0.007
Observations: 1526				
Number of Failures: 1438				
Likelihood Ratio $\chi 2$: 115.82				

Table 3.20: Player Age at Time of Draft; Players Reaching MLB Only (‘66-‘97)

	Average Age	Standard Deviation	# of Players
No Classification	20.06	2.1438	16
High School	17.78	0.6097	1327
4-Year College	21.03	0.7173	1937
JUCO	19.27	0.9463	277
Overall Average	19.68	1.6925	3557

Table 3.21: Predicted Probs. and Career Duration from Artificial Player Profiles

Player Profile	Predicted Probability of Reaching MLB
High School, LHP, 1st Round, 1st Pick in Round, East South Central, Ave. Height and Weight	0.2946
4-Year College, LHP, 1st Round, 1st Pick in Round, East South Central, Ave. Height and Weight	0.2777
High School, LHP, 10th Round, 1st Pick in Round, East South Central, Ave. Height and Weight	0.1610
4-Year College, LHP, 10th Round, 1st Pick in Round, East South Central, Ave. Height and Weight	0.1501
4-Year College, OF, 3rd Round, 1st Pick in Round, East South Central, Ave. Height, Weight = 205, Bats R, Throws R	0.2149
4-Year College, OF, 3rd Round, 1st Pick in Round, East South Central, Ave. Height, Weight = 180, Bats R, Throws R	0.1040
High School, SS, 2nd Round, 10th Pick in Round, East South Central, Height = 70, Ave. Weight, Bats L, Throws L	0.6450
4-Year College, SS, 2nd Round, 10th Pick in Round, East South Central, Height = 75, Ave. Weight, Bats L, Throws L	0.4635

Note: Average height = 73.61 inches; average weight = 195.17 pounds

Player Profile	Predicted Hazard Ratio
High School, LHP, 1st Round, 1st Pick in Round, East South Central, Ave. Height and Weight	0.4049
4-Year College, LHP, 1st Round, 1st Pick in Round, East South Central, Ave. Height and Weight	0.5490
High School, LHP, 10th Round, 1st Pick in Round, East South Central, Ave. Height and Weight	0.3950
4-Year College, LHP, 10th Round, 1st Pick in Round, East South Central, Ave. Height and Weight	0.5355
4-Year College, OF, 3rd Round, 1st Pick in Round, East South Central, Ave. Height, Weight = 205, Bats R, Throws R	0.4086
4-Year College, OF, 3rd Round, 1st Pick in Round, East South Central, Ave. Height, Weight = 180, Bats R, Throws R	0.4466
High School, SS, 2nd Round, 10th Pick in Round, East South Central, Height = 70, Ave. Weight, Bats L, Throws L	0.4485
4-Year College, SS, 2nd Round, 10th Pick in Round, East South Central, Height = 75, Ave. Weight, Bats L, Throws L	0.4540

Note 1: Each profile assumes player is in his 10th MLB season

Note 2: Average height = 73.54 inches; average weight = 192.57 pounds

Figure 3.1: Drafted Players by Training Classification and Round (1966-2005)

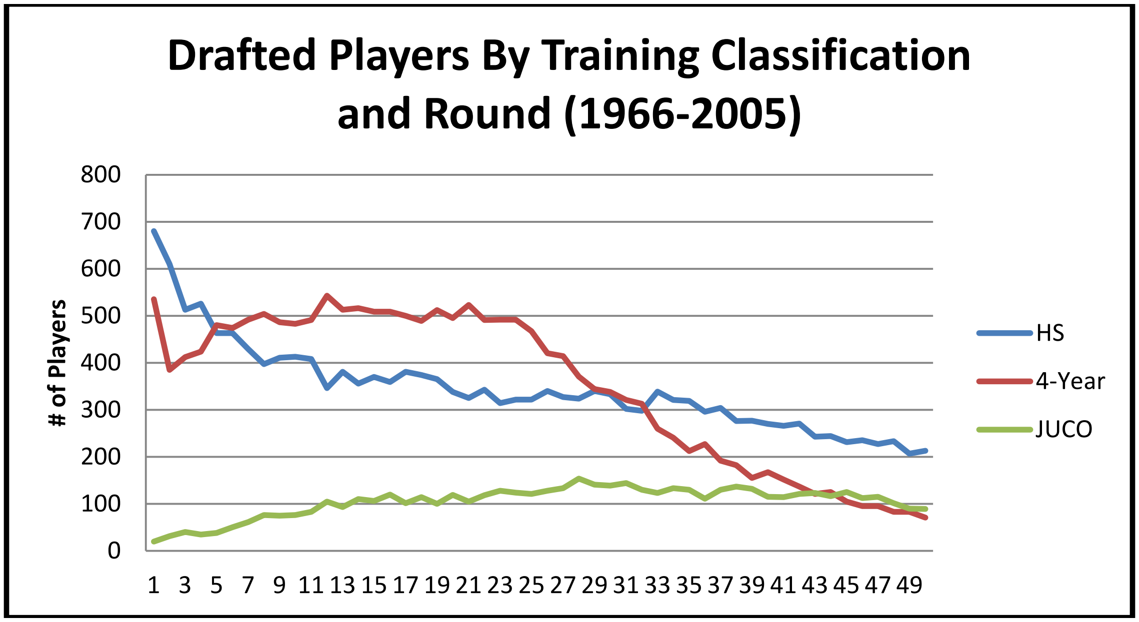


Figure 3.2: Drafted Players by Training Classification and Round (1966-1975)

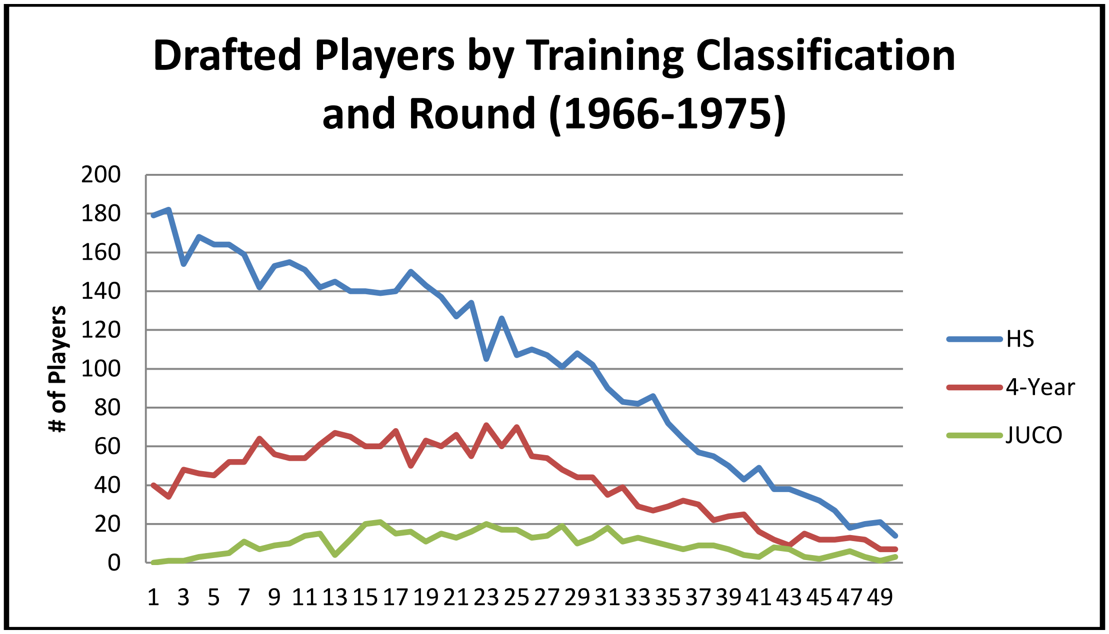


Figure 3.3: Drafted Players by Training Classification and Round (1976-1985)

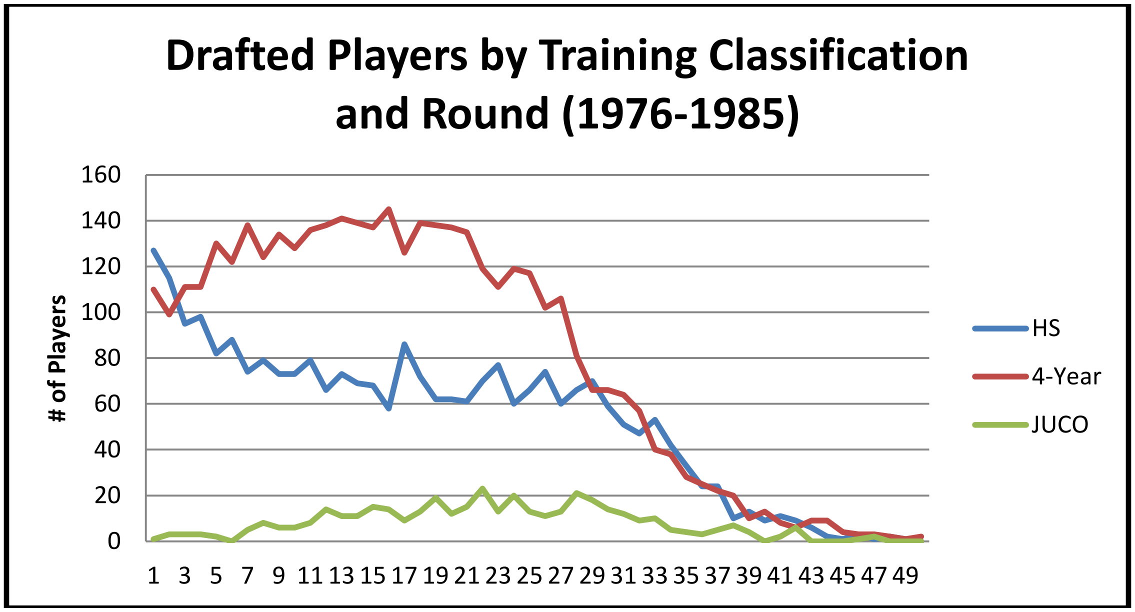


Figure 3.4: Drafted Players by Training Classification and Round (1986-1995)

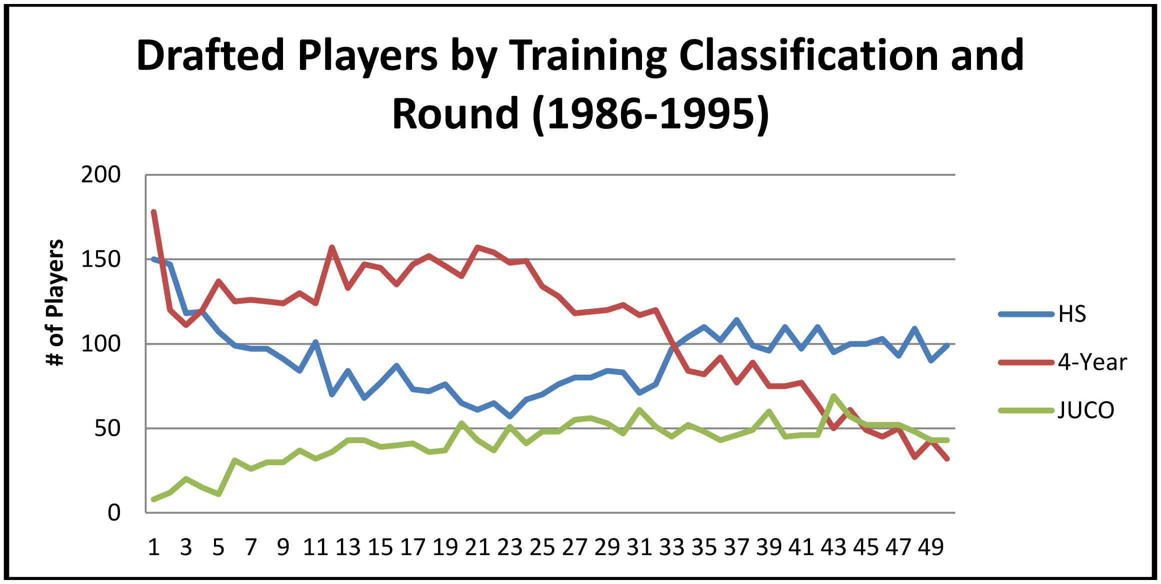


Figure 3.5: Drafted Players by Training Classification and Round (1996-2005)

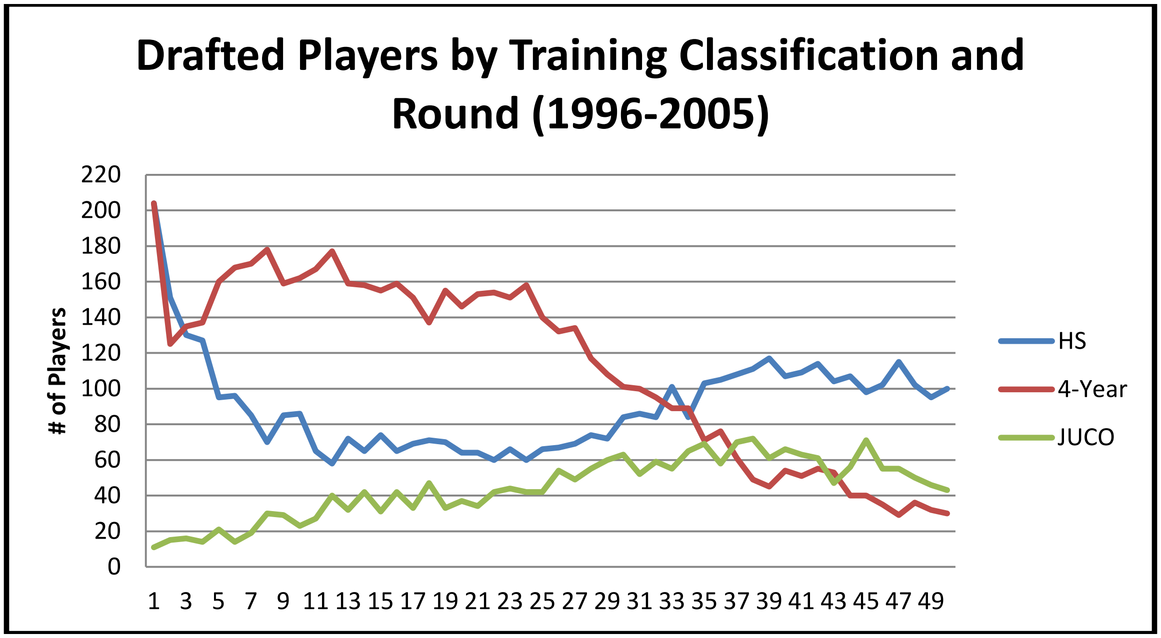


Figure 3.6: Differences in Prob. (4-Year College - High School) of Reaching MLB ('66-'05)

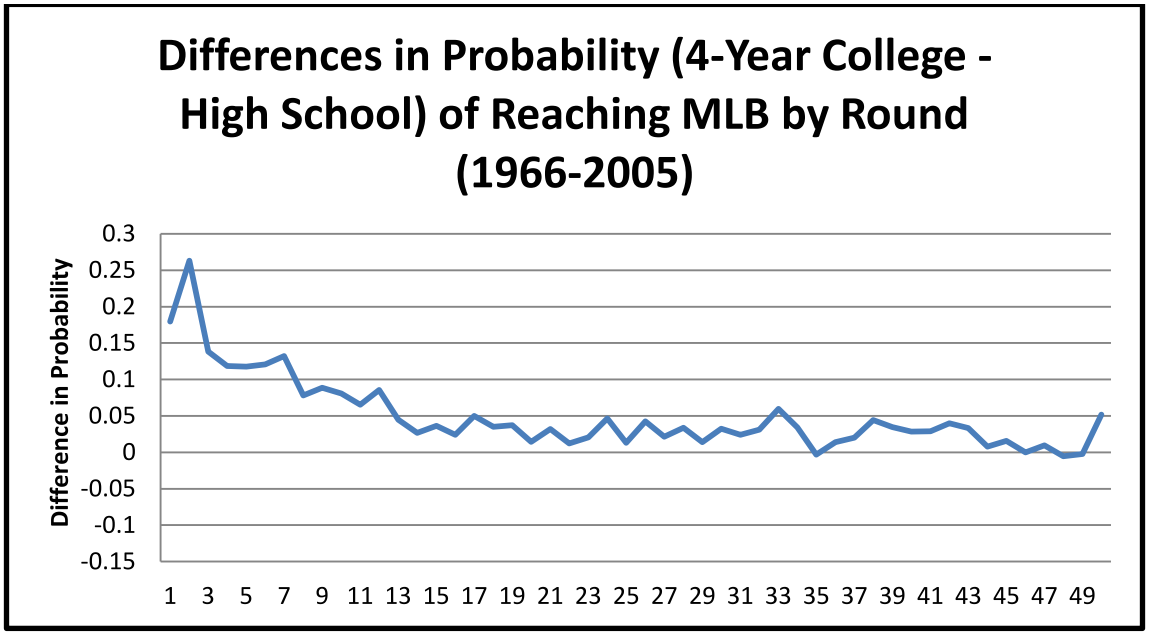


Figure 3.7: MLB Career Duration by Round Drafted and Training Classification

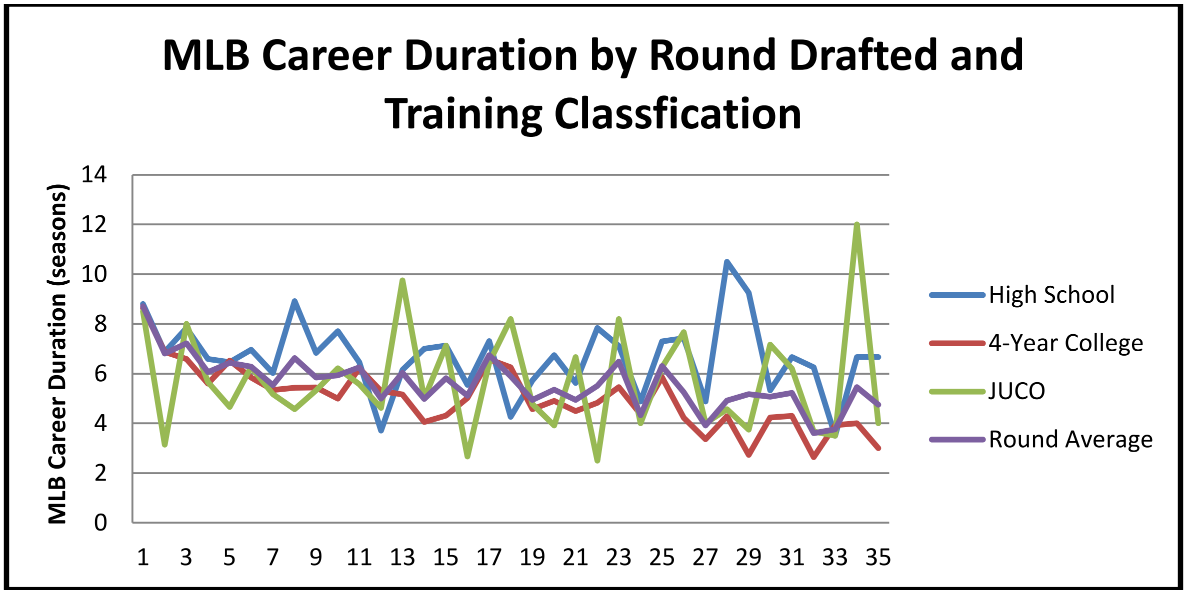


Figure 3.8: Est. Cumulative Hazard Function: Comparison of Training Class. (‘66-‘97)

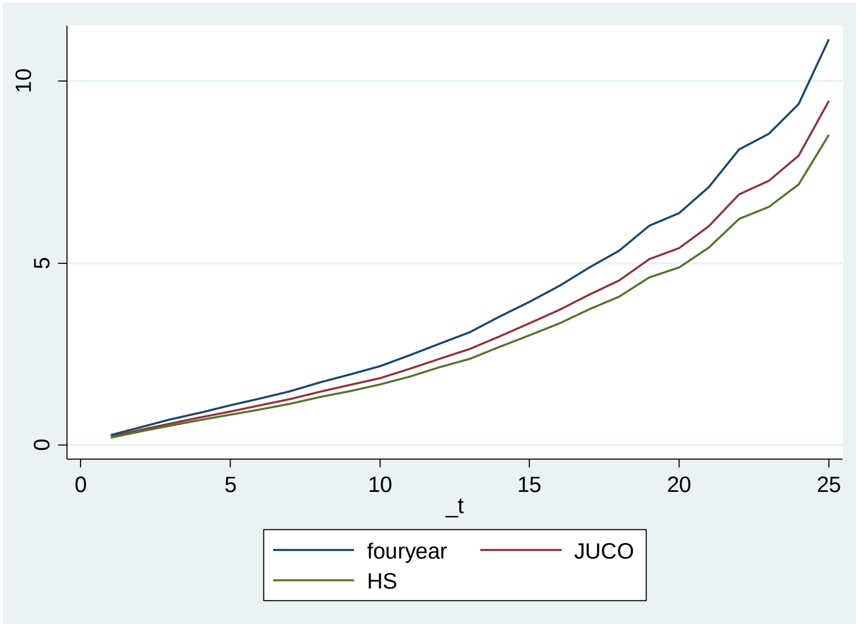
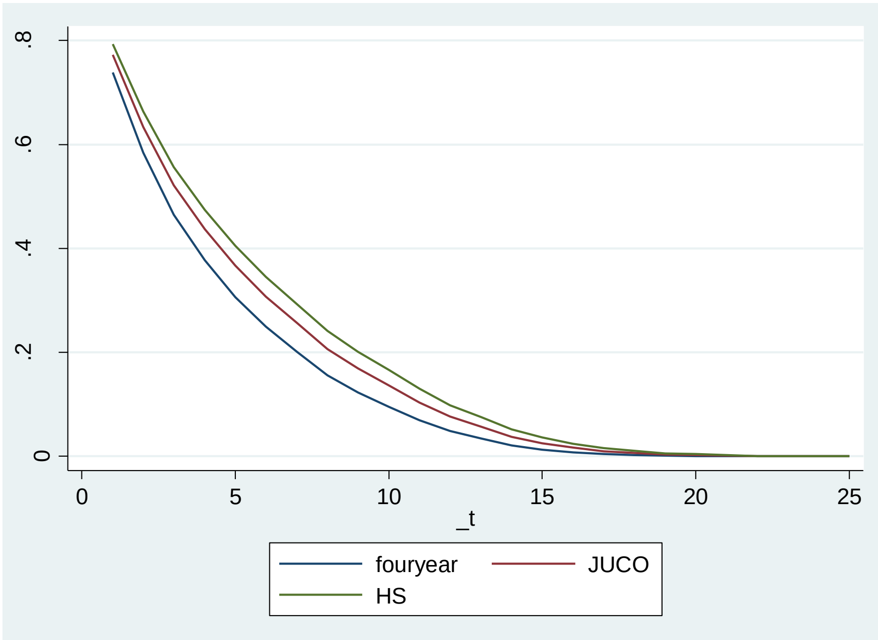


Figure 3.9: Est. Cumulative Survivor Function: Comparison of Training Class. (‘66-‘97)



CHAPTER 4

Competitive Balance in College Football

4.1 Introduction

Research on competitive balance is a centerpiece topic in the sports economics literature (Fort, 2006b). This chapter examines an ignored area in the competitive balance literature, namely, competitive balance in big time college sports, here confined to the conferences comprising the modern Football Bowl Subdivision (FBS, previously known as Division I-A for football)—the Atlantic Coast Conference (ACC), the Big 12, Big East, Big Ten, Pacific-12 (Pac-12), and Southeastern Conference (SEC). Balance is an important area in sports economics since the introduction of the uncertainty of outcome hypothesis (UOH) by Rottenberg (1956). Essentially, if fans like close contests and championship races, then runaways reduce the value to fans and to their ultimate servants, in this case, college football conference members. Thus, paying attention to balance is important because, if the hypothesis is verified, balance matters to fans and conferences.

If enhanced uncertainty does indeed increase live attendance, broadcasting revenues or ancillary revenues, then a conference has a vested interest in promoting competitive balance. Despite the fact that the majority of work measuring competitive balance has focused on North American professional sports and European football, analysis of balance in NCAA revenue generating sports is no less important. After all, the highest level of NCAA athletics is undoubtedly big business. The 14-year, $10.8 billion broadcasting contract the NCAA signed with CBS and Turner Broadcasting in April 2010 for the rights to the NCAA Basketball Tournament is a prime example (NCAA, 2010).

Furthermore, college football is of paramount importance because of the ability for an institution's football program to be the driver in generating revenues that support the entire athletic department. Hence, a degree of competitive balance is important to the bottom line of NCAA conferences and individual athletic departments.

Now, competitive balance can be tracked for its own sake or it can be an important element in the analysis of fan demand (Fort and Maxcy, 2003). This chapter involves simply tracking the behavior of balance over the lifespan of the six modern FBS conferences (while the analysis of fan demand in pro sports are legion, the only examples in college sports are Price and Sen, 2003, and DeSchriver and Jensen, 2006). Tracking serves the important functions of simply informing those interested about its behavior but also is useful in first-level assessment of the possible causes for changes in balance. In a sense, tracking balance also tracks changes in the FBS environment that might be predicted to change balance. However, tracking balance using sophisticated time series analysis can also help researchers to avoid some analytical pitfalls in their approaches to using competitive balance data.

The approach taken in this chapter is straightforward: review the literature and historical background, define the balance *concepts* of interest, define the *measures* of balance that capture these concepts, identify the data, specify the methodology that will be applied to the data, and finally, present the results discovered by that application.

Each section of this chapter follows through on this approach, culminating with conclusions and suggestions for future work.

Briefly, the work on FBS balance is episodic at best. The balance concepts involve the different levels where fans are interested in balance—at the level of close games, at the level of close season outcomes getting into the post-season, and at the level of uncertainty about champions across-seasons (dynasties). The measures involve the distribution of winning percentages and margin of victory (game level), years per championship (post-season level), and the correlation of year-to-year winning percentages (dynasties). The data on winning percentages and championships are readily available for every conference throughout their existence.

The methodology is found in the most recent work on tracking competitive balance (Lee and Fort, 2005, 2012; Fort and Lee, 2007). First, it is always informative to simply chart the data examine decade average behavior (Quirk and Fort, 1992). However, much more sophisticated time series approaches allow the researcher to examine the time series for unit root (stationarity) and structural break points. This gives the added advantages of 1) a more in-depth understanding of the behavior of the time series, 2) the avoidance of spurious correlations if data are analyzed across structural breaks, and 3) the chance to associate the history of events with the break points (or absence of a break point for some interesting historical events). This last allows one to gauge the effectiveness of competitive balance policy interventions and check on another of Rottenberg's (1956) hypotheses that has come to be called the Invariance Principle (IP).

The IP states that a change that simply reallocates the value created in an activity among the participants will not change the allocation of resources devoted to that activity. Rottenberg (1956) presented this idea relative to free agency in pro sports. Since free agency didn't exist at his writing, he hypothesized the following. It is in the best interests of the team owners in a pro league to allocate talent in order to maximize profit. Under the then prevailing reserve clause (prior to free agency, players were bound to the team that had their contract and could exercise no mobility at all), players moved to their highest valued use across the league and owners kept the value of the move through contract sales. If free agency were put in place, so that players could move at their discretion, then the same value of moves would not go to players instead of owners. The important observation, however, is that the players would still move and still to the highest valued use in the league (where players would still receive their highest return). Thus, reasoned Rottenberg, the distribution of the player talent resource would be invariant to the mechanism that distributed the reward from that distribution.

Moving to the FBS case, current NCAA restrictions bind players to the original athletic department where they signed their national letter of intent. This, along with the amateur requirement, guarantees that all of the revenue generated by players stays in the athletic department. FBS members care about the revenue generated so the allocation of players in the FBS occurs more-or-less to maximize that revenue subject to any other constraints on that behavior due to the interesting and unique nature of the relationship between athletics and universities. The IP would predict that, as long as FBS members still care about revenues, any change in the rules that would allow players to earn more of the revenue they generate would not alter the distribution of players.

Tracking competitive balance provides a simplistic ability to check this version of the IP. One policy intervention that changed the returns for players was the institution of grant-in-aid where athletes receive money to cover tuition, books, and room and board. The IP would predict that balance would be no different after the grant-in-aid policy went into effect than it was before that policy. Two other institutional changes in NCAA business structure relevant under the IP are the introduction of the G.I. Bill and the death of the College Football Association (CFA). The extension of Rottenberg's IP to college sports is something that has yet to be done in the academic literature.

This type of policy assessment from the full analysis of the behavior of the time series of competitive balance in the FBS has important implications. The utilization of time series techniques allows for the ability to make assessments on series stationarity and also measure the existence of structural breaks within a series. This permits the researcher to 1) match structural breaks with historical events which allows for an analysis of Rottenberg's IP and 2) assists in identifying break points so that spurious correlations can be avoided from analyzing data across structural breaks.

In regards to historic balance measurement, the data show that game uncertainty has increased over time but the distribution of team winning percentages and the year-to-year correlation between team winning percentages have remained relatively unchanged over history. However, the margin of victory in individual games relative to the degree of scoring has improved substantially, suggesting that game closeness has improved along with the advancement in the offensive sophistication of the sport. The work also shows that individual conferences have been controlled by a small, elite group of programs over the history of each conference. Overall, these results suggest that while individual games may have become more competitive over time the power programs have collectively ruled their given conferences over the sport's history. So despite the occasional short-term run of success from a non-powerhouse program, a granular look at the data shows that the usual suspects have controlled conference play over time regardless if the measure used is conference championships, conference winning percentage, or margin of victory.

In addition, time series techniques reveal mixed support for the IP in college football. Specifically, nine of fifteen competitive balance metric series that could be created are stationary without break points. The remaining six metric series are stationary with either one or two breaks. This result implies that competitive balance within FBS conferences has been relatively stable over time. A cursory look at conference championship history also illustrates that a small number of programs have dominated their given conferences and that the distribution of championship titles has not changed substantially over time. These results fail to reject Rottenberg's IP. On the other hand, three of the nine total break points identified match approximately with one of the three key events presented (GI Bill, grants-in-aid, and the demise of the CFA). In each case the identified break point is followed by a subsequent enhancement of balance. This provides evidence supporting the hypothesis that the institutional changes identified in this section are associated with structural changes in competitive balance. These findings reject Rottenberg's IP.

The final section contains a summary of the conclusions and a roadmap for the type of research indicated by the findings in this chapter.

4.2 Literature Review and Background

The sports economics literature contains many contributions focused on competitive balance (for examples highlighting the attention paid to balance please see Borland and Macdonald, 2003; Szymanski, 2003; Fort, 2006b). Despite the attention paid to the topic, the vast majority of the academic work in the area has focused on North American professional leagues and European football. While there are a few contributions dealing with the FBS, they have examined the short-term behavior of competitive balance. These works have measured balance in response to events viewed as key business changes at either the NCAA level or the individual conference level. Specifically, past research has measured the response of various forms of competitive balance to television deregulation (Bennett and Fizel, 1995), changes in scholarship limits (Sutter and Winkler, 2003), conference realignment (Quirk, 2004), and NCAA rules enforcement (Eckard, 1998; Depken II and Wilson, 2006). However, there has been no long-term examination of the behavior of competitive balance over the history of college football. This chapter addresses this hole in the literature by examining the behavior of competitive balance over the history of the FBS.

More precisely, there is almost no documented analysis regarding the overall behavior of balance in college football. Particularly, how has balanced behaved in general over time? Have individual games and conference races become more or less competitive? To what degree has the same team or a small number teams dominated a specific conference over history? Is there any correspondence between identifiable historical events and the behavior of competitive balance? On this last, the IP and an identifiable set of relevant events seem of particular interest.

Empirical tests of the IP are common in the sports economics literature, but previous analyses have focused solely on professional team sports leagues. Traditionally, these tests have measured changes in competitive balance in response to alterations in the access to playing talent (reserve clause versus free agency; the amateur player draft), various forms of revenue sharing among teams in a given league, and cross-subsidization mechanisms (salary caps and luxury taxes).[18] The findings in these works are mixed regarding whether the IP holds in response to changes in the institutional configuration of a league.[19]

In order for analysis of the invariance proposition to be appropriate, asymmetries in team quality must be evident (Szymanski, 2003) and the FBS clearly fits this requirement. Programs are located in different geographic locations that have long-standing allegiance to different conferences, a sure prescription for differences in the ability to acquire talent that subsequently determines team quality. The college football research on competitive balance cited above leaves a pair of significant episodes "lying on the sidewalk" as tests of the IP. The first is the immediate post-World War II period with the introduction of the G.I. Bill in 1946 and the formal imposition of athletic grant-in-aid in 1956 by the NCAA. The second is the demise of the CFA. Each of these represents settings appropriate for empirical tests of the IP in college football.

[18]Theoretical analyses of league policies on competitive balance also are extensive in the sports economics literature. See El-Hodiri & Quirk (1971), Quirk & El-Hodiri (1974), Quirk & Fort (1992), Fort & Quirk (1995), Vrooman (1995), Marburger (1997), Kesenne (2000; 2005), Szymanski (2004), and Szymanski & Kesenne (2004).

[19] Daly and Moore (1981) invoked information asymmetry and their empirical work found that free agency and decreased balance by strengthening larger-revenue market teams, counter to the IP. Fort and Quirk (1995) found evidence consistent with the IP for free agency, championship balance, the salary cap in the NBA, and (less convincingly to the authors) the draft in MLB and the NFL.

4.2.1 Episode 1: The G.I. Bill, the Sanity Code, 12-Point Code, and Grants-in-Aid

The first episode is a college parallel to "free agency" and player pay involving the immediate post-World War II rearrangement of revenues toward players. World War II caused an exodus of playing talent out of college football following the attack on Pearl Harbor in December of 1941 (Reimann, 2004). Before the end of 1943, Congress had required males between the ages of 18 and 45 to register with the Selective Service. Shortly thereafter, the draft was put into effect and 18 to 20 year old males were forced into service. Military obligations were mandated to last the duration of the War plus an additional six months (Moskos, 1988). This chain of events caused the widespread exit of both varsity level high school and college playing talent away from athletic competition and into military service. The impact on college football programs was substantial—39% of universities that fielded a major college football program ceased operations for at least one year during the War. At the lower levels of college football, a massive 82% of institutions halted programs between 1943 and 1946 (Boda & Claasen, 1961).

Following the end of the War, military veterans and a class of high school talent all flocked to college football in a single year, 1946, largely due to the enactment of the G.I. Bill (Andrews, 1984; Reimann, 2004). The G.I. Bill entitled any veteran with ninety or more days of service time to one year of college education. Each additional month of active duty service time netted an additional month of schooling, with a maximum of 48 months. The Bill paid up to $500 per year in tuition, fees and supplies, an amount exceeding the cost of the most expensive institutions at the time. The G.I. Bill also granted single veterans a stipend of $50 per month and married veterans $75 (Haydock, 1996).

With the war in the rear view mirror and the G.I. Bill in place, veterans returned to U.S. institutions of higher education *en masse*. With both the 1942 senior class of high school athletes and those who were already playing college football at the onset of the War returning home, the nature of college recruiting changed significantly. In conjunction with the returning influx of playing talent, the lenient restrictions associated with the G.I. Bill sparked a recruiting rampage. G.I. Bill regulation allowed those veterans who played only one year of college football prior to the War to attend any institution of their choosing upon return *without losing any eligibility*. Major college football programs which had previously suspended play due to a lack of numbers now not only had enough athletes to field competitive teams, but fierce recruiting battles for their services ensued (Andrews, 1984; Reimann, 2004). This limited type of "free agency" definitely sent more of the revenue that they generated to athletes than before the G.I. Bill.

The nature of college football recruiting changed dramatically following the introduction of the G.I. Bill. Prior to the War, recruiting was primarily a regional activity. The combination of increased access to air travel, enhanced recruiting budgets and larger bowl payouts following the troops' return elevated the process to the national stage (Falla, 1981; Reimann, 2004; Byers, 2005). Institutions were aware that their ability to attract talent and subsequently field competitive football programs could enhance their national recognition. This scenario resulted in some returning servicemen "selling" their services to the highest bidder. Thus, football players enjoyed a bit of “free agency” that did not exist prior to the passage of the GI Bill, precisely the setting for a test of the IP.

While exact details of the arrangements are not clear, the perception is that the offers were very generous for that period of time (Andrews, 1984). The nature of the recruiting business became so intense that the NCAA established the “Sanity Code” in 1948 in an attempt to regulate student-athlete compensation. The Sanity Code did not succeed due to enforcement issues and was abandoned only three years later in 1951 (Falla, 1981, p.134). While it is clear that the nature of college football recruiting changed dramatically following the conclusion of World War II, its roots can clearly be traced back to the introduction of the G.I. Bill. For the purposes of this chapter, relative to the operation of college programs prior to its passage, athletes received a larger share of the revenues they produced for their football program with the passage of the GI Bill. The IP would predict no change in competitive balance with this limited form of “free agency” occurring after the imposition of the GI Bill.

It is commonly noted that, after the demise of the Sanity Code, the passing of the NCAA “12-Point Code” in 1952 was the turning point in NCAA regulation following the failure of the Sanity Code (Falla, 1981; Eckerd, 1998). Included in this new legislation were two items focused on student-athlete compensation. The first was point number seven in the 12-Point Code, which was to “limit the number and amount of financial grants to athletes.” The second was an excerpt in a section titled “Principle Governing Financial Aid” which stated, “any athlete who receives financial assistance other than that administered by his institution, shall not be eligible for intercollegiate athletic competition” (Falla, 1981, pp. 135-136). This legislation passed by the NCAA did prohibit outside entities from providing financial assistance to athletes, but it does not set specific limits on financial aid or compensation that can be provided to an athlete by their institution. Therefore, without specific compensation limits set and enforced by the NCAA, it is unreasonable to assume that programs across the country would uniformly be providing compensation packages of equal value to athletes. Based on the specifics of the 12-Point Code, this 1952 change in NCAA regulation is not considered to be the most appropriate measurement point to test the invariance proposition.

Instead, the effectual turning point in the regulation of NCAA athlete compensation occurred in 1956, with the formal adoption of athletic grant-in-aid (Byers, 1995). This ruling established guidelines for student-athlete compensation across the NCAA and ended a roughly 30-year period of either non-regulation or unsuccessful enforcement of regulations where significant variation in compensation among institutions was the norm. This policy change was ratified at the 1956 NCAA Convention. The grant-in-aid program allowed for institutions to compensate undergraduate athletes regardless of their financial need or academic potential. It provided them with “commonly accepted educational expenses,” which included tuition, fees, room and board, books and $15 per month for laundry. Grants were provided for a maximum of four years and could not be annulled even if an athlete decided to remove himself from the athletic program. The goal was to provide athletes with only what they would need in order to bring compensation back to levels appropriate with amateur status (Byers, 1995).

Establishment of the grant-in-aid program was due largely in part to the explosion of lucrative offers made to athletes following World War II and the failure of the Sanity

Code. Officials from the Southern, Southeastern and Southwest conferences lobbied for the new system, while the traditional football powers in the Ivy League and Big 10 Conference supported the status quo. Supporters of the grant-in-aid program, consisting largely of the southern schools and traditional non-powers believed that the shift would level the playing field in terms of the ability to recruit talent (Byers, 1995). Under the new athletic grant system, institutions would be able to provide compensation only up to the levels set forth in NCAA bylaws. The previous structure was classified by erratic levels of athlete compensation based largely on an institution's desire to produce a quality football program and on the levels of booster and alumni contributions (Andrews, 1984; Byers, 1995). The shift to the grant-in-aid program did however eliminate direct payments to athletes and their parents by athletic boosters and alumni. Instead, the new athletic grant system resulted in an arrangement where boosters paid the institution directly and in turn those contributions were used to fund athletic grants (Byers, 1995).

Previous literature marks the 1952 21-Point code as the point in time where large scale enforcement of NCAA regulations actually began to materialize (Falla, 1981; Eckerd, 1998). While the historical documents largely appear to support this stance, further evidence points to the 1956 introduction of athletic grant-in-aid as the event marking a tangible shift in the manner in which student-athletes were compensated. The 1952 legislation was the catalyst leading to more stringent enforcement of regulations following years of ineffective regulation. However, it is clear based on the account of Walter Byers, NCAA Executive Director from 1951 to 1988, that the establishment of athletic grant-in-aid in 1956 was the event that normalized compensation and largely eliminated direct payments to student-athletes (Byers, 1995).

In this chapter, the 1956 formal adoption of athletic grant-in-aid will be used as the event that formally standardized student-athlete compensation. As with the GI Bill, the institution of grants-in-aid altered the share of revenue generated by athletes that they were able to keep for themselves. Alterations in student-athlete “pay” of this variety also make grants-in-aid appropriate for testing of the IP.

In summary, the immediate post-war period offers two test-points for the IP. The first is the period prior to the start of World War II and the passage of the G.I. Bill up to 1941 and an the period from 1946 onward. The second is an amount of time prior to the imposition of grants-in-aid up to 1956 and the period after, 1957 onward.

4.2.2 Episode 2: The Demise of the CFA

Changes in league revenue sharing arrangements have also been used as tests of the IP in the professional team sports league literature. An important parallel occurred in the NCAA in the mid-1990s with a significant change in the way television broadcast revenue was collected and distributed in big time college football. This shift transpired following the 1995 season with the conclusion of the CFA national television contract. Following this event, the television broadcasting model shifted from a single contract dominated by the CFA to the current characterization where each individual conference negotiates their own deals and shares those revenues with member institutions.

The introduction and growth of television significantly changed the revenue collection structure of NCAA athletic departments. The first televised intercollegiate football game was a 1940 matchup between Pennsylvania and Maryland from Franklin Field in Philadelphia. Once broadcasts of live games became commonplace in the late

1940s it was evident that this new form of entertainment was a legitimate threat to the gate revenues of institutions and was a strong substitute for game-day attendance. By 1950, unrestricted live broadcasting of college football games was the norm and attendance figures in many locations declined significantly. In fact, some colleges saw a 25% year-over-year decrease in attendance in 1950. In 1952, the NCAA unveiled a plan limiting live broadcasts to one national game each Saturday afternoon with each college permitted only one television appearance per season. Ninety-two percent of NCAA member institutions voted in favor of the proposal. Following its ratification, the NCAA sold the rights to televise the twelve games of the 1952 season to NBC for $1.14 million (Falla, 1981, p. 104-108). Member institutions now had a legitimate revenue source to compliment gate receipts.

From 1952 to 1984, the NCAA controlled the live broadcasting of college football games. Throughout this period, the typical revenue distribution arrangement consisted of the NCAA keeping between three and eight percent of the rights fees with the remaining amount collected by the programs featured on television each week. Over time, the NCAA allowed the broadcast of regional games, increasing the number of programs appearing on television. In 1968, the NCAA loosened its restrictions and allowed teams to appear live twice per season as part of a “wild-card” game selected by the network partner. In 1977, as part of a new four-year contract signed with ABC, which paid the NCAA an average of $30 million per year, teams were allowed to appear up to three times per season (Falla, 1981, pp. 111-119). While the number of teams appearing on television did increase over time, the most attractive programs were selected for broadcast by the networks more regularly and subsequently reaped the majority of the financial benefits of the broadcasting plan.

The College Football Association was formed in 1976, with 62 of the major college football programs, excluding the Big Ten Conference and Pacific Ten Conference members, joining the organization. With the NCAA firmly entrenched as the single entity controlling college football television rights, the CFA was established largely to gain influence over the broadcasting process. In 1981, the CFA negotiated a separate television deal with NBC that provided more exposure and was more lucrative than the deal constructed by the NCAA (Greenspan, 1988). But the NCAA threatened CFA members with severe penalties including expulsion from the NCAA, exile from participation in the NCAA Men's Basketball Tournament and elimination of Bowl Game affiliations if any institution were to sign the broadcasting deal with NBC. In response, the CFA collectively declined to enter into contract with the network (Siegfried & Burba, 2004).

By the early 1980s, NCAA control over college football telecasts was in the midst of a serious challenge by some of college football's most successful programs. Despite reaping the majority of the financial benefits of the existing NCAA system, the big-time programs were fighting for additional television exposure, which was being artificially restricted by the NCAA. The institutions claimed that the NCAA was restraining trade, which in turn limited the ability for athletic departments to generate revenues from additional television appearances. In June of 1984, the dispute was settled in Supreme Court in *The Board of Regents of the University of Oklahoma, et al. v. the NCAA*. It was determined that the NCAA was in violation of the Sherman Act and the ruling granted individual institutions the right to negotiate their own broadcast deals (Greenspan, 1988). The verdict also resulted in the voidance of the NCAAs existing television contracts, worth $280 million (Siegfried & Burba, 2004).

Following the Supreme Court's ruling, the number of televised college games increased dramatically as the CFA, several conferences, and individual institutions all began signing broadcasting deals.[20] Due to the increase in the supply of games available for purchase by the networks and cable stations, rights fees decreased dramatically (Fort and Quirk, 1999). The short-term result was that the majority of big-time programs generated less revenue under the new unrestricted arrangement as compared to the regulated NCAA plan. Additionally, many smaller Division I-A (now FBS) programs that were previously televised under the NCAAs regional broadcasting plan found them untelevised and generating less television revenue under the new system (Greenspan, 1988). The impact was dramatic as 1984 cumulative television broadcasting revenue collected by NCAA programs was $31 million. Meanwhile, the original NCAA package would have generated $74 million for member institutions (Siegfried & Burba, 2004).

More importantly, the fallout from the Supreme Court ruling marked the emergence of the CFA as the leading entity in television negotiations.[21] The CFA negotiated four separate deals with multiple network and cable partners from 1984 to 1995. Under the CFA contracts, member schools collected revenues based on two factors. First, each program received a direct payment for being a CFA member. Over time, the per-program amount of this participation payment grew from $50,000 to $150,000 per

[20] The Big Ten and (then) Pac-10 (now Pac-12) were not a part of the CFA and were the only large conferences with television broadcasting deals.

[21] The Big Ten and Pac-10 conferences remained unaffiliated with the CFA and negotiated separate television deals.

year. Revenue was also distributed based on number of television appearances. Programs were paid on a per-appearance basis and these payouts consisted of 75% to 80% of the total contract amounts. The CFA deals allowed for programs to appear on television more frequently as compared to the NCAA reign. Naturally, the strongest programs were again the financial beneficiaries of this decision, but some conferences did share appearance revenues between their members (Siegfried & Burba, 2004).

Despite the backing of the majority of the most successful programs in the nation, the CFA began to weaken in 1990. The first large scale blow to the CFA came early that year when Notre Dame, the independent member of the CFA with the strongest national appeal, left the organization to sign a four-year, $38 million broadcasting deal with NBC.[22] The fallout from the move arrived when ABC mandated a $25 million reduction to the CFA in their upcoming broadcasting deal (Sandomir, 1991). This was a substantial setback, but it was the events of 1995 that put the nail in the coffin of the CFA. With the current CFA television deal expiring at the end of the year, CBS made an aggressive move to acquire the rights to SEC football. Despite previous overtures from ABC in the late 1980s that the conference declined, the SEC decided to withdraw from the CFA and accept the five-year, $85 million offer from CBS. With both Notre Dame and the SEC gone from the CFA, the organization’s collective bargaining power was virtually eliminated and the CFAs role in negotiating broadcasting deals ended with the conclusion of the 1995 contract (Siegfried & Burba, 2004).

The demise of the CFA marked the shift from a single entity collectively negotiating television deals to the current characterization where conferences individually

[22] At the time, a large and successful contingent of independent programs including Miami (FL), Florida State, Penn State, Louisville, Virginia Tech, Syracuse, South Carolina, Boston College, West Virginia and Pittsburgh were all CFA members

make contracts. Following the conclusion of the 1995 CFA television deal, 1996 marked a massive reorganization of partnerships between conferences and the major networks and cable companies. In addition to the SEC contract, CBS also acquired the rights to Big East conference contests.[23] The ACC followed by signing multiple five-year deals with both ABC and ESPN for the 1996-2000 seasons. The ABC deal paid the ACC approximately $50 million while the ESPN deal grossed the conference $30 million (Associated Press, 1994). The Big 12 inked an eight-year deal with ABC and Liberty/Prime Sports and a secondary broadcasting deal with Fox Sports (Brooks, 1995).[24] The Pac-10 also signed a deal with Fox beginning in the 1996 season. In addition to these new partnerships, the Pac-10 and Big Ten both had existing contracts with ABC that were arranged during the reign of the CFA.

Additionally, the demise of the CFA ushered in the three-tier broadcasting rights structure (Fort, 2006a). This system (now technically not the same with the advent of conference TV networks like the Big Ten Network) was characterized by a conference entering into multiple contracts with multiple media providers over the same period of time. The first-tier provider was traditionally a national over-the-air network who paid the largest rights fees and had the first choice of which conference game they will select for broadcast on a weekly basis. The second-tier was typically a national cable provider such as ESPN or Fox Sports who paid a smaller rights fee compared to the network provider and had the second choice of which conference game they will show.

---

[23] The Big East Conference began play in football during the 1991 season with eight members (Boston College, Miami, Pittsburgh, Rutgers, Syracuse, Temple, Virginia Tech, West Virginia).

[24] The Big 12 Conference was formed on February 12, 1994 and football play began in 1996, joining all members of the Big 8 Conference (Colorado, Iowa State, Kansas, Kansas State, Missouri, Nebraska, Oklahoma, Oklahoma State) with 4 members of the Southwest Conference (Baylor, Texas, Texas A&M, Texas Tech).

Traditionally, the third tier was a regional broadcasting company who paid an even smaller rights fee to the conference in exchange for the third selection each week.

An example of this broadcasting structure is visible in the ACCs broadcasting partnerships for the 1996-2000 seasons. ABC and ESPN were the ACCs first- and second-tier partners respectively. The regional network Jefferson Pilot Sports served as the conference's third-tier partner (Associated Press, 1994).

The result of this shift away from a unified CFA television deal marked a significant change in the way revenue was distributed among college football's largest programs. The CFA deals consisted of membership and appearance payments that subsidized all members, but clearly benefitted the nation's most popular programs. Once the CFA's reign ended, the new system enabled conferences to distribute broadcasting revenues as each saw fit. This resulted in a discernible change from the CFA era where a single contract was negotiated and each member was bound to the membership and appearance payout structure collectively determined by the CFA. Under this revenue distribution structure, a maximum of 25% of the yearly television revenue was split evenly among programs as a payment for being a CFA member. Consequently, at least 75% of the television revenue collected by the CFA was distributed based on number of television appearances.[25] As expected, the more successful and traditional powerhouse programs were the financial beneficiaries of the CFAs payout scheme.

Naturally, each conference has not used the same exact formula in determining how to distribute revenues among its members. While exact terms of revenue sharing

[25] The percentage of revenues distributed based on CFA membership and number of television appearances changed over the course of the CFA's power. Percentages of total revenue paid out based on membership ranged from 20% to 25% and percentages paid out based on the yearly number of television appearances fluctuated between 75% and 80% (Siegfried & Burba, 2004).

arrangements are difficult to obtain, information does exist on how certain conferences choose to allocate revenues. Generally, these distributions are determined by a vote of university administrators or the executive committee of the conference and change over time. The 2009-2010 fiscal year television revenue sharing arrangements for the FBS power conferences are shown in Table 4.1. Four of the six conferences share football television revenues equally among member institutions.[26] The exceptions are the Big 12 and the (then) Pac-10 (now, Pac-12), which share a percentage of television revenues equally and then distribute the remaining funds based on the number of television appearances by each institution.[27] This type of arrangement benefits programs that are more appealing to network and cable partners, as more popular teams produce higher ratings, which allows broadcasters to charge higher advertising fees to firms.

So, as with the post-World War II episode, the demise of the CFA represents an alteration in the distribution of the value created, but none have any wish for the total to change in any way. In this situation, the IP again predicts no change in competitive balance after the demise of the CFA, from 1996 onward.

4.3 Competitive Balance Concepts

It is somewhat straightforward to characterize fans caring about three balance concepts—at the level of close games, at the level of close season outcomes getting into the post-season, and at the level of uncertainty about champions across-seasons

[26] Beginning in 2010, the Big East reorganized to a system where each football program shared football television revenues equally. Previous arrangements included both a membership payment and a payment based on television appearances. Percentages of total revenue paid out based on the number of television appearances changed over time (Furfari, 2010).

[27] In October of 2010, the Pac-10 Conference voted to share football television revenue equally among its 12 members beginning with the addition of Colorado and Utah to the league on July 1, 2011. Colorado and Utah will be subjected to a three-year period of reduced revenue sharing (Matuszewski, 2010).

(dynasties). These concepts reveal “competitive balance” as multi-faceted and therefore, no single metric is appropriate for measurement. Sloane (1976) introduced these concepts and over time his contributions have morphed into three distinct categories of competitive balance – game uncertainty (GU), playoff uncertainty (PU), and consecutive-season uncertainty (CSU) (Cairns, 1987). Empirical literature in the area has produced multiple metrics for each category.

Metrics assessing GU have examined the dispersion of team winning percentages and the degree of game closeness. Playoff uncertainty metrics are grouped into variables accounting for the extent to which teams are in contention for the playoffs and the degree of concentration of league championships. Measures of CSU are less developed and the existing metrics have attempted to capture the degree of variance in team quality over successive seasons. The metrics used here are detailed below.

4.4 Competitive Balance Measures

The measures involve the distribution of winning percentages and margin of victory (game uncertainty, or GU), years per championship (post-season uncertainty, or PU), and the correlation of year-to-year winning percentages (consecutive season uncertainty, or CSU). There are multiple competitive balance metrics and this chapter will utilize measures from each of the three categories – GU, PU and CSU. This will allow for a comprehensive view of college football balance over the history of the sport.

Measures of game uncertainty are divided into measures of winning percentage dispersion and game closeness (Fort, 2006b). This chapter will utilize the well-known ratio of standard deviations of winning percentages (RSD) to proxy for a measure of winning percentage dispersion. This metric is the ratio of the standard deviation of winning percentages in an actual league to the standard deviation of winning percentages in the idealized or perfectly balanced league. RSD was introduced by Noll (1988) and first utilized by Scully (1989). Let SD = standard deviation and WP = winning percentage. The RSD metric is defined below:

RSD = SD of WP in actual league / SD of WP in idealized league

The numerator of the RSD metric is calculated by taking the standard deviation of the year-end winning percentages of each team in a conference. For the denominator, an "idealized" league is one where the probability that any team beats any other is literally 0.5. The denominator is simply equal to $(.5)(\sqrt{N})$ (the derivation is in Fort and Quirk, 1995), where N is the number of conference games played in a season. Because the denominator accounts for changes in season length, the more games in a league schedule, the lower the idealized standard deviation (Quirk & Fort, 1992). The closer the RSD value is to one, the more balanced the league. Alternatively, a league with less balance will show larger positive RSD values.

Measures of game closeness are absent for the sport of American football despite their existence in other sports including baseball, hockey and world football (Fort, 2006b). Based on the need for a metric to be applicable over the entire history of the sport, a new measure was created here. In order to capture the changes in the level of uncertainty in individual games, a "Margin of Victory Ratio" (MVR) was developed. This metric is the ratio of the margin of victory in single contest to the total points scored in a single contest. Let MV = margin of victory and TPS = total points scored. MVR for a single game is defined below:

$$MVR = MV\ in\ a\ Contest\ /\ TPS\ in\ a\ Contest$$

This measure offers an advantage over simply tracking historical MV values over the history of a conference, as it accounts for the changes in total points scored over time. The number of conference games played in a season has also fluctuated over time. In order to account for these changes, a margin of victory ratio is calculated for every game in a conference season. Those values are summed and then divided by the total number of conference games in that particular season. Therefore, when n is the number of conference games played in a season, the MVR metric for conference i in year j is defined as:

$$MVR_{i,j} = \sum MVR / n$$

The benefit of using the metric specified is that MV is standardized despite significant changes in the levels of scoring over the history of the sport. For example, from 1900 to 1910 the average TPS in a single Big Ten conference game was 25.07. As the game evolved, so did offensive prowess and from 2000 to 2010 the average TPS in a Big Ten game increased to 51.49. The value of the MVR metric is bounded between zero and one, with tied contests, including 0-0 ties, producing a value of zero. Alternatively, shutouts generate a value of one. MVR closer to zero means more balance in terms of individual game uncertainty. MVR closer to one represent less uncertainty in individual contests.

In order to capture PU by way of championship concentration, we will utilize the years per championship (YPC) metric first employed by Quirk and Fort (1992). Let ECY = eligible conference years. The years per championship metric is detailed below:

$$YPC = ECY\ /\ Number\ of\ Conference\ Championships$$

This metric is beneficial as it provides a measure of the degree to which a single team or a small number of teams in a conference has dominated championships. However, the metric suffers because it provides a single value for a program over its entire conference playing history as opposed to an annual measure. Programs with lower values in the years per championship metric are teams that have regularly won the conference championship during their tenure. Larger values will be associated with programs that infrequently win their conference championship.

Lastly, the correlation of year-to-year winning percentages (WPC) for teams in a conference will be used to capture the CSU category of competitive balance. Let $WP_{i,t}$ = WP for team *i* in year *t*. WPC is defined as:

$$WPC = Correlation(WP_{i,t}, WP_{i,t-1})$$

This metric was originally utilized by Butler (1995) and Lee and Fort (2008) and is employed here to determine the degree of churning in the season-to-season conference standings. The WPC metric is bounded between -1 and 1 with -1 suggesting negative correlation, 0 suggesting no correlation, and 1 suggesting positive correlation. Lower WPC values suggest more churning in the conference standings and increased balance within a conference while higher values suggest less churning in conference positioning and therefore less balance.

4.5 The Data and Other Background

The data on winning percentages and championships are readily available for every conference throughout their existence. Quirk and Quirk (2012) provide complete end-of-season records for every major college football program since 1894. The rest of the required data including the game-level data needed to calculate the MVR metric were found at www.sports-reference.com (their college section). The data utilized to calculate the measures in the last section will cover the entire playing history of each conference. Each conference began play at a different point in time, with the Big Ten being the earliest conference to begin formal play in 1896. Table 4.2 shows the historical playing periods of each conference.

But collecting the data appropriately also requires a bit of understanding of the history of the evolution of the FBS conferences. Due to the amount of churning that has taken place in Division I conferences over the history of the sport, there is a need to clarify the specifics regarding the historical examination periods for each conference. Some conferences, such as the ACC, Big East and SEC have operated under the same name throughout their existence. Therefore, the examination period for each of these conferences is clear and begins at the time of original conference formation and continues through the completion of the 2010 season. Meanwhile, the Big Ten, Big 12 and Pac-10 conferences have operated under multiple names over their respective histories, so clarification is warranted.

The current Big Ten operated as the Western Conference from 1896 to 1952 and then adopted its current name in 1953. The Big 12 was formally established in 1996 when Texas, Texas A&M, Texas Tech and Baylor of the Southwest Conference merged with the Big Eight Conference. Prior to the formation of the Big 12, this analysis will follow the Big Eight beginning in with the 1928 season. The Big Eight operated as the Big Six Conference from 1928 to 1947 and the Big Seven Conference from 1948 to 1959. The conference became the Big Eight in 1960 until it merged into the Big 12 beginning with the 1996 season. Lastly, the Pac-10 began play in 1916 but the conference operated as the Pacific Coast Conference, the Athletic Association of Western Universities, and the Pacific-8 Conference prior to officially becoming the Pac-10 in 1978. For a detailed historical account of conference churning and patterns of conference stability in college football, please see Quirk (2004).

The raw data on records and championships were collected from two primary sources. A portion of the historical winning percentage and conference championship data are from Quirk and Quirk (2012). The remainder of the data was collected from the college football pages at www.sports-reference.com. The measures in the last section were collected or calculated from this raw data. In order to control for differences in out-of-conference scheduling between teams within a single conference, only conference games are included in the analysis.

4.6 Methodology

The methodology is found in the most recent work on tracking competitive balance (Lee and Fort, 2005, 2012; Fort and Lee, 2007). First, it is always informative to simply chart the data examine decade average behavior (Quirk and Fort, 1992). However, much more sophisticated time series approaches allow the researcher to examine the time series for unit root (stationarity) and structural break points. This gives the added advantages of 1) a more in-depth understanding of the behavior of the time series, 2) the avoidance of spurious correlations if data are analyzed across structural breaks, and 3) the chance to associate the history of events with the break points (or absence of a break point for some interesting historical events). This last allows one to gauge the effectiveness of competitive balance policy interventions and check on another of Rottenberg's hypotheses about policy intervention that has come to be called the Invariance Principle (IP).

The aggregation and organization of raw data along with the interpretation of summary statistics are adequate to generate a baseline understanding of the history of balance in a particular conference or sport. Subsequently, summary statistics are presented which outline the behavior of each identified competitive balance metric over the playing history of each conference.

More sophisticated time series techniques are utilized to evaluate the historical behavior of competitive balance over the history of each conference. The objective is to determine whether break points exist in the selected competitive balance metrics and whether or not these breaks match with the key historical events identified. This approach begins with a test of each competitive balance metric series against a null hypothesis of a unit-root. This is completed through the use of Augmented Dickey Fuller (ADF) and Philips-Perron (PP) unit-root tests. For any series with unit root, the second step is to assess series stationarity with break points accomplished through the use of a two-break minimum Lagrange Multiplier (LM) unit-root test and a one-break minimum Lagrange Multiplier unit-root test. A total of four tests are completed – two ADF tests and two PP tests – each specifying a constant only and then once again specifying both a constant and trend. The number of lags specified in the ADF tests is selected from the Swartz-Bayesian criterion and the number of lag truncations in the PP tests are selected from the work of Newey and West (1994).

Determining series stationarity, or stationarity with break points, is important for researchers who wish to perform regression analysis on a time series of data. If the series is found to be nonstationary, then biased estimates can result from utilizing traditional regression analysis. The usual approach is to take first differences of the data in the series to eliminate this bias.

Two-break minimum LM tests and one-break minimum LM tests are conducted to account for the possibility that breaks may occur near the ends of a given time series. The two break test is carried out first with the one-break test completed only if a series is not rejected at the highest critical level in the two-break test (Perron, 1989; Lee and Strazicich, 2001; 2003; 2004). The two-break test is illustrated as follows and is taken directly from Fort and Lee (2006):

> Based on the LM principle, a unit-root test statistic can be obtained from the following regression:
>
> $$\Delta y_t = d' \Delta Z_t + \phi S_{t-1} + \Sigma\gamma_i \Delta S_{t-1} + \varepsilon_{t.,}$$
>
> where Δ is the difference operator and S_t is a detrended series such that $S_t = y_t - \psi_x - Z_t\delta$, $t = 2,\ldots,T$. δ is a vector of coefficients in the regression of Δy_t on ΔZ_t and $\psi_x = y_1 - Z_1\delta$. ε_t is the contemporaneous error term and is assumed i.i.d. N(0, σ^2). Z_t is a vector of exogenous variables. Corresponding to the two-break equivalent of Perron's (1989) most general model (level and trends allowed to vary), with two changes in level and trend, Z_t is described by $[1, t, D_{1t}, D_{2t}, DT_{1t}, DT_{2t}]'$ where $D_{jt} = 1$ for $t \geq T_{Bj} + 1$, $j = 1, 2$ and zero otherwise, $DT_{jt} = 1$ for $t \geq T_{Bj} + 1$, $j = 1, 2$ and zero otherwise, and T_{Bj} stands for the time period of the breaks.
>
> The unit-root null hypothesis is described in the equation above by $\varphi = 0$ and the test statistic is a *t* statistic for this null. To endogenously determine the location of two breaks ($\lambda_t = T_{Bj} / T$, $j = 1, 2$), Lee and Strazicich (2003) use a grid search to find a minimum *t* statistic. Therefore, the critical values correspond to the location of the breaks (see Lee and Strazicich (2003) for more detail and the critical value tables).
>
> To implement this test, Lee and Strazicich first determined the number of augmentation terms S_{t-j} $j = 1,\ldots,k$, that correct for serial correlation in the equation above. At each combination of break points $\lambda = (\lambda_1, \lambda_2)'$ in the time interval [.1T, .9T] where T is the sample size, Lee and Strazicich determine *k* by following a general-to-specific procedure described by Perron (1989). Start with

> an upper bound k_{max} for k. If the last included lag is significant, choose k = k_{max}. If not, reduce k by 1 until the last lag becomes significant. If no lags are significant, set k = 0.

In addition, for researchers that wish to subsequently perform regression analysis on level-data (e.g., panel data sets over a number of years within the time series), the identification of break points is important. Running level-data regressions on data that span the break points invites spurious correlations.

The sequential tests presented in Bai and Perron (2003) follow the LM tests. This "BP Method" is useful in that it tests for the optimal number of breaks within a series based on sequential tests of the Sup F_t (l + 1/l) test. Following the LM tests, if needed, individual regressions are estimated through the use of the BP Method for any metric series with identified breakpoints in the aforementioned methodology. The BP Method is useful because it allows both the level and the trend of a series to change (Perron, 1989). The following equation specifies the general form of the BP Method regressions estimated:

$$METRIC_{t,c} = \alpha x_t + \beta_i z_t + \varepsilon_{t,c,}$$

where $t = T_{i-1} + 1, \ldots T_i$, and i = 1,…m + 1, and $METRIC_{t,c}$ is the specified historical competitive balance metric series in year t and conference c, i indexes the regime number, $T_1, \ldots T_m$ are the unknown specified break points, x_t (p x 1) and z_t (q x 1) are covariates with corresponding coefficients α and β_i which have the ability to vary over time. The α and β_i coefficients are the constant and trend terms and ε is the error term. A bit of clarification is needed regarding the relationship between the number of break points and regimes. A series with two identified break points will produce three total regimes – one regime prior to the first break point, a second regime between the two identified break points and a third regime following the second identified break point.

The BP Method allows for a determination on the statistical significance of identified breaks from the previous tests. This information combined with the updated trend values at each break is useful in explaining the historical behavior of the competitive balance metrics. This allows for a comparison of the breaks identified against the three key historical events outlined in earlier in the chapter. If the break points identified through the methods utilized here match up with the key historical events it would provide evidence against the IP. Alternatively, if the estimated break points do not match up with the specified events or if no break points exist in a series then supports exists for the IP.

4.7 Results: Historical Competitive Balance in College Football

As opposed to onerous and lengthy tables with the historical yearly values of each metric in each conference, Figures 4.1 through 4.18 are preferred due to their concise visual nature. Additionally, Tables 4.3 through 4.8 demonstrates conference averages by decade and by time period for the RSD, MVR and WPC metrics. Tables illustrating historical conference championship outcomes are shown in Tables 4.9 through 4.14.

Figures 4.1 through 4.6 illustrate the historical behavior of the RSD metric for each of the six equity conferences. In general, RSD has been relatively stable over the history of each conference with no clear trend in the behavior of the metric over time. Normal variation both above and below the conference mean is evident in each respective series. This suggests that over the history of each conference balance has been relatively stable as measured by the RSD metric. Of course, later with the more sophisticated time series analysis this is determined more precisely.

Table 4.3 provides conference RSD values by decade. This table shows that over the history of each respective conference the ACC and Pac-10 have been the most balanced as measured by the RSD metric based on overall RSD values of 1.50 and 1.51, respectively. Alternatively, the Big Ten, Big East and SEC have been the least balanced with RSD values of 1.58 – a difference of roughly five percent. Also of interest is the most recent ten-year period from 2000 to 2010 where all RSD averages with the exception of the Pac-10 were below the conference historical averages. This suggests enhanced balance as measured by RSD in the most recent ten-year period as compared to historical levels. On the other hand, the period from 1980 to 1999 represents an era of reduced year-to-year balance in both the Big Ten and Big 12. During this time period, the Big Ten was controlled handily by Michigan and Ohio State while Northwestern, Indiana and Minnesota were basement regulars. In the Big 12 (the Big 8 up until 1996) Nebraska, Oklahoma and Colorado dominated conference play resulting in the elevated RSD values during this twenty-year period.

When examining the behavior of RSD over the last twenty years evidence does exist supporting a general increase in balance (measured by a decrease in RSD) from 1990 to 2010. Five of the six BCS conferences had RSD values that declined by at least 7% over this time period. The only holdout was the Big East, which began conference play in 1991. This provides limited preliminary evidence that the distribution of winning percentages in BCS conferences may be compressing. However, caution is recommended based on the historical behavior of this metric – as fluctuation both above and below a conference's mean RSD value is common over the history of play (please see Figures 4.1 through 4.6).

Table 4.4 demonstrates conference RSD averages by time period. This table shows that the Pac-10 has shown substantially lower RSD values in each period as compared to the other five equity conferences. In addition to the findings in Table 4.3, this result suggests that on a season-to-season basis the Pac-10 has shown the most balance of all the equity conferences according to the RSD metric.

The historical behavior of the MVR metric for each conference is displayed in Figures 4.7 through 4.12. Clearly evident in each graph is a gradual downward trend of the MVR measure over time. In fact, in the period from 1960 to 2010, MVR has declined by minimum of 22.4% in the ACC all the way to a maximum of 35.0% in the SEC. This suggests substantial increases in balance as measured by individual game closeness in conference play. MVR also shows less overall variation around its moving average as compared to the RSD metric. All else equal, it is clear that balance as measured by game closeness through MVR has increased over each conference's history.

Conference MVR averages by decade are illustrated in Table 4.5. The lowest average value is seen in the ACC, which suggests it has the closest relative games of the six conferences. Opposite of the ACC is the Big Ten, which possesses a historical MVR value 30.16% larger than the ACC over the history of both conferences. This suggests the most uneven individual contests and less game uncertainty in the Big Ten. However, the Big Ten's overall MVR average is hampered based on the length of its playing history as all of its largest values are seen prior to 1940. This brings us to Table 4.6, which shows MVR by time period. In the most recent twenty-year period from 1990 to 2010, the Pac-

10 shows the lowest MVR value with the ACC, Big Ten and SEC approximately 5% to 7% higher. This implies individual games that are of relatively equal closeness based on the degree of scoring in each respective conference. The lowest levels of individual game closeness during this period are evident in the Big 12 and Big East – with MVR values approximately 16% larger than the Pac-10.

Figures 4.13 through 4.18 demonstrate the behavior of the WPC metric over the history of each conference. Unlike RSD and MVR, the WPC measure shows more volatility in both the spread of values among conferences and the stability of the metric over time. Visible from the Big 12 figure is the high levels of year-to-year correlation in team winning percentages from the mid-1970s to the late 1990s. This corresponds roughly to the period of domination by Nebraska, Oklahoma and Colorado that was referenced earlier in the discussion of RSD. The 1990-1999 period is also a prime example of the lack of balance in the Big 12. During this period, the conference's WPC metric stood at .7909, which is 21.9% greater than the next closest conference. Essentially, there was almost no year-to-year change in the conference standings in the Big 12 over the decade. The Big Ten also exhibits higher than average WPC values from 1970 to 2000. This era again corresponds to conference domination by Michigan and Ohio State and consistently poor conference performance from Northwestern, Indiana and Minnesota.

Conference WPC averages by decade are shown in Table 4.7. The Big Ten leads in terms of historical balance as measured by the degree of churning within conference standings. The Pac-10 follows slightly behind – 7.51% lower than the Big Ten. On the other end of the spectrum, the Big 12 produces the largest WPC value, or the highest degree of year-to-year correlation of team winning percentages – a value 23.28% larger than the Pac-10. Table 4.8 demonstrates WPC averages by time period for the six equity conferences. Evident again is the high degree of win percentage correlation in the Big 12 both in the most recent twenty year period and over the history of the conference. This suggests the Big 12 has had the least amount of churning in conference standings over time. Similar to the other metrics, the Pac-10 shows enhanced balance as compared to the other conferences as is evidenced by a WPC value of .4442 for the 1990 to 2010 time period.

Tables 4.9 through 4.14 provide historical conference championship data on each equity conference. In the ACC, Clemson has been the most successful program from the founding group that began play in the 1953 season, claiming a championship once every five seasons. Virginia has been the least successful program in the ACCs history, with zero outright titles and only two shared championships over 57 years of play. Despite only nineteen total ACC seasons, Florida State has earned ten conference championships, good for second all-time in the ACC. They have clearly dominated the conference over the past two decades.

The Big 12 conference championship history is shown in Table 4.10. This conference has been dominated like none of the other five BCS conferences with Oklahoma and Nebraska winning 64.83 championships in 83 total league years. The only other marginally competitive teams as measured by number of conference titles are Missouri with one crown every fifteen years and Colorado with one title every thirteen seasons. However, once the Big Eight became the Big 12 in 1996, Nebraska's dominance has become muted with Oklahoma continuing its supremacy and Texas taking over for the Cornhuskers.

In the Big Ten, Michigan and Ohio State have been the most successful programs over the history of the conference. The Wolverines have earned 27.24 total conference championships in 105 seasons –a title every 3.85 seasons. Meanwhile, Ohio State has earned 26.16 championships, but has done so in only 98 years of play. Indiana has been the program with the least success in the Big Ten with only 1.33 championships in 111 conference years. Other programs such as Illinois and Penn State have had marginal championship success, but the Big Ten has been controlled handily by Michigan and Ohio State over its history. Indeed, this is still largely the case with Ohio State earning at least a share of the Big Ten crown in each season from 2005-2010.

In only twenty years of Big East play, the conference has seen its share of changes in membership. Despite its instability, Miami has historically controlled the conference, winning 6.33 titles in only thirteen years of play. Beyond the Hurricanes, West Virginia, Syracuse, Virginia Tech and Cincinnati have all won multiple titles in the Big East's short existence. On the other end of the spectrum, Pittsburgh has only shared in a single conference title in twenty years of play despite its history of producing quality NFL players. Rutgers has been the conference's least successful program in terms of championships as the school is still searching for its first Big East crown.

Southern California has captured the most crowns in the Pac-10 with 32.66 in 89 years of play. Following the Trojans, UCLA, Washington and California have all won more than ten titles each. Stanford and Oregon follow with 9.83 and 6.83 championships, respectively. In fact, all of the conference's current members have captured at least a portion of a conference crown. In all, this makes the Pac-10 the most historically balanced in terms of the distribution of conference titles.

Lastly, Alabama has captured 18.83 SEC titles in 78 years of conference play, the most of any team. Tennessee (10.66), Georgia (10) and Florida (8) all follow Alabama in conference championships. Despite their status as founding members of the SEC, both Kentucky and Mississippi State have claimed only one outright title each in 78 years of play. Vanderbilt, another founding member has yet to claim any portion of a SEC crown. Overall, the SEC has not been dominated by a single team but there is a clear distinction between the "haves and the have-nots" in terms of championship success.

The data presented here reveal detailed information regarding the historical behavior of competitive balance in college football. It is useful to begin with a brief discussion of the general behavior of the identified metrics over time. A visual inspection of the RSD measure in the six BCS conferences illustrates that balance as measured by the dispersion of winning percentages in a conference has not changed much over this history of each conference. There have been within conference fluctuations in the value of this metric over time, but there has been no systematic change in RSD over the examination period.

Alternatively, the relative closeness of individual games as measured by MVR shows a clear trend towards enhanced levels of balance in all six conferences. From the origination of each conference through 2010, the MVR ratio has decreased, suggesting closer games relative to the level of scoring in each conference. In terms of the WPC measure, there is no clear pattern of a general increase or decrease in the metric over time. Individual conference WPC values are at roughly the same levels currently as compared to the early years of each conference. However, there is more within-conference variation in this metric as compared to both RSD and MVR.

A look at the data at the individual conference level provides an even more detailed understanding of the behavior of balance in college football. It is clear based on the four competitive balance metrics presented here that the Pac-10 has been the most balanced conference over the examination period. The group is undoubtedly the most balanced in terms of championship equality and is at the top or in the top pair of conferences in the RSD, MVR and WPC metrics. Additionally, over the past twenty years, the conference has boasted the lowest RSD, MVR and WPC metrics. Clearly, the Pac-10 has been the most balanced conference in the modern era. The title of the least balanced conference belongs to the Big 12 with its aforementioned 1990-1999 period where its WPC metric stood at .7909. Essentially, there was almost no year-to-year change in the conference standings in the Big 12 in the 1990s.

In general, the ACC, Big Ten, Big East and SEC fall in the middle of the pack in terms of balance relative to the other BCS conferences. However, these conferences share a similar trait to both the Pac-10 and Big 12 – in general there has been relatively little variation at the top end of each conference over time. A granular look at the data shows that only a select number of programs in each conference have enjoyed any type of sustained success. As previously noted, Nebraska and Oklahoma have owned the Big 12 along with Michigan and Ohio State in the Big 10 and Clemson and Florida State in the ACC. Southern Cal has almost two and a half times the number of conference titles as UCLA, the second most successful program in the Pac-10. The same can be said in the SEC where Alabama has enjoyed conference supremacy. Tennessee, Florida and Georgia make up the rest of a historically strong group in the SEC. While the Big East has little extended history to speak of the conference has still experienced a reign of supremacy at the hands of Miami. Only after the departure of the Hurricanes have other programs flourished. The take home message here is that at the individual conference level history tells us that surprise programs make their presence felt on an occasional basis, but over the long haul a small number of elite teams dominate their respective conferences.

4.8 Results: Stationarity and Break Points

Table 4.15 illustrates the results of ADF and PP tests for the RSD, MVR and WPC metric series for each conference. YPC measure does not produce a metric for each conference year so it is of no use in this part of the analysis. The ADF tests almost uniformly reject the unit-root null hypothesis for each conference metric, which suggests series stationarity. The results of the PP test are even more consistent with all but one of the series showing significance at the 99% critical level. Together the ADF and PP unit-root test results strongly suggest series stationarity in each balance metric.

The next step conducts two-break minimum LM unit-root tests on each competitive balance metric. Table 4.16 displays these results for the RSD metric and shows that the unit-root null is rejected at the 95% level in each series. Both break points in the ACC, Big Ten and Pac-10 series are also significant at the 99% level suggesting that the two-break test is appropriate for the RSD metric in these conferences. Table 4.17 provides two-break LM tests results for the MVR metric. All five conferences again show statistical significance at the 95% level or better which allows for a rejection of the unit-root null. Both break points in the Big Ten and SEC are statistically significant suggesting the two-break LM test as appropriate for this these two conferences. Two-break LM test results for the WPC metric are illustrated in Table 4.18. Four of the five conferences show statistically significant test statistics with the Pac-10 as the only holdout. The ACC, Big Ten and Big 12 also exhibit two significant break points each, which supports the use the two-break LM test in these conferences.

One-break minimum LM tests are also conducted for each metric series in each conference. Similar to that of the two-break tests, the one-break tests are significant in each of the five conferences. However, the one-break minimum tests have more applicability to the conferences where both breaks are non-significant in the two-break tests. Specifically to the RSD measure, the Big 12 and SEC are of particular interest. Table 4.19 shows that the one-break tests in both conferences produce significant breaks, suggesting that the one-break test is more appropriate. Table 4.20 displays the one-break LM tests for the MVR measure. Again, all five conferences show significant test statistics at the 95% level or greater. Since the two-break LM test did not produce two significant break points in the Big 12 or Pac-10, results from the one-break test are needed. In this case, the one-break test produces a significant break point in both conferences, which can be interpreted that the one-break LM test is more appropriate for the MVR metric in the Big 12 and Pac-10.

Lastly, the one-break LM results for the WPC metric are provided in Table 4.21. Unlike RSD and MVR, the WPC one-break test statistics are not uniformly significant as the Big 12 and Pac-10 values are not significant at standard levels. However, the Pac-10 and SEC are the two conferences of interest here based on non-significant break points in the two-break tests. The one-break test is clearly more appropriate for the SEC based on a significant break point. In summation, the results of the ADF, PP, two-break, and one-break unit root tests strongly support series stationarity in the competitive balance metrics evaluated. However, the results of the LM tests also suggest the presence of structural breaks within a number of series. Sequential tests will be performed to determine the number of structural breaks, if any, in each particular competitive balance metric series.

Tables 4.22 through 4.24 present the results of sequential tests presented in Bai and Perron (2003) for the RSD, MVR and WPC metrics, respectively. The upper portion of each table presents test statistics for the Sup $F_t(k)$ tests which specify the statistical significance for a given number breaks (k) against a null of zero breaks. The UD_{max} and WD_{max} tests are maximum tests for the presence of any structural breaks up to an unknown maximum against the null of zero structural breaks. The lower portion of each table provides test statistics for the sequential Sup F_t (l + 1/l) tests, which allow for a comparison of the significance of a series with l +1 breaks against a series with l breaks.

Table 4.22 displays the results of the sequential tests for the conference RSD metrics. In the ACC, each of the five Sup $F_t(k)$ are significant which suggests the presence of at least one break. A look at the sequential test results in the lower portion of the table shows statistical significance on the Sup F(2/1) test suggesting that two break points are preferred over a single break. The Big Ten also displays a significant result for each Sup $F_t(k)$ test with the Sup $F_t(1)$ test significant at the 99% critical level. Combined with the lack of significance in any Sup F_t (l + 1/l) test, this confirms only a single break in the Big Ten RSD series. The Big 12 RSD series shows statistical significance for each Sup $F_t(k)$ test but at lower critical levels. Based on a Sup $F_t(1)$ test significant at only the 90% level and no significant Sup F_t (l + 1/l) tests, this suggests a lack of break points in the Big 12 series. The Pac-10 and SEC share a similar RSD series profile with no significant Sup $F_t(k)$ or Sup F_t $(l + 1/l)$ tests. This is interpreted as a lack of break points in the Pac-10 and SEC RSD series.

The results of the sequential tests for the conference MVR metrics are demonstrated in Table 4.23. The ACC and Pac-10 show five significant Sup $F_t(k)$ tests and significant Sup F_t $(l + 1/l)$ tests suggesting two breaks for the ACC and a single break for the Pac-10. The Big 12 shows four of five significant Sup $F_t(k)$ tests but a non-significant Sup $F_t(1)$ test which infers a lack of breaks. The Big Ten and SEC display no significant Sup $F_t(k)$ tests and therefore have no break points in their respective MVR series.

Table 4.24 exhibits conference WPC metric series results for the sequential tests. The Big Ten and Big 12 both show five Sup $F_t(k)$ tests with significance levels at 95% or above which suggests the presence of at least one break in those two conferences. The Big Ten shows zero significant Sup F_t $(l + 1/l)$ tests, which infers only a single break in the conference's WPC series. However, the Big 12 illustrates a Sup F(2/1) test which is significant at the 99% critical level. Despite a Sup F(3/2) test significant at the 90% level, the statistical power of the Sup F(2/1) test suggests two breaks in the Big 12 metric. Rounding out the group is the ACC, Pac-10 and SEC which show no evidence of breaks based on the results of the sequential tests. Table 4.25 conveniently provides a single illustration highlighting the number of break points in each balance metric, the actual year of the estimated break and 90% confidence intervals for each. A total of nine breaks in six different series are found.

BP Method regression results for conference metrics with estimated break points are displayed in Table 4.26. Trend and intercept coefficient values, which estimate the behavior of the series over time, are denoted by α_i and β_i respectively. In the ACC RSD metric series, each of the three intercepts and two of the three trend coefficients are significant at the 99% level. The ACC model produces a R^2 value of .345. In the Big Ten, both trend variables and one of the two intercepts are significant at the highest level and the model produces a R^2 value of .216. On the surface, it is tough to internalize the impact of the significance of these coefficients. Therefore, graphs that plot the actual behavior of the RSD metric over time against the fitted behavior of the metric which models the breakpoints, trends and changes in intercept are provided in Figures 4.19 through 4.23. The corresponding graph for the ACC RSD metric is shown in Figure 4.19 and the Big Ten is shown in Figure 4.21.

MVR BP Method regression results are also highlighted in Table 4.26. The ACC model produces a R^2 value of .653 based on three trend and three intercept coefficients significant at the 99% level. Graphs demonstrating actual versus fitted values for the MVR metric are provided in Figures 4.24 through 4.28 with the ACC shown in Figure 4.24. Each of the Pac-10 coefficients is also significant at the highest level and the model produces a R^2 value of .769. The Pac-10 MVR visual is provided in Figure 4.27.

Lastly, BP regression results for conference WPC metrics are presented in Table 4.26 with the corresponding graphs shown in Figures 4.29 through 4.33. The Big 12 model produces a R^2 value of .327. Two of three intercept coefficients and two of three trend coefficients are significant in this model. The visual description of the fitted Big 12 WPC metric is shown in Figure 4.30. Meanwhile, the Big Ten model produces a R^2 value of .175 with both intercept coefficients and neither trend coefficients showing statistical significance. The fitted WPC graph for the Big Ten is provided in Figure 4.31.

4.9 Results: The Invariance Proposition

Before evaluating the results for the IP from the BP Method, a preliminary look at the behavior of the competitive balance metrics in pre-event and post-event periods will be completed. This follows in line with the work of Fort and Quirk (1995) who use this methodology to evaluate the behavior of the distribution of winning percentages in response to a host of changes in league configuration in the major North American team sports leagues.

Tables 4.27, 4.28 and 4.29 show the behavior of the RSD, MVR and WPC metrics both before and after the introduction of the G.I. Bill in 1946. The pre-event and post-event periods are evaluated against each other using paired t-tests. Table 4.27 shows the RSD metric in the pre and post G.I. Bill periods. The only significant differences are seen in the Pac-10 and SEC. However, the Pac-10 shows an increase in RSD while the SEC shows a decrease in RSD. Combined with non-significant pre-event to post-event changes in the Big Ten and Big 12 there is no evidence of a systematic change in RSD following the introduction of the G.I. Bill. This result would support Rottenberg's IP.

Table 4.28 displays the behavior of MVR in response to the G.I. Bill and shows significant reductions in MVR from the pre-event to post-event period in all four conferences. This clear and consistent result suggests that balance as measured by game closeness has increased following the G.I. Bill. However, a closer look at the raw data shows that MVR has systematically decreased over time in each of the BCS conferences.

Therefore, this significant change may simply be part of a systematic reduction in the behavior of MVR and not necessarily attributable to the introduction of the G.I. Bill. Time series analysis will provide more information on this finding.

Table 4.29 presents the WPC metric both pre and post G.I. Bill periods. In three of the four conferences WPC decreases in the post-G.I. Bill period, however none of these reductions are statistically significant. The only significant change is in the positive direction in the Big 12 and is interpreted as a reduction in balance. In summation, no clear pattern of behavior exists in the WPC metric in response to the introduction of the G.I. Bill, which would support the IP.

Competitive balance behavior in response to the formal institution of athletic grant-in-aid in 1956 is shown in Tables 4.30, 4.31 and 4.32. These three tables vary in that they contain standard pre-event and post-event periods but also an alternative post-event period, which allows for a four year buffer period following the introduction of the regulation. Table 4.30 displays RSD in response to grant-in-aid. In the standard post-event period the only significant effect is seen in the Pac-10 with a sharp reduction in RSD. This same significant effect is also visible for the conference in the alternate post-event period. However, a significant increase in RSD is evident in the Big 12 in the alternate post-event period. Again, no consistent significant changes are seen in the RSD metric following the introduction of grant-in-aid.

The behavior of MVR and WPC in response to formal grant-in-aid is displayed in Tables 4.31 and 4.32, respectively. Interestingly, despite the gradual decline of MVR over time as shown in Figures 4.7 through 4.12, there are no significant changes in any conference from the pre-event to post-event periods. These findings would support

Rottenberg's IP. With WPC, a significant increase in the measure is seen in the SEC in the pre-event to post-event periods. When comparing the pre to alternate post-event periods significant increases are seen in the SEC and Pac-10 while a significant decrease occurs in the Big 12. Once again there is no consistent behavior in the metric, which makes claims of any changes in the metric due to the occurrence of athletic grant-in-aid invalid.

Tables 4.33 to 4.35 illustrate the behavior of RSD, MVR and WPC in response to the death of the CFA in 1995. The only significant change in RSD is seen in the Pac-10 as the metric rose from 1.37 to 1.56. The Pac-10 also displayed the only significant change in MVR with a 15.15% decrease in the measure. Taken together these two results suggest that over the twenty-year examination period surrounding the death of the CFAs the spread of Pac-10 winning percentages increased while individual games within the conference became more competitive. Lastly, five of the six BCS conferences showed no significant change in WPC with the exception of the Big 12. Based on the lack of significant changes in the competitive balance metrics in the pre-event to post-event periods, there is little evidence for any substantial changes in balance following the cessation of the CFA. These results would support the IP.

Because the use of the pre-event versus post-event methodology is not statistically feasible on the YPC metric, a qualitative examination of the historical behavior of the measure will have to suffice. Tables 4.9 through 4.14 illustrate the championship history in each BCS conference. In the ACC, Florida State has dominated since its entry into the conference in 1992. In its first nineteen years of conference play the Seminoles captured eleven total championships. Even more recently, Virginia Tech captured four titles in only seven years. Prior to the runs of Florida State and Virginia Tech, Clemson controlled the conference with twelve conference championships from 1953 to 1991. The Big 12 shows little difference as compared to the ACC as a small number of teams have swiped the lion's share of titles. Oklahoma and Nebraska have essentially passed the Big 12 championship trophy back and forth from Norman to Lincoln as the two schools have combined to share 64.83 of the 83 total conference crowns.

Championship history in the Big Ten is slightly more dispersed than the Big 12 with Michigan and Ohio State claiming roughly half of the titles in the 115-year history of the conference. Illinois and Minnesota have both claimed over ten crowns but both squads sit at roughly one title every ten seasons – which equates to the number they would earn through randomly distributing the title on a yearly basis among all conference teams. The Big East is the youngest BCS member and also has an enhanced distribution of conference titles. Due in part to the excessive turnover in the conference five programs have won multiple titles in the twenty years of Big East play.

In the Pac-10, Southern California has claimed 32.66 of 95 possible titles – displaying a clear dominance over the rest of the conference. California, UCLA and Washington have also been competitive over time. Lastly, Alabama has been the dominant team in terms of conference championships in the SEC. Tennessee, Georgia and Florida are the only other conference programs averaging a conference title at a better clip than once every ten seasons.

The take home message stemming from a basic analysis of BCS conference championship history is twofold. First, it is very clear that with the exception of the Big East the BCS conferences have been dominated by a very small number of teams in each conference. While the average program may have a short-run episode of success, this rarely holds over time. What is evident is that over the history of these conferences there is a clear divide between the power programs and the rest of the programs in the conference. Secondly, there is very little evidence that any of the three key historical events identified here are associated with changes in the historical distribution of conference championships. Clearly, different programs have had successful stretches over various periods, but there is little evidence to suggest there are any patterns or changes in distribution tied to the G.I. Bill, grant-in-aid or the death of the CFA.

In summary, the "before and after" results suggest little in the way of any type of consistent change in any of the competitive balance metrics following the G.I. Bill, athletic grant-in-aid or the death of the CFA. The only statistically significant change in any of the metrics was seen by way of a reduction in the margin of victory ratio following the initiation of the G.I Bill. However, a general look at the behavior of this measure over time shows a clear and consistent reduction in MVR over the history of the sport – a finding that would largely eliminate the G.I. Bill as an explanation for the significant decrease in the measure. Likewise, little evidence exists to support any sort of systematic change in the distribution of conference championships in response to the three events discussed here. Overall, this provides no support for changes in competitive balance in response to the three events identified. These results all support the IP.

The time series methodology utilized confirms that nine of the fifteen historical competitive balance metric series are stationary without break points. The remaining six metric series are stationary with either one or two breaks. The 60% of metrics that are stationary without breaks undoubtedly support the IP – as a lack of breaks is interpreted as no structural changes in the specified balance measure. In the remaining six series where breaks were identified, a closer look is needed to determine whether or not any breaks match with the events identified.

Again, Table 4.25 illustrates the series containing break points along with break dates and corresponding 90% confidence intervals for each break. Beginning with the RSD metric, the second break in the ACC is estimated to take place in 1993. The 90% confidence interval for this break ranges from 1987 to 1994, which misses the death of the CFA by only a year. The corresponding trend following this break is a reduction in the RSD metric, which is interpreted as an increase in balance. If the assumption is made that this break aligns with the death of the CFA, this would be evidence against the IP.

The first and only break in the Pac-10 MVR series is estimated to take place in 1947 which is the year following the introduction of the G.I. Bill. However, the 90% confidence interval spans from 1944 to 1952 and clearly includes this event. Following this break a slight shift down in the metric occurs and a negative trend also continues. This change would suggest an improvement in balance following the introduction of the G.I. Bill and provides evidence against the IP.

Another look at Table 4.25 shows that the first Big 12 break point in the WPC series occurs in 1950 with the 90% confidence interval covering the period from 1949 to 1967. This interval includes the introduction of formal athletic grant-in-aid in 1956. Previous to this break point, WPC in the Big 12 was increasing steadily which is a reduction in balance. Following the break, there is a downward shift in the metric and a leveling of WPC over time until the second break in the series. The aforementioned break is interpreted as an increase in balance as measured by WPC and would provide evidence against the invariance proposition. Each of the three break points outlined above match approximately but not directly with one of the key historical events presented. This provides limited evidence of a match between the given historical event and a statistically significant change in balance.

The remaining three series with identified break points do not match with any of the historical events identified here. However, three of these six total remaining breaks fall within a short period of time (1963-1966) and a fourth falls in 1970. In addition to proximity to each other, all four breaks are associated with a decrease in balance. This suggests that a key institutional change in college football may be associated with these breaks. It is possible that these breaks match with the general social turbulence of the time period as the 1960s were rife with social movements and the early 1970s were marred by general opposition against the Vietnam War. If these events altered attendance which would subsequently impact revenues then these events could have an impact on competitive balance. However, as a more likely event relationship, these dates match approximately with the integration of African-American football players in the south.

More specifically, in 1963 Darryl Hill of Maryland became the ACCs first black college football player and Kentucky was the first SEC school to integrate in 1966 (Sawchick, 2010). Prior to this point some southern schools fought to keep African-Americans of their teams or refused to play teams with African-American players (Lapchick, 2008). The turning point came in 1970 when USC and their all African-American backfield were the first fully integrated team to play at the University of Alabama. USC handily beat the non-integrated and Bear Bryant led Crimson Tide 42-21. Many historians mark this as the turning point in which the southern states accepted full integration of college football (Everson, 2009). By 1971, every SEC team had at least one African-American player (Lapchick, 2008).

Meanwhile, schools in the north and northeast were years ahead in this regard with Michigan State being the first large conference school to fully integrate (Sawchick, 2010). Syracuse University is another example as two-time All-American tailback Ernie Davis led the Orangemen to the 1959 NCAA Division I-A National Championship. While integration is not the focus of this study, there appears to be evidence suggesting that integration in the southern football schools may be associated with changes in competitive balance. In fact, Lee and Fort (2005) account for this possibility in their time series analysis of competitive balance in MLB. While a cursory look at the northern versus southern schools during this era does not illuminate any obvious disparities in balance or domination of play by schools from a particular region, this is certainly a topic worthy of future investigation.

In summation, based on the analysis here mixed evidence exists on Rottenberg's IP. In the pre versus post methodology there is little substantiation for any changes in balance in response the events identified. There is also little change in the distribution of conference championships over time. However, in regards to the time series methodology, three break points match up approximately with the key historical events described. Furthermore, following each of these three break points significant improvements in balance are seen. The remaining twelve series either do not include breaks or have breaks that do not match with the events highlighted here. These cases would provide support for the existence of the IP.

4.10 Conclusions and Suggestions for Future Research

This study contributes to the academic literature in sports economics by examining the historical behavior of competitive balance in college football. An examination of the four competitive balance metrics presented illustrates an increase in game closeness over time as measured by a substantial reduction in the margin of victory ratio. The correlation of team winning percentages in BCS conferences has remained relatively stable over time suggesting little change in consecutive season balance. The distribution of team winning percentages in a conference as measured by RSD has also remained relatively consistent over time despite recent data which suggests that year-end balance may actually be increasing. Only additional data will confirm whether or not this reduction in RSD is permanent or simply a short-term fluctuation around a long-run average.

This work also shows that individual BCS conferences have historically been dominated by a single team or a very small group of elite programs. Mid-tier and basement dwelling programs have had periods of sustained success, but these runs are almost always short-term in nature. A detailed look at the championship data, year-to-year conference standings and individual game level data shows that over time the largest conferences in the NCAA FBS have been controlled by a select group of programs. Despite this, NCAA college football remains a supremely popular spectator sport in the United States. This raises the question of how important competitive balance really is to the long-term financial viability of NCAA conferences – an idea originally developed by Rottenberg (1956) in the original paper on sports economics.

Turning to the IP, This section contributes to the sports economics literature in a number of ways. First, it advances the relatively scant scholarship on competitive balance in college football by providing a long-run analysis on the behavior of balance in response to three key events which altered the business structure of the sport. The section also extends Rottenberg's IP to college sports – something that has yet to be done in the academic literature. The use of time series techniques allows for the ability to examine the time series for unit root (stationarity) and structural break points. This gives the added advantages of 1) a more in-depth understanding of the behavior of the time series, 2) the avoidance of spurious correlations if data are analyzed across structural breaks, and 3) the chance to associate the history of events with the break points. This last allows for the ability to gauge the effectiveness of competitive balance policy interventions and evaluate whether or not Rottenberg's IP holds. Furthermore, this method can be extended to football at lower levels where finances are quite different. Finally, since the point of tracking is just that, there also remains the more micro-level analysis of the possible explanations offered by pairing up the history of FBS events with the behavior of the time series.

From an empirical standpoint, this section produces analysis which unearths support both for against the existence of the IP in the world of college football. Specifically, time series techniques show that nine of fifteen competitive balance metric series are stationary without break points. This result tells us that competitive balance within BCS conferences has been relatively stable over time and that the three events described here have not significantly influenced balance in the majority of cases. A cursory look at conference championship history also illustrates that a small number of programs have dominated their given conferences and that the distribution of championship titles has not changed substantially over time. On the other hand, three of the nine total break points identified match approximately with one of the three key events presented. In each case the identified break point is followed by a subsequent enhancement of balance. This provides evidence supporting the hypothesis that the institutional changes identified in this section are associated with structural changes in competitive balance. This result would oppose Rottenberg's IP.

It is also important to note the limitations associated with this chapter. At this point in time there are no other academic contributions which have outlined the existence of the IP in the context of college sports. Based on the lack of research in the area, certain assumptions have been made regarding the objective function of university athletic departments. In the professional sports context, franchise owners are largely considered to be profit maximizers. However, in the college sports literature, there is no work which models the objective function of NCAA athletic departments. Anecdotal evidence suggests that Athletic Directors may be revenue maximizers as opposed to profit maximizers. Future work should focus on clarifying this possibility which would shed further light on the place of the IP in the college sports context.

Secondly, this chapter identifies three key institutional changes in the business structure of NCAA college football and evaluates changes in balance in response to each event. Claims are made that balance shows either no significant change or a significant change in response to each event. However, it is important to note that other events could also be impacting the behavior of balance over time.

As for directions for future work, research on the economics of college sport is scant compared to that of professional sport. There are a limited number of contributions which examine competitive balance in college football and those cover only limited time periods. This work contributes to the literature by producing historical competitive balance data on the sport and also explains how balance has behaved at the highest level of NCAA football. This contribution will assist future researchers who wish to utilize these competitive balance metrics when investigating consumer demand for college football or the relationship between revenues and balance in the sport.

A natural follow up to this study would analyze long-run attendance demand for NCAA football at the FBS level. The UOH line of competitive balance research is focused on the relationship between balance in a league and consumer interest in that league. As previously mentioned, there are sparse contributions focused on consumer demand for NCAA football and none of these control for the effect of balance on demand. Now that a substantial collection of balance metrics exist, the ability exists to properly estimate attendance demand for the sport while also accounting for all three aspects of competitive balance – game uncertainty, playoff uncertainty and consecutive season uncertainty.

Future research on the IP in college sports is largely unrestricted. A natural extension to this work would be an analysis of competitive balance in NCAA college basketball along with the identification of key historical events which have altered the business structure of the sport. A time series analysis of balance in response to the highlighted events would offer an interesting application of the IP in college basketball – another largely unexplored area in the field of sports economics.

4.11 References

Andrews, D.S. (1984). The GI Bill and college football: The birth of a spectator sport. *Journal of Physical Education, Recreation & Dance*, 55(7), 23-26.

Associated Press. (1994, February 15). ACC develops major tv deal. *The Augusta Chronicle*. p. C10.

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica, 66*, 47-78.

Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, *18*, 1-22.

Bennett, R.W., & Fizel, J. L. (1995). Telecast deregulation and competitive balance: NCAA Division I football. *American Journal of Economics and Sociology*, 54, 183-199.

Boda, S., & Claasen, H. (1961). Ronald encyclopedia of football. NY: The Ronald Press, Co.

Brooks, B.G. (1995, December 2). Big 12 presidents back longer tv deal. *Rocky Mountain News*. p. 2B.

Butler, M.R. (1995). Competitive balance in major league baseball. *American Economist, 39*, 46-52.

Byers, W. (1995). Unsportsmanlike conduct: Exploiting college athletes. Ann Arbor, MI: University of Michigan Press.

Cairns, J. (1987). Evaluating changes in league structure: The reorganization of the Scottish Football league. *Applied Economics*, 19, 259-275.

Daly, G., & Moore, W.J. (1981). Externalities, property rights and the allocation of resources in Major League Baseball." *Economic Inquiry*, 19, 77-95.

Depken II, C.A., & Wilson, D.P. (2006). NCAA enforcement and competitive balance in college football. *Southern Economic Journal*, 72(4), 826-845.

DeSchriver, T.D., & Jensen, P.E. (2006). Determinants of spectator attendance at NCAA Division II football contests. *Journal of Sport Management*, 16(4), 311-330.

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

El-Hodiri, M., & Quirk, J. (1971). An economic model of a professional sports league. *Journal of Political Economy*, 70, 1302-1319.

Everson, D. (2009, December 4). The game that changed Alabama. *The Wall Street Journal*. Retrieved June 15, 2012 from: http://online.wsj.com/article/S B10001424052748704107104574572340901355488.html.

Falla, J. (1981). NCAA: The voice of college sports. Mission, KS: National Collegiate Athletic Association.

Fort, R. (2006a). Sports economics (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Fort, R. (2006b). Competitive balance in North American professional sports. In J. Fizel (Ed.) Handbook of Sports Economics Research (pp. 190-206). Armonk, NY: M.E. Sharpe, Inc.

Fort, R., & Lee, Y.H. (2006). Stationarity and MLB attendance analysis. *Journal of Sports Economics, 7,* 408-415.

Fort, R. & Lee, Y.H. (2007). Structural change, competitive balance, and the rest of the major leagues. *Economic Inquiry, 45*, 519-532.

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

Fort, R., & Quirk, J. (1995). Cross-subsidization, incentives, and outcomes in professional team sports leagues." *Journal of Economic Literature*, 23, 1265-1299.

Fort, R., & Quirk, J. (1999). The College Football Industry. In John Fizel,. Elizabeth Gustafson and Lawrence Hadley (eds.) *Sports Economics: Current Research* (Westport, CT: Praeger Publishers), pp. 11-26

Furfari, M. (2010, February 1). Big East pays out. *The Journal*. Retrieved December 8, 2010, from: http://www.journal-news.net/page/content.detail/id/531210/Big-East-pays-out.html?nav=5009.

Greenspan, D. (1988). College football's biggest fumble: The economic impact of the Supreme Court's decision in National Collegiate Athletic Association v. Board of Regents of the University of Oklahoma. *The Antitrust Bulletin*, 33(1), 1-65.

Haydock, M.D. (1996). The GI Bill. *American History*, 31, 52-70.

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

Kesenne, S. (2005). Revenue sharing and competitive balance: Does the invariance proposition hold? *Journal of Sports Economics*, 6, 98-106.

Lapchick, R. (2008, February 20). Breaking the college color barrier: Studies in courage. *ESPN.com*. Retrieved May 16, 2012, from: http://sports.espn.go.com/espn/ blackhistory2008/columns/story?columnist=lapchick_richard&id=3254974.

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

Lee, Y.H., and Fort, R. (2008). Attendance and the uncertainty-of-outcome hypothesis in baseball. *Review of Industrial Organization*, *33*, 281-295.

Lee, Y.H., and Fort, R. (2012). Competitive balance: Time series lessons from the English Premier League. *Scottish Journal of Political Economy*, 59(3), 266-282.

Lesmerises, D. (2010, June 1). Big Ten expansion would be about opportunities, says OSU's Gene Smith. *The Cleveland Plain Dealer*. Retreived December 8, 2010, from: http://www.cleveland.com/osu/index.ssf/2010/06/big_ten_expansion_wo uld_be_abo.html.

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

Matuszewski, E. (2010, October 21). Expanded Pac-10 Conference to share revenue, hold football championship. *Bloomberg*. Retrieved December 8, 2010, from: http://www.bloomberg.com/news/2010-10-21/expanded-pac-10-conference-to-share-revenue-hold-football-championship.html.

Moskos, C. (1988). A call to civil service: National service for country and community. NY: The Free Press.

National Collegiate Athletic Association. (2010, April 22). NCAA signs new 14-year TV deal for men's DI basketball. *NCAA.org*. Retrieved December 8, 2010, from: http://www.ncaa.org/wps/portal/ncaahome?WCM_GLOBAL_CONTEXT=/ncaa/ NCAA/NCAA+News/NCAA+News+Online/2010/Association-wide/NCAA+signs+new+14year+TV+deal+for+DI+mens+basketball_NCAANe ws_04_22_10.

Newey, W., and West, K. (1994). Automatic lag selection in covariance matrix estimation. *Review of Economic Studies, 61*, 631-653.

Noll. R.G. (1988). Professional basketball. *Studies in Industrial Economics Paper No. 144*. Stanford University, Stanford, California.

Quirk, J. (2004). College football conferences and competitive balance. *Managerial and Decision Economics*, 25, 63-75.

Quirk, J., & El-Hodiri, M. (1974). The economic theory of a professional sports league. In R.G. Noll (Ed.), *Government and the Sports Business* (pp. 33-80). Washington, DC: Brookings Institution Press.

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

Quirk, J.P. & Quirk, Q.T. (2012). College football history: Records and game scores for the major college football programs from 1869 onward. CDROM under copyright.

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica, 57*, 1361-1401.

Price, D.I., & Sen, K.C. (2003). The demand for game-day attendance in college football: An analysis of the 1997 Division I-A season. *Managerial and Decision Economics*, 24(1), 35-46.

Reimann, P.A. (2004). The G.I. Bill and collegiate football recruiting after World War II. *International Sports Journal*, 8(2), 126-133.

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

Sandomir, R. (1991, August 25). Notre Dame scored a $38 million touchdown on its tv deal. *The New York Times*. Retrieved November 16, 2010, from: http://query.nytimes.com/gst/fullpage.html?res=9D0CE0D91038F936A1575BC0A967958260&pagewanted=print.

Sawchick, T. (2010, November 10). Clemson's first black football player has spent a lifetime laying foundations. *The Post and Courier*. Retrieved May 16, 2012, from: http://www.postandcourier.com/article/20101110/PC1602/311109944.

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

Siegfried, J.J., & Burba, M.G. (2004). The College Football Association television broadcast cartel. *The Antitrust Bulletin*, 49, 799-819.

Sloane, P.J. (1976). Sporting equality: Labour market versus product market control – A comment. *Journal of Industrial Relations*, 18, 79.

Sutter, D., & Winkler, S. (2003). NCAA scholarship limits and competitive balance in college football. *Journal of Sports Economics*, 4(1), 3-18.

Szymanski, S. (2003). The economic design of sporting contests. *Journal of Economic Literature*, 41(4), 1137-1187.

Szymanski, S. (2004). Professional team sports are only a game: The Walrasian fixed-supply conjecture model, contest-Nash equilibrium, and the invariance principle. *Journal of Sports Economics*, 5, 111-126.

Szymanski, S., & Kesenne, S. (2004). Competitive balance and gate revenue sharing in team sports. *Journal of Industrial Economics*, 52, 165-177.

Ubben, D. (2010, June 2). Big 12 Blog. *ESPN.com*. Retrieved December 8, 2010, from: http://espn.go.com/blog/big12/post/_/id/13032/unequal-revenue-sharing-will-remain.

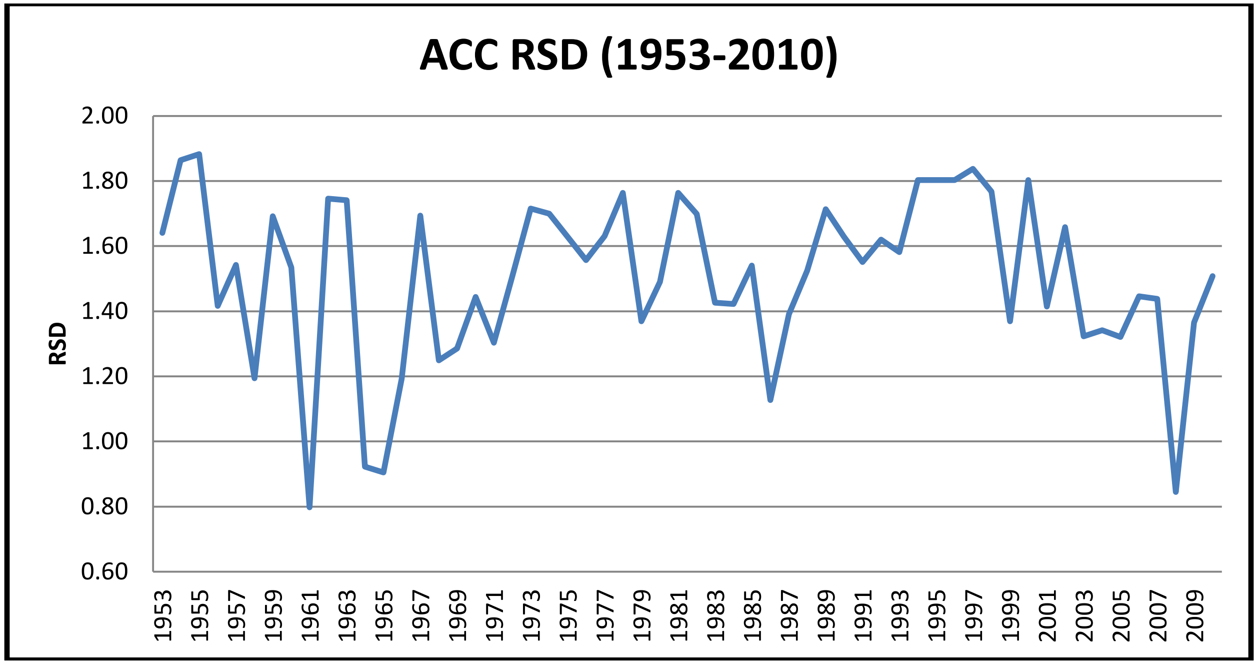
Vrooman, J. (1995). A general theory of professional sports leagues. *Southern Economic Journal*, 61, 971-990.

Wellman, R. (2010, August 4). Ron Wellman's column: Conference realignment. *Wake Forest Deamon Deacon Athletics*. Retrieved December 8, 2010, from: http://wake forestsports.cstv.com/genrel/080410aaa.html.

Wilner, J. (2010, August 12). Pac-10 football: Conference revenue and per-school distribution Figures. *Mercury News*. Retrieved December 8, 2010, from: http://blogs.mercurynews.com/collegesports/2010/08/12/pac-10-football-conference-revenue-and-per-school-distribution-figures/.

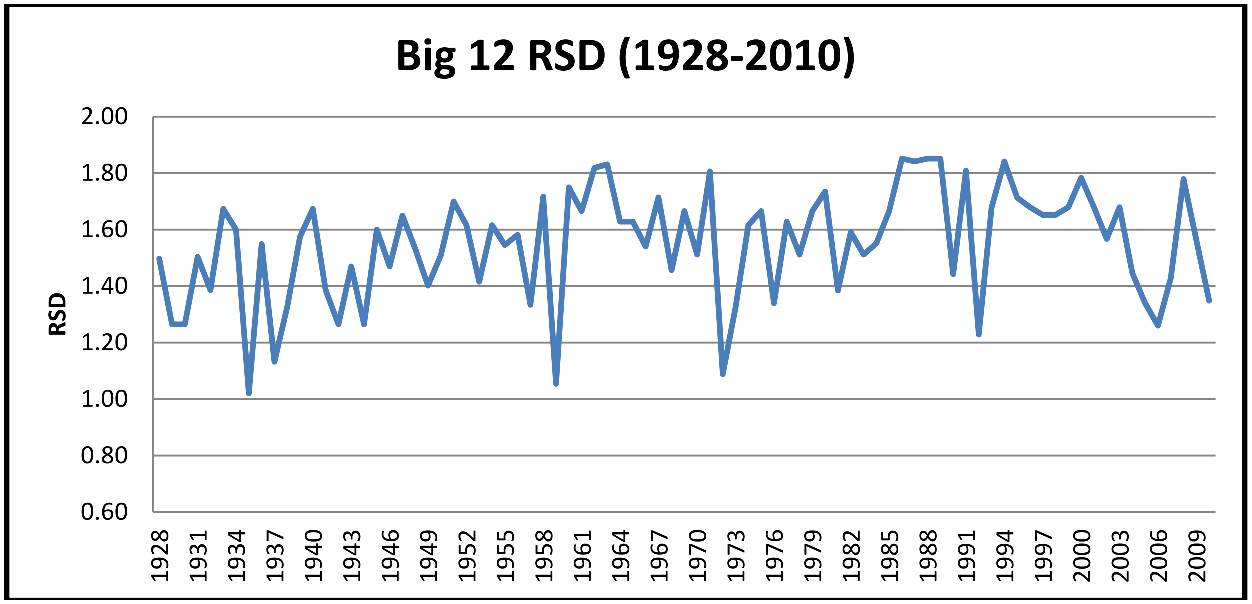
Withers, B. (2010, October 18). USC, UCLA would each receive $2 million more than other Pac-12 teams in new revenue-sharing deal, until threshold is reached. *The Seattle Times*. Retrieved December 8, 2010, from: http://seattletimes.nwso urce.com/html/husky football/2013196435_pac19.html.

Figure 4.1: Historical RSD Metric: ACC (1953-2010)



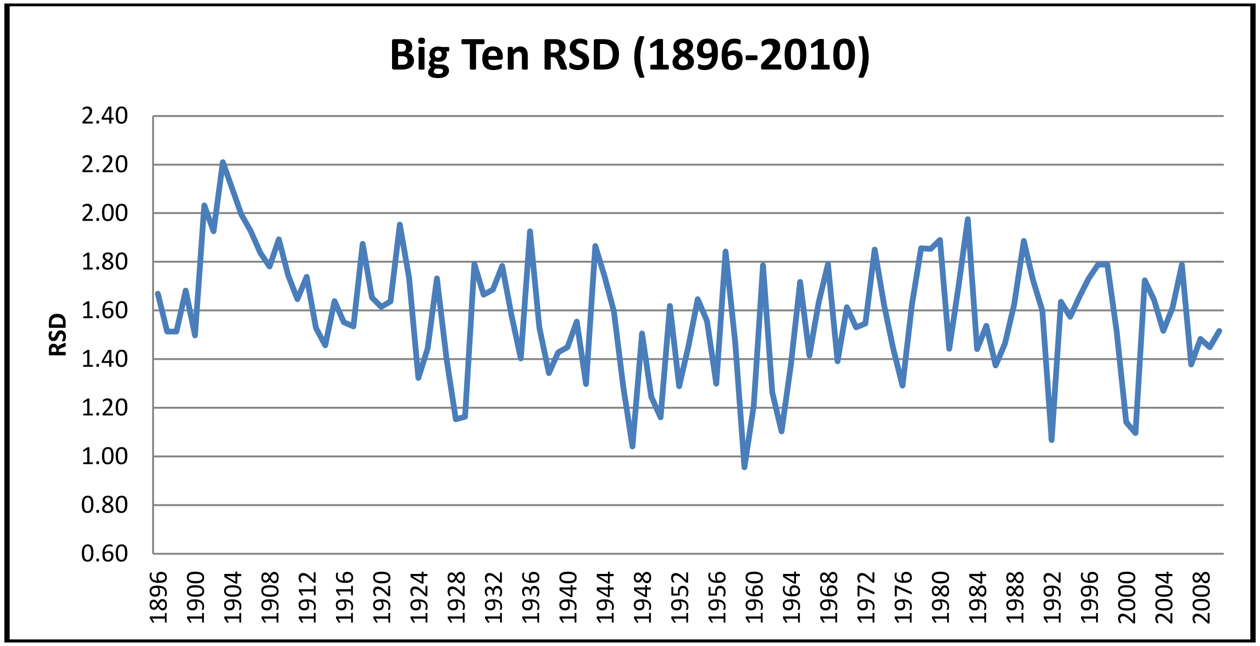
Note: Average = 1.50

Figure 4.2: Historical RSD Metric: Big 12 (1928-2010)



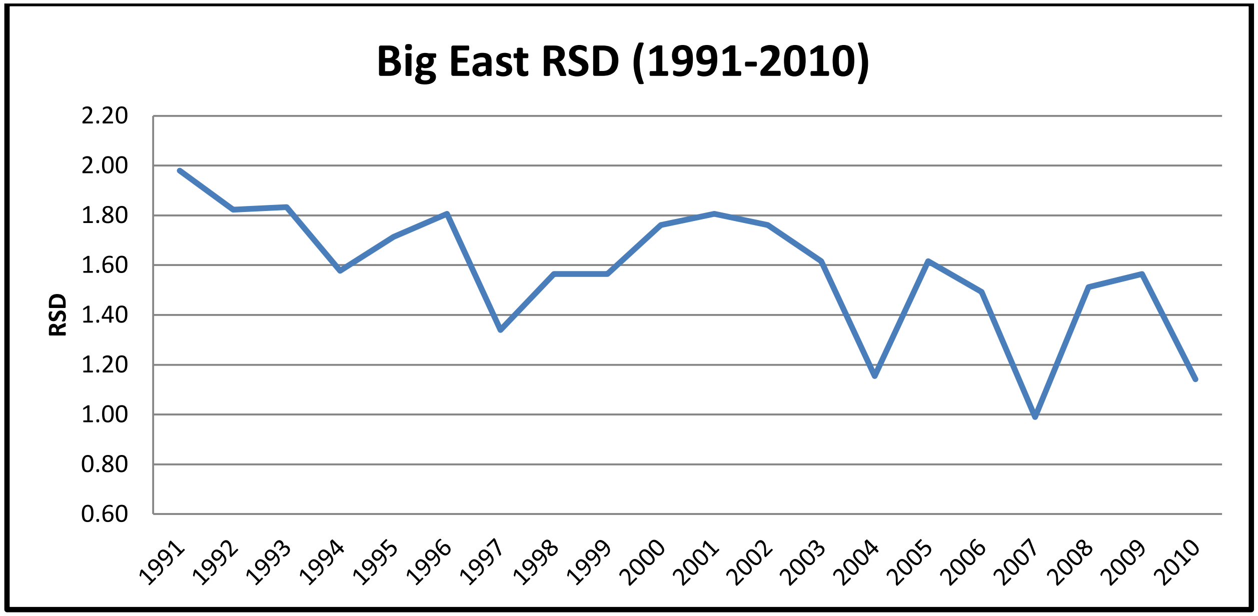
Note: Average = 1.55

Figure 4.3: Historical RSD Metric: Big Ten (1896-2010)



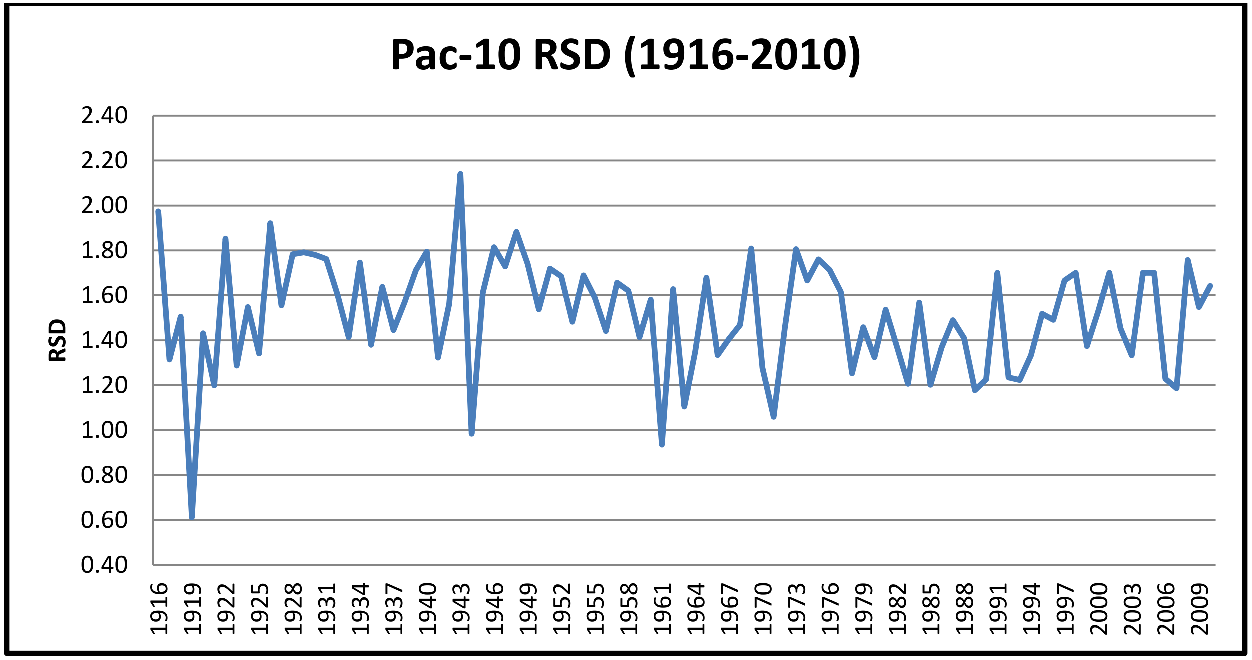
Note: Average = 1.58

Figure 4.4: Historical RSD Metric: Big East (1991-2010)



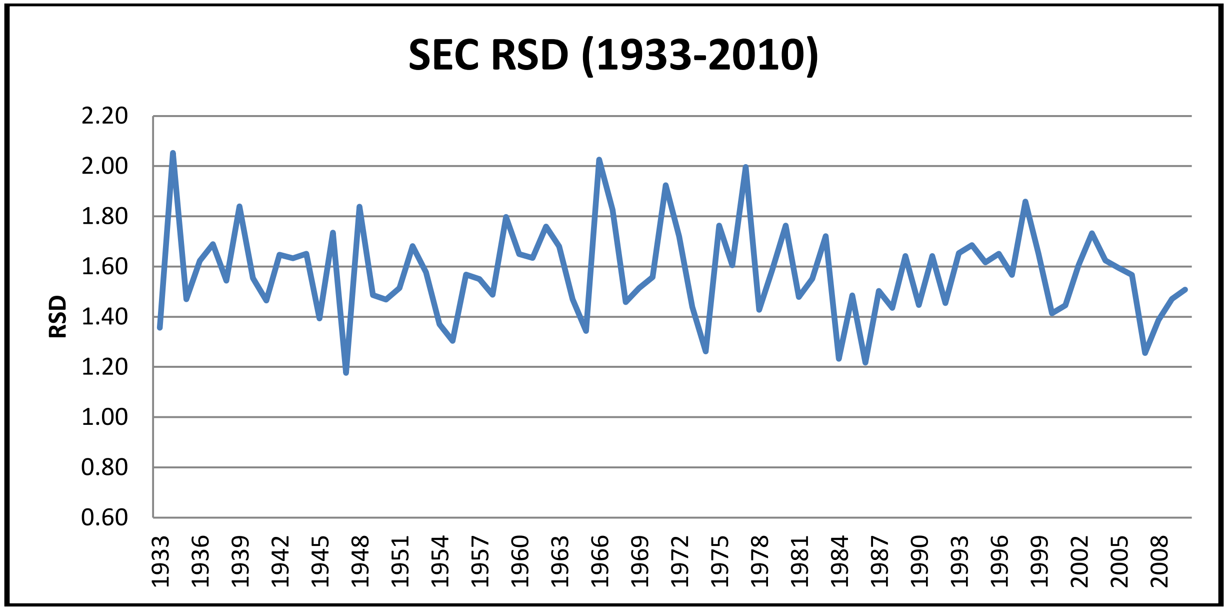
Note: Average = 1.58

Figure 4.5: Historical RSD Metric: Pac-10 (1916-2010)



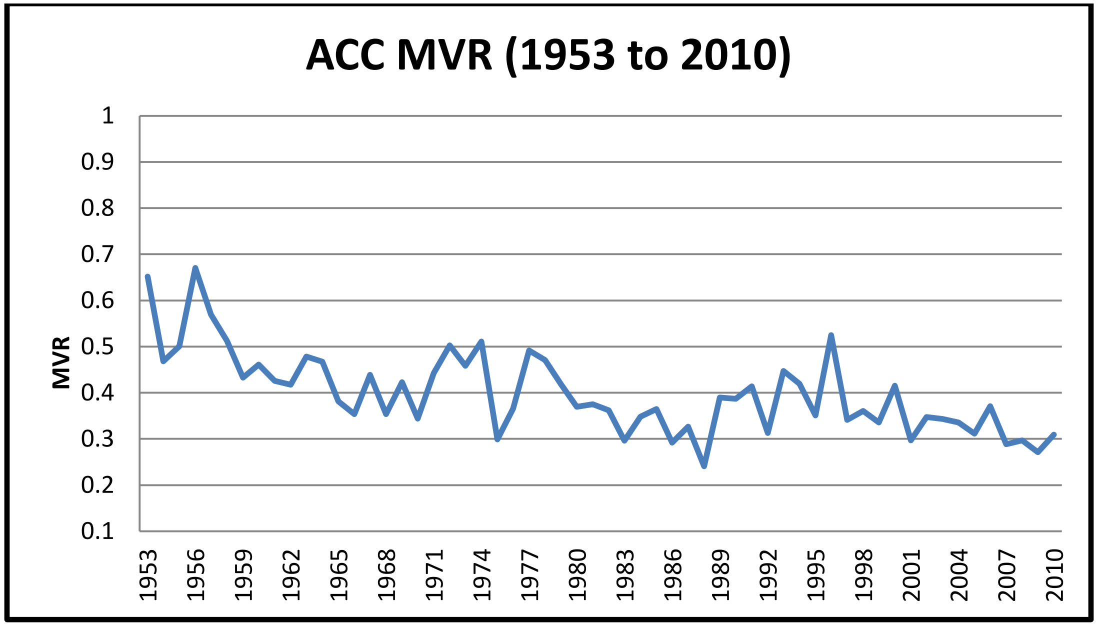
Note: Average = 1.51

Figure 4.6: Historical RSD Metric: SEC (1933-2010)



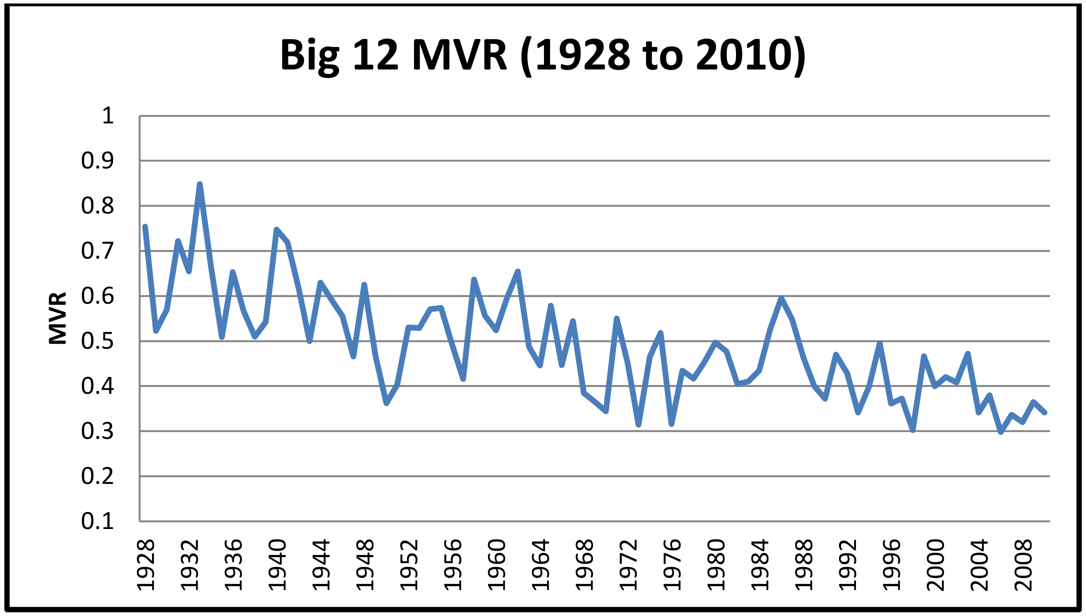
Note: Average = 1.58

Figure 4.7: Historical Margin of Victory Ratio: ACC (1953-2010)



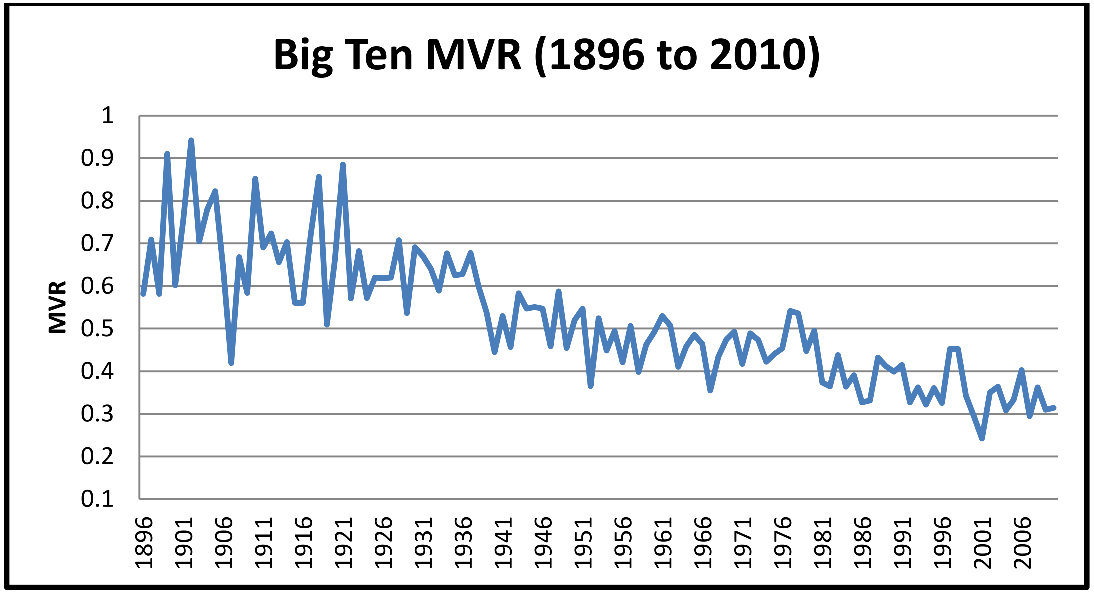
Note: Average = 0.3992

Figure 4.8: Historical Margin of Victory Ratio: Big 12 (1928-2010)



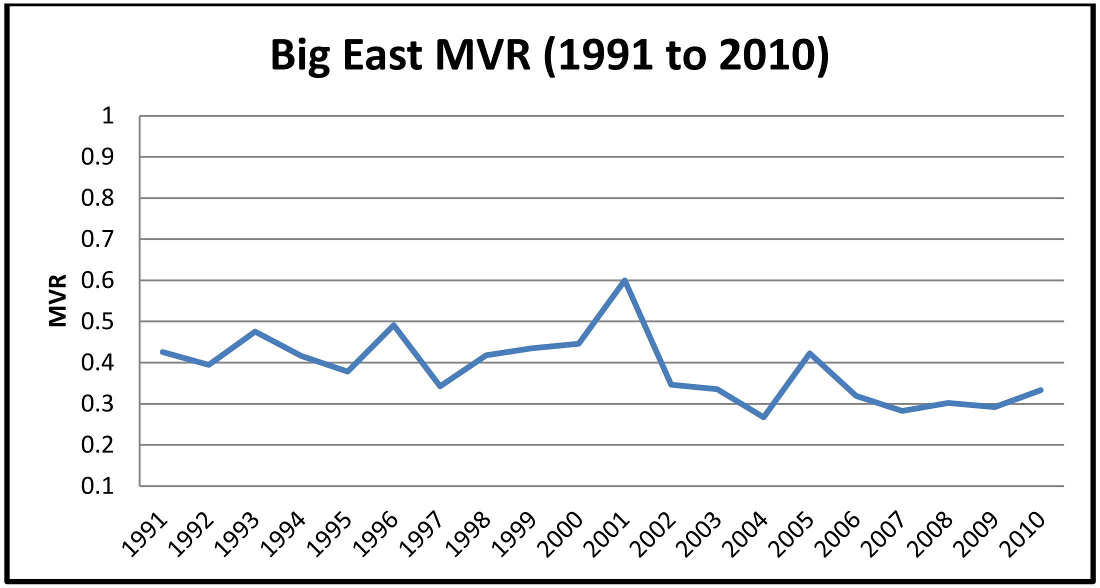
Note: Average = 0.4896

Figure 4.9: Historical Margin of Victory Ratio: Big Ten (1896-2010)



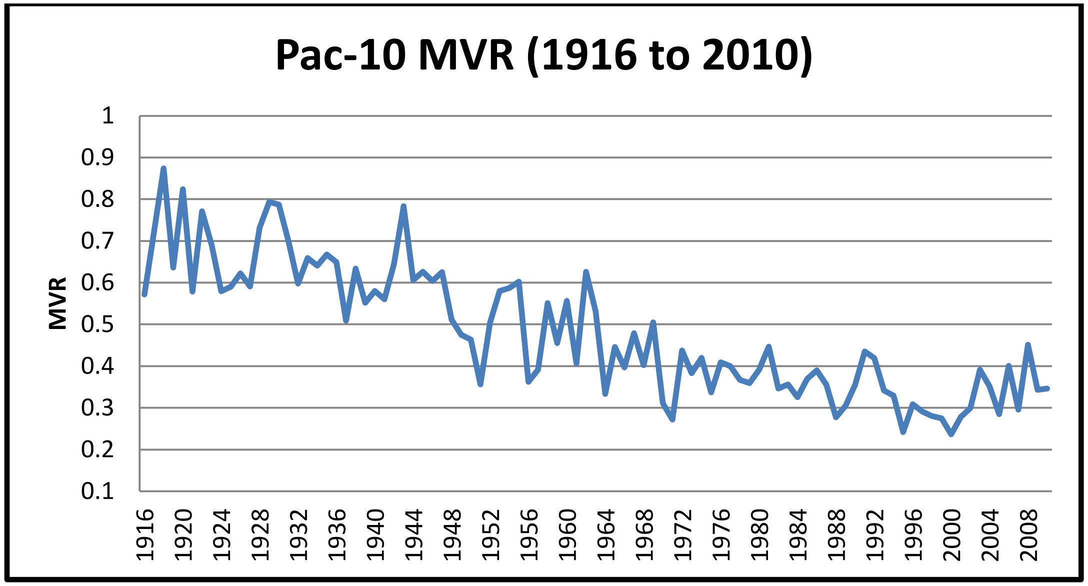
Note: Average = 0.5196

Figure 4.10: Historical Margin of Victory Ratio: Big East (1991-2010)



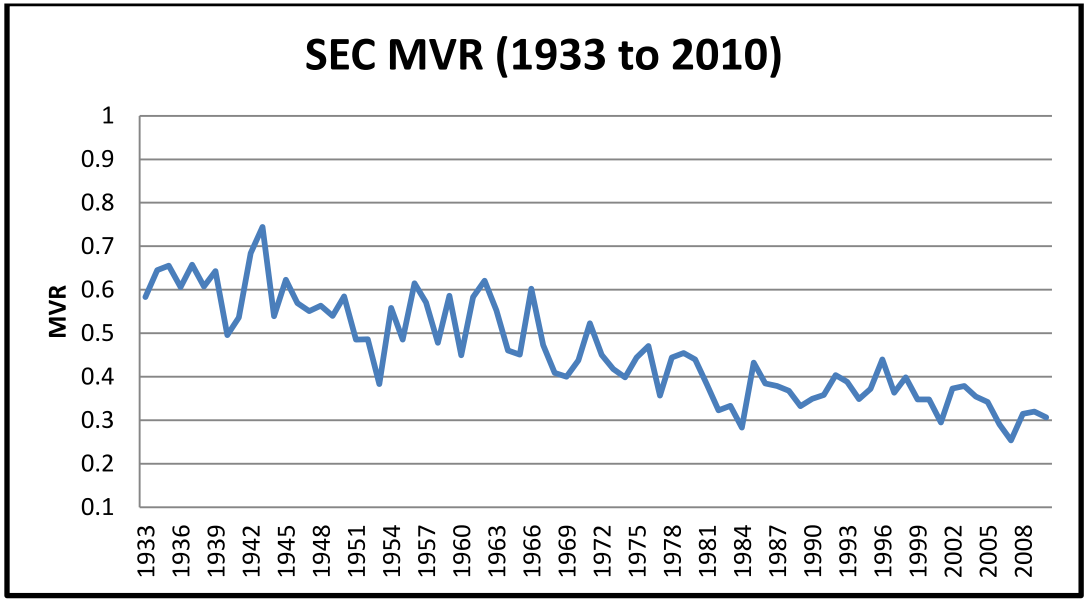
Note: Average = 0.3862

Figure 4.11: Historical Margin of Victory Ratio: Pac-10 (1916-2010)



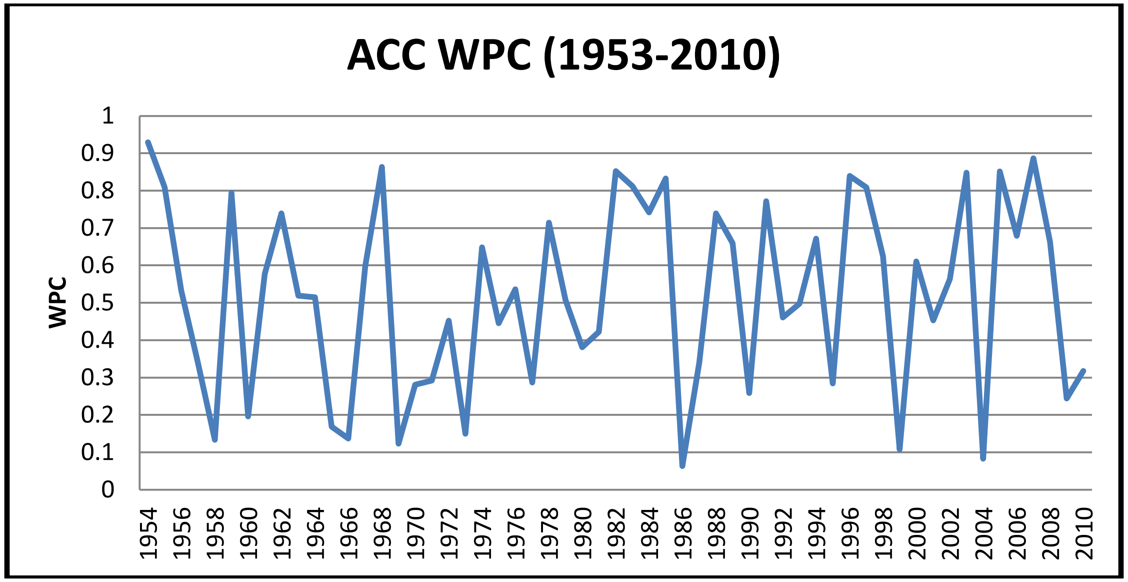
Note: Average = 0.4815

Figure 4.12: Historical Margin of Victory Ratio: SEC (1933-2010)



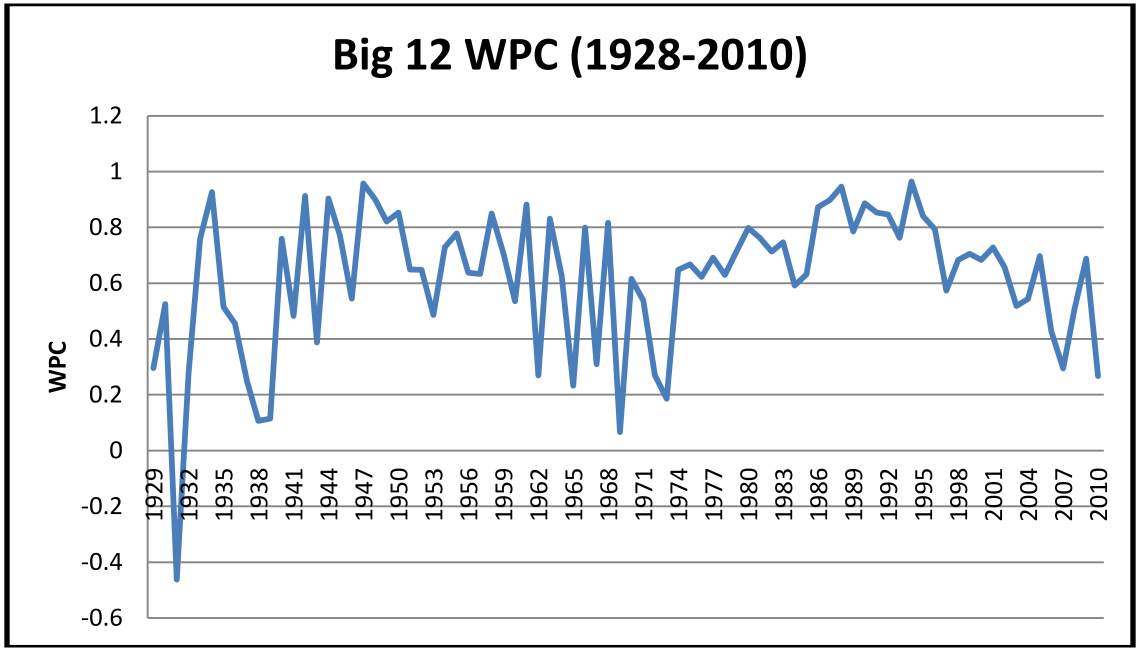
Note: Average = 0.4600

Figure 4.13: Historical Team Winning Percentage Correlation: ACC (1953-2010)



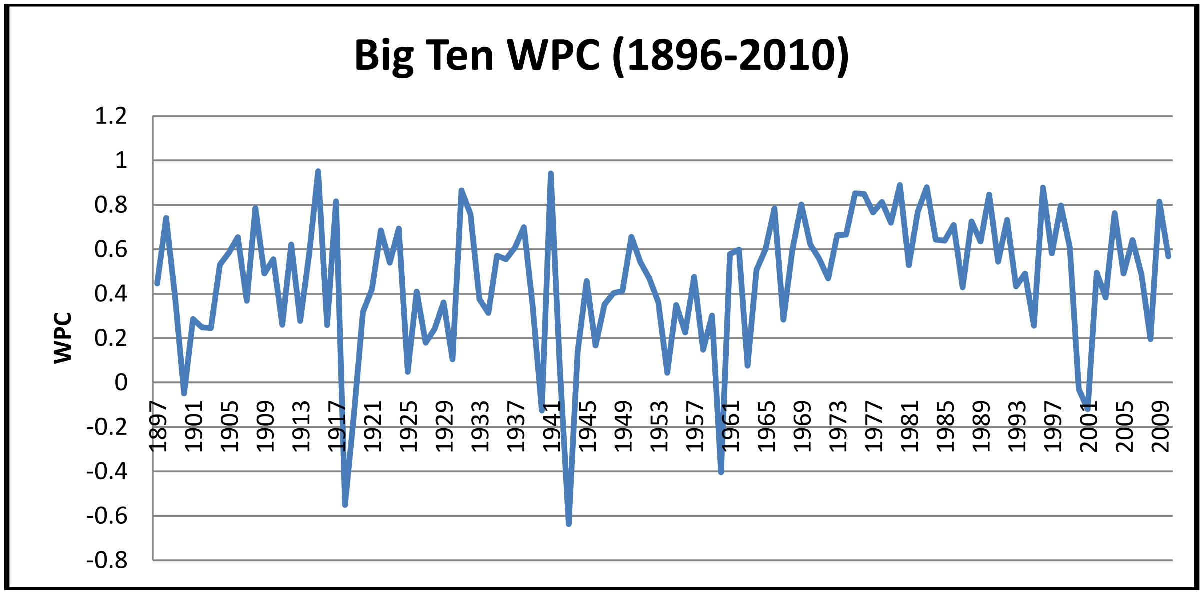
Note: Average = 0.5202

Figure 4.14: Historical Team Winning Percentage Correlation: Big 12 (1928-2010)



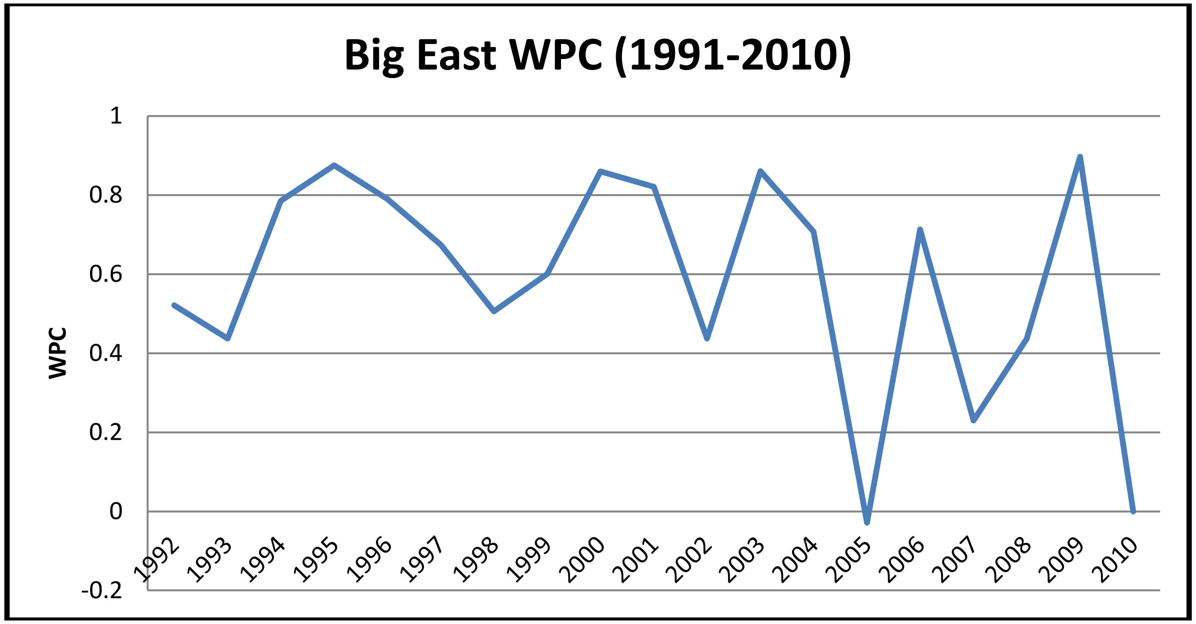
Note: Average = 0.6195

Figure 4.15: Historical Team Winning Percentage Correlation: Big Ten (1896-2010)



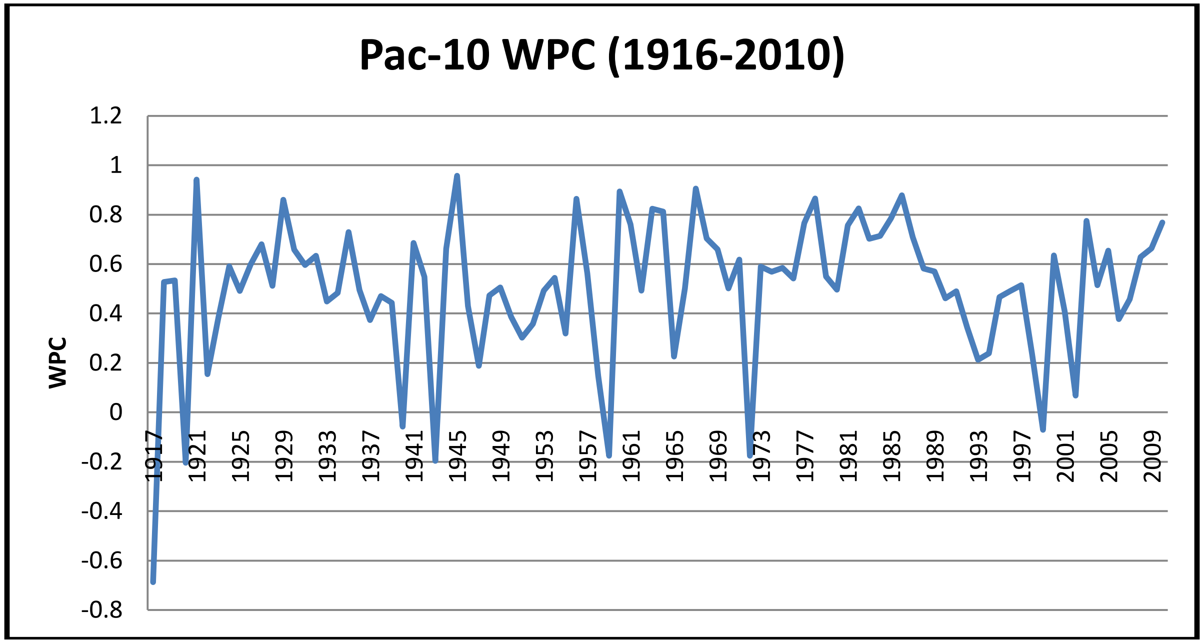
Note: Average = 0.4674

Figure 4.16: Historical Team Winning Percentage Correlation: Big East (1991-2010)



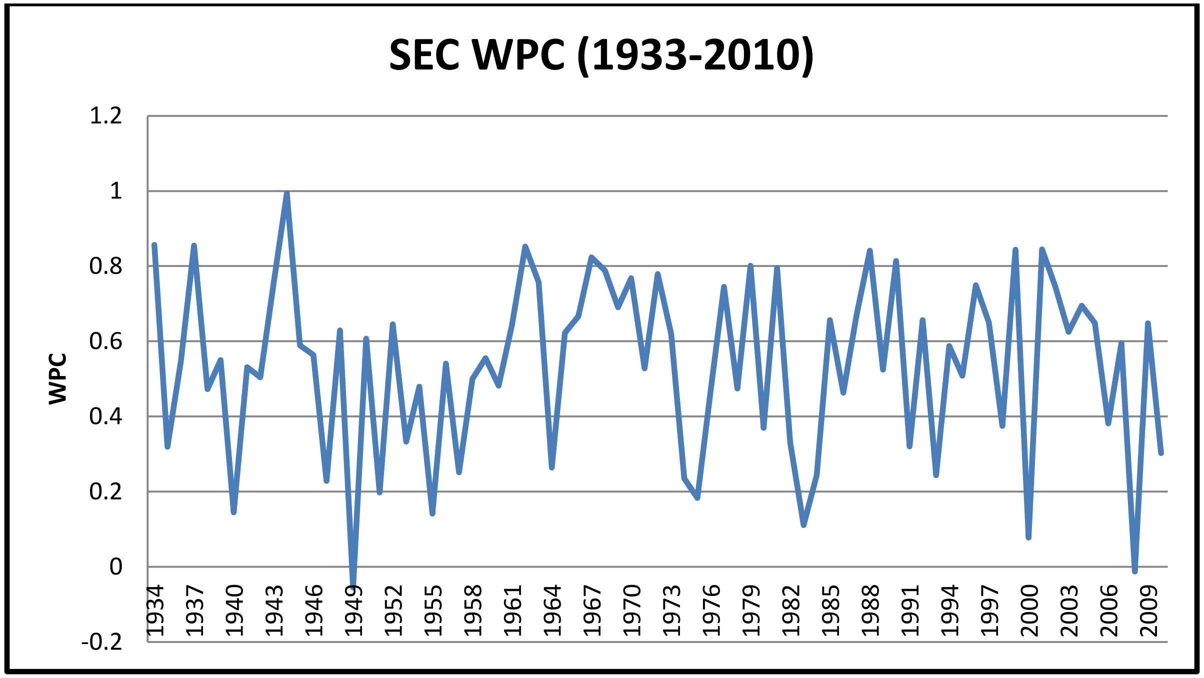
Note: Average = 0.5883

Figure 4.17: Historical Team Winning Percentage Correlation: Pac-10 (1916-2010)



Note: Average = 0.5025

Figure 4.18: Historical Team Winning Percentage Correlation: SEC (1933-2010)



Note: Average = 0.5356

Figure 4.19: ACC RSD: Actual versus Fitted (1953-2010)

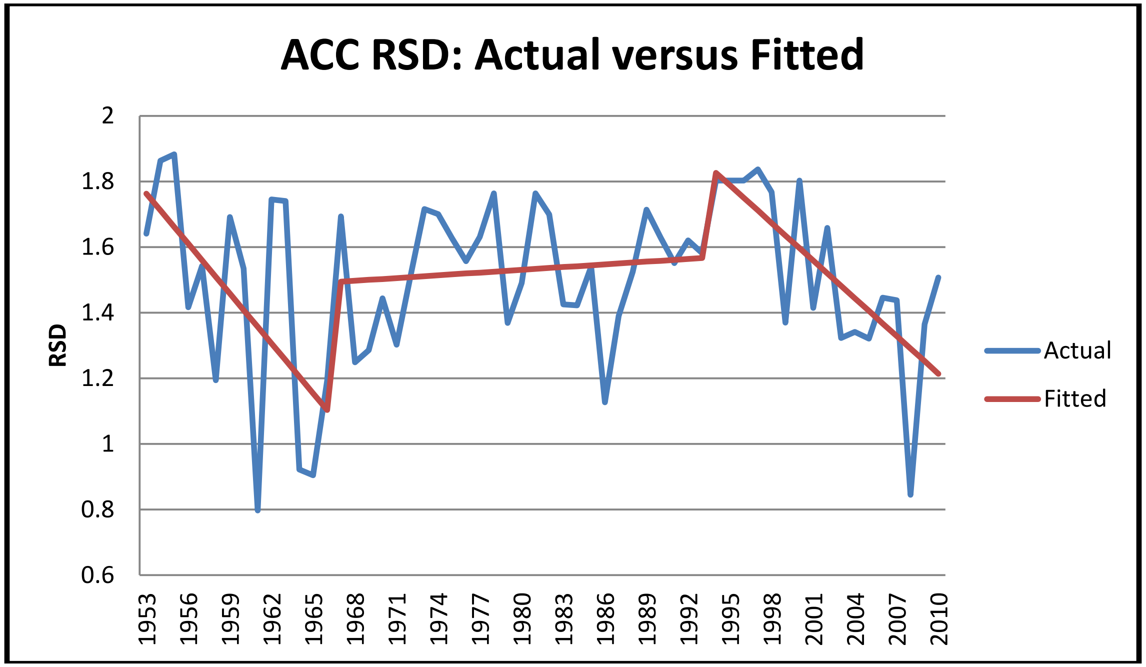
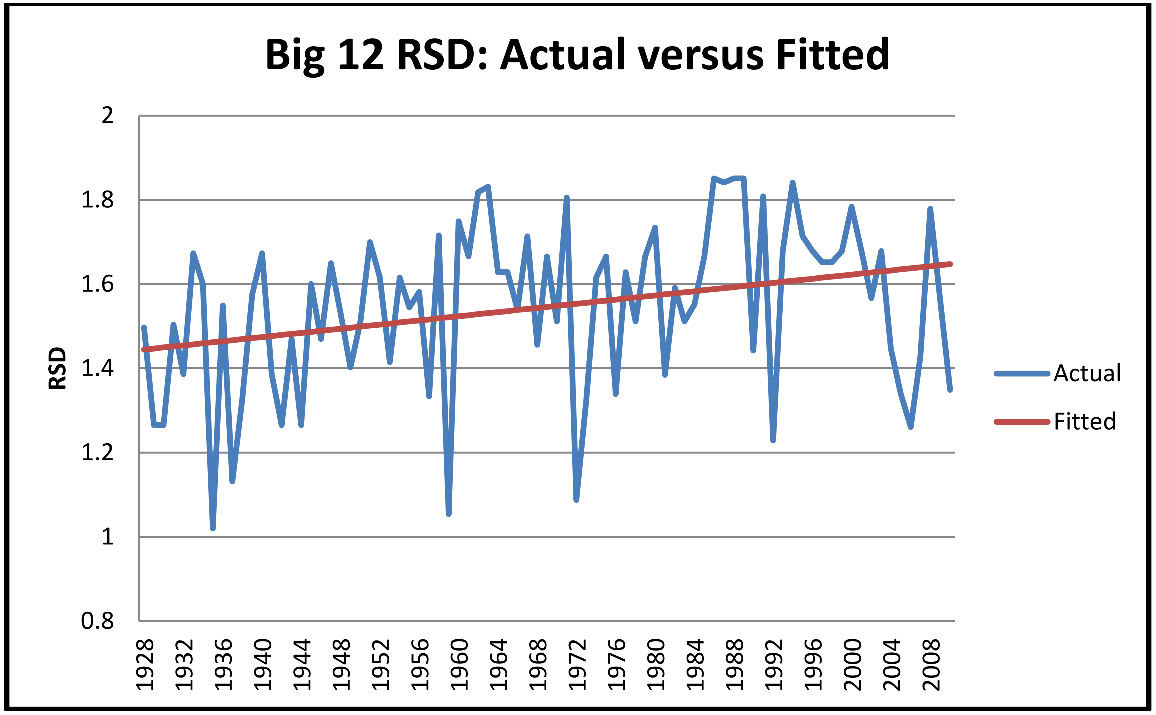


Figure 4.20: Big 12 RSD: Actual versus Fitted (1928-2010)



Note: Stationary without breaks.

Figure 4.21: Big Ten RSD: Actual versus Fitted (1896-2010)

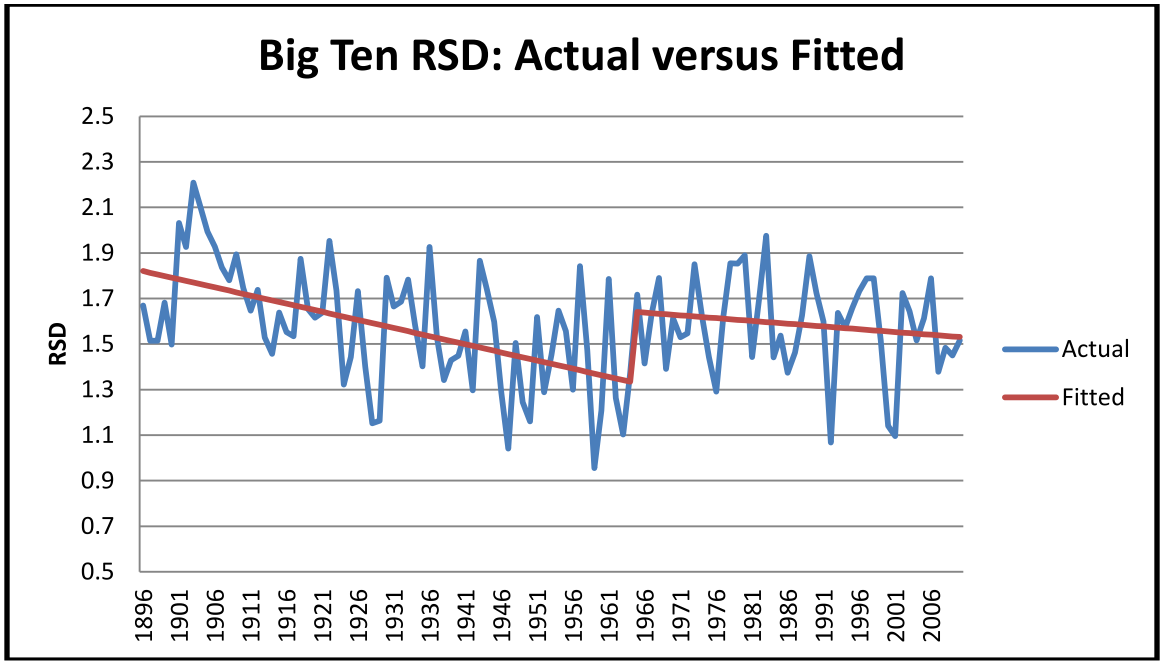
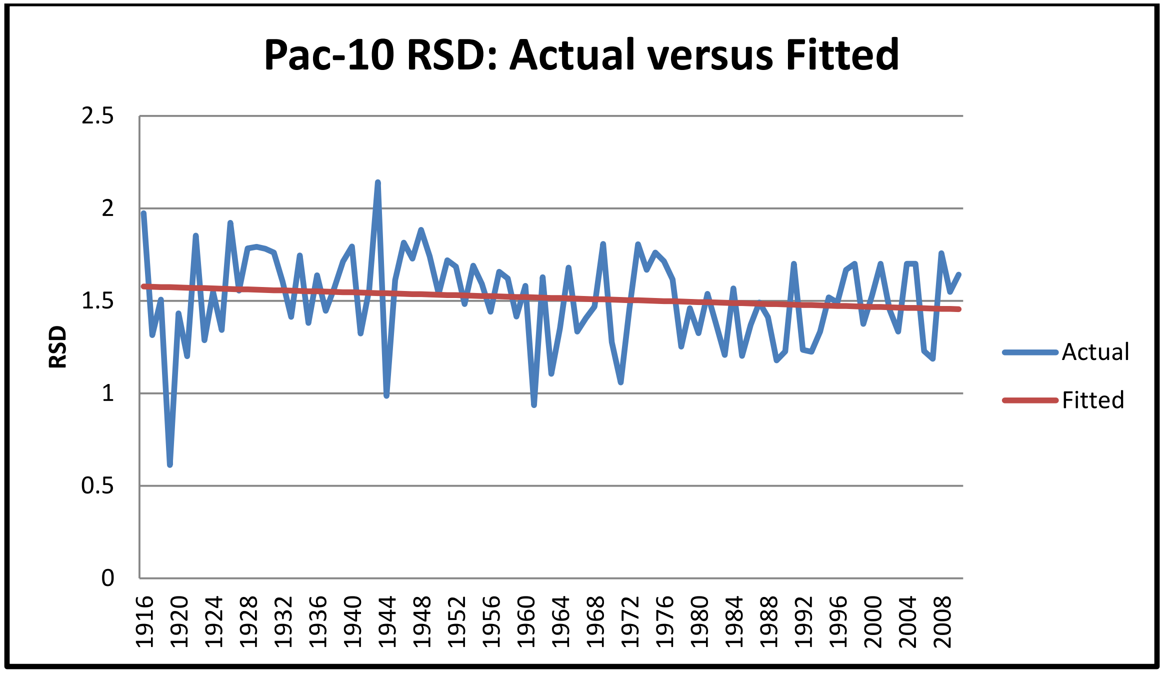
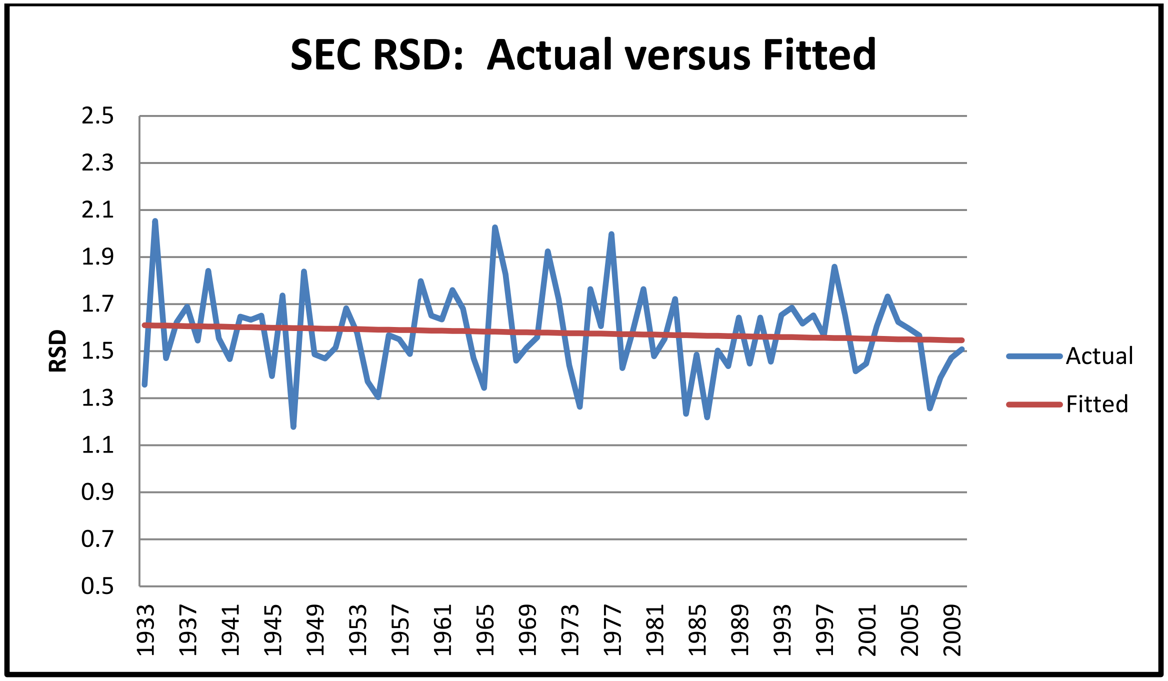


Figure 4.22: Pac-10 RSD: Actual versus Fitted (1916-2010)



Note: Stationary without breaks.

Figure 4.23: SEC RSD: Actual versus Fitted (1933-2010)



Note: Stationary without breaks.

Figure 4.24: ACC MVR: Actual versus Fitted (1953-2010)

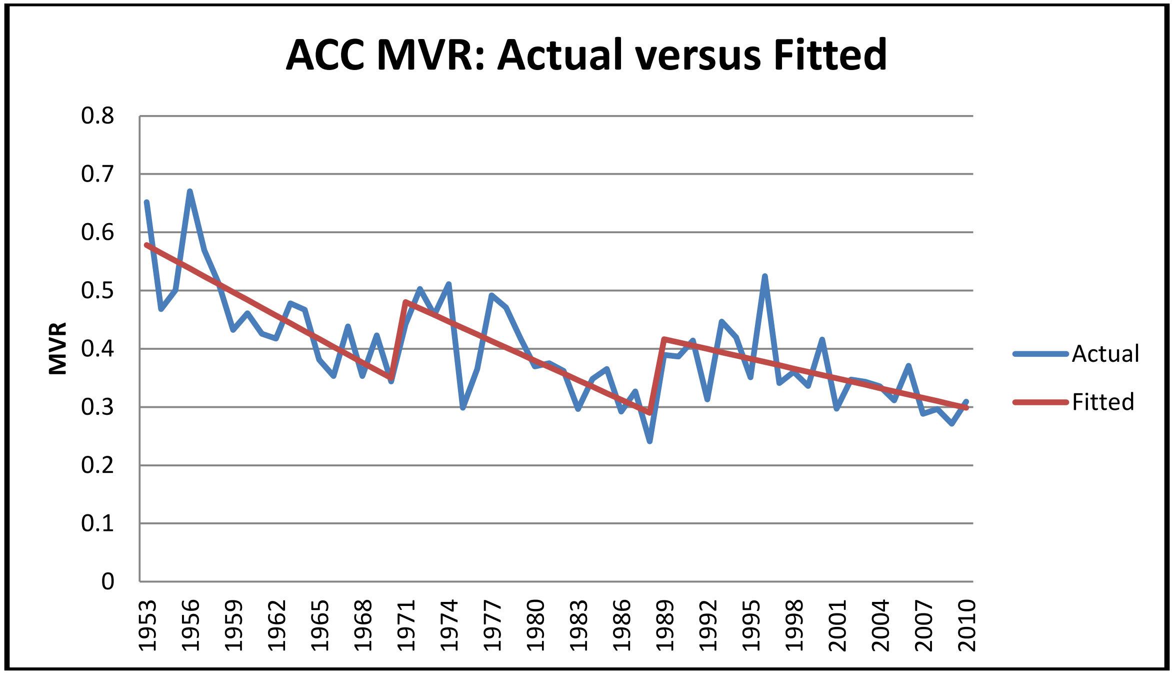
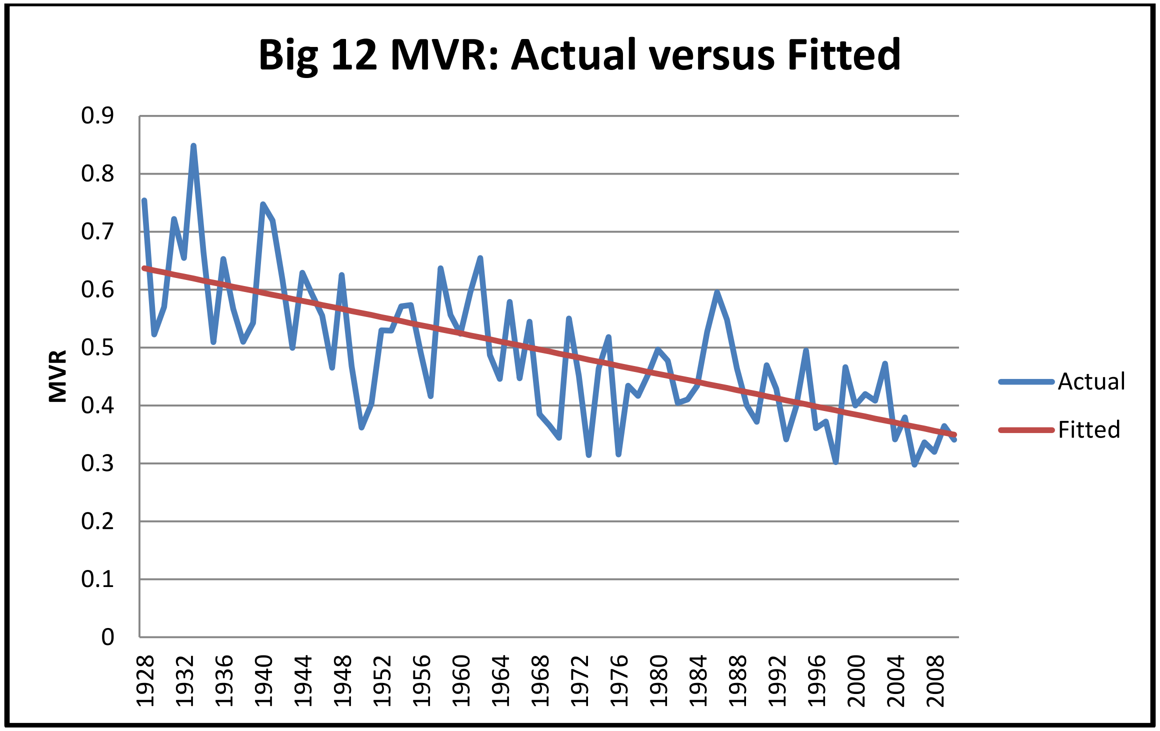
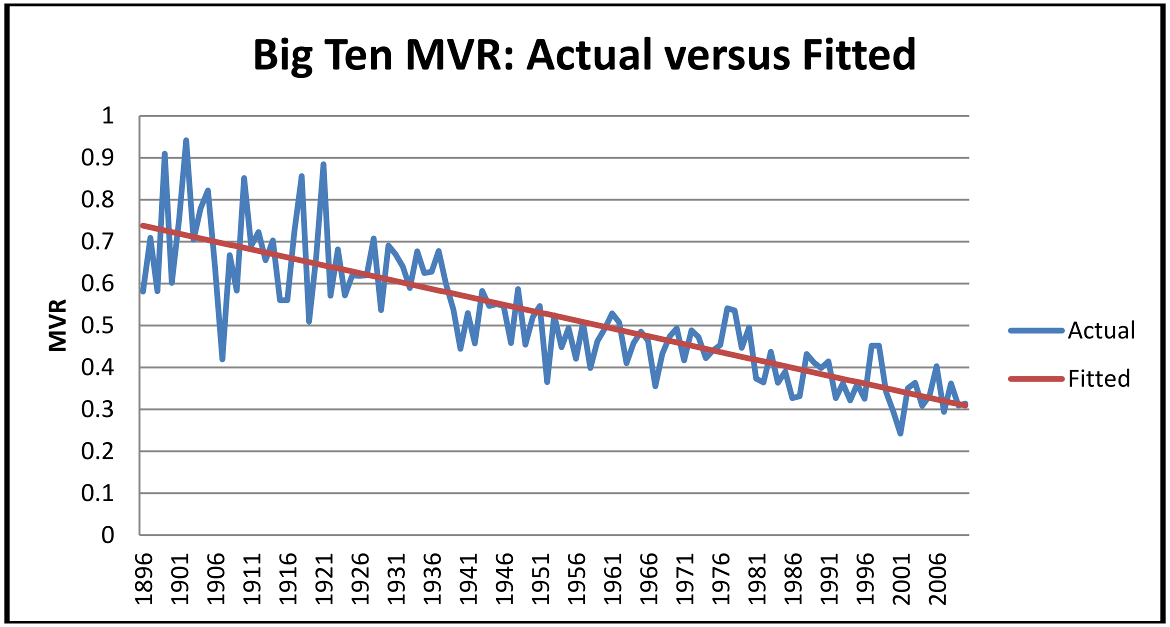


Figure 4.25: Big 12 MVR: Actual versus Fitted (1928-2010)



Note: Stationary without breaks.

Figure 4.26: Big Ten MVR: Actual versus Fitted (1896-2010)



Note: Stationary without breaks.

Figure 4.27: Pac-10 MVR: Actual versus Fitted (1916-2010)

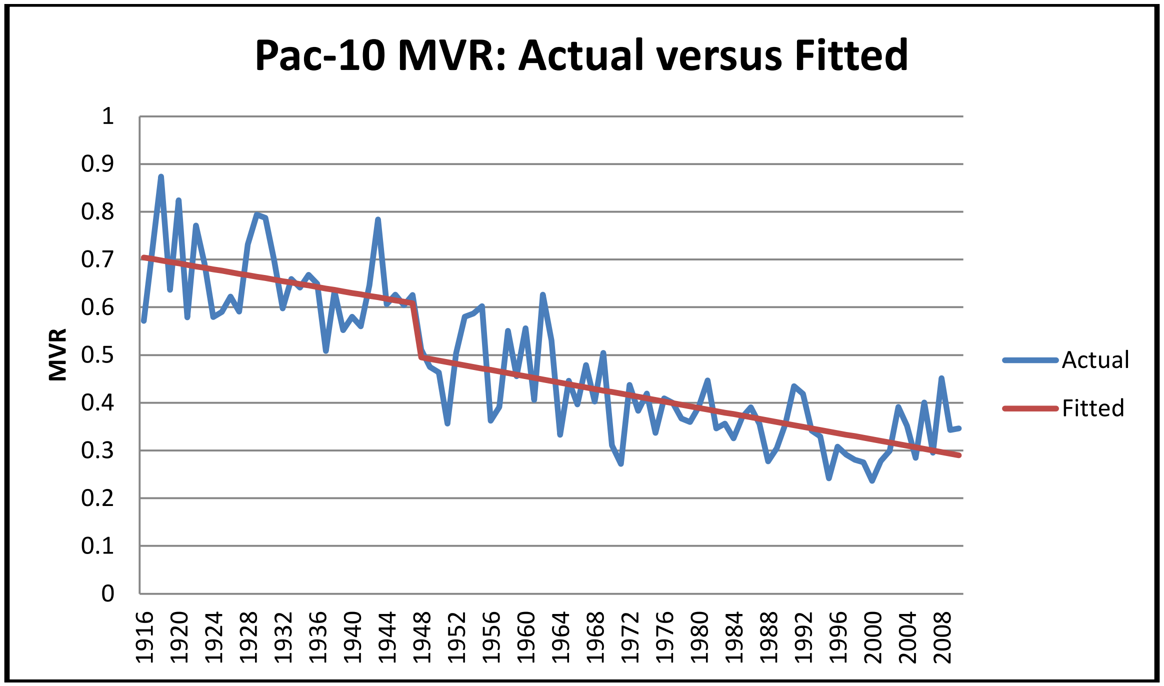
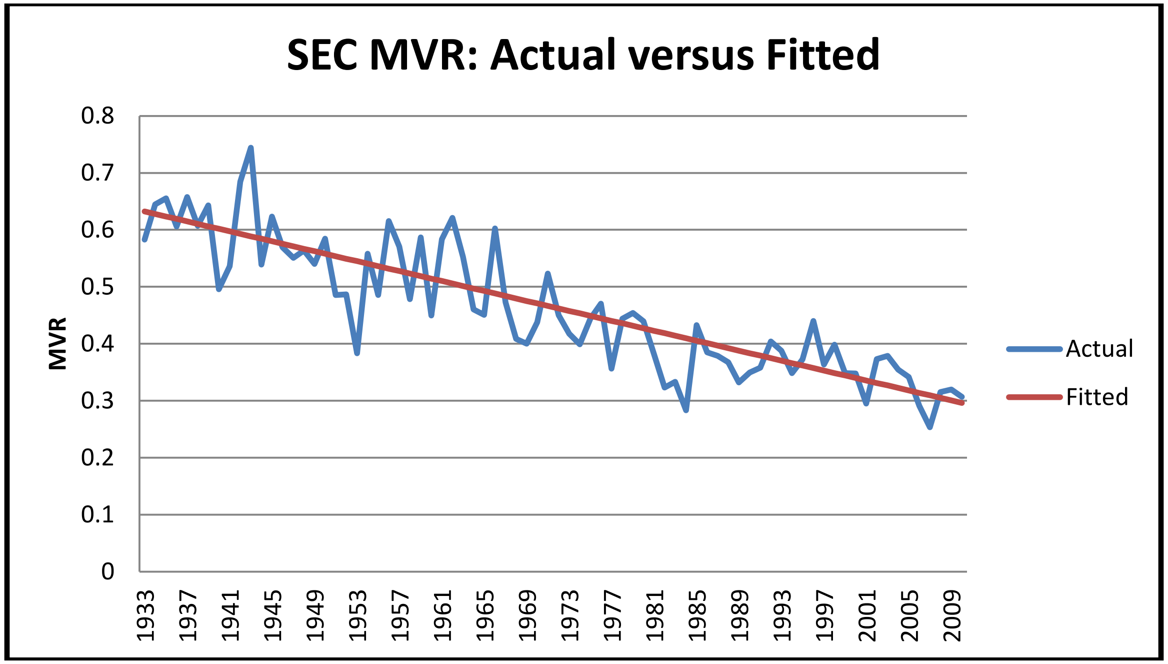
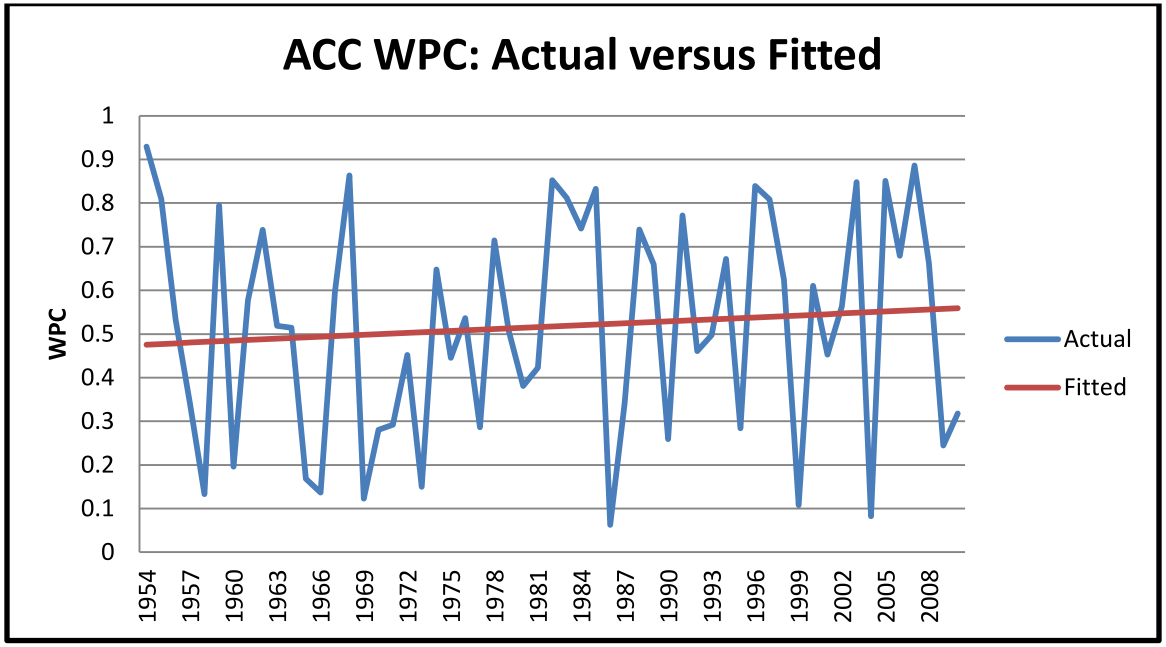


Figure 4.28: SEC MVR: Actual versus Fitted (1933-2010)



Note: Stationary without breaks.

Figure 4.29: ACC WPC: Actual versus Fitted (1953-2010)



Note: Stationary without breaks.

Figure 4.30: Big 12 WPC: Actual versus Fitted (1928-2010)

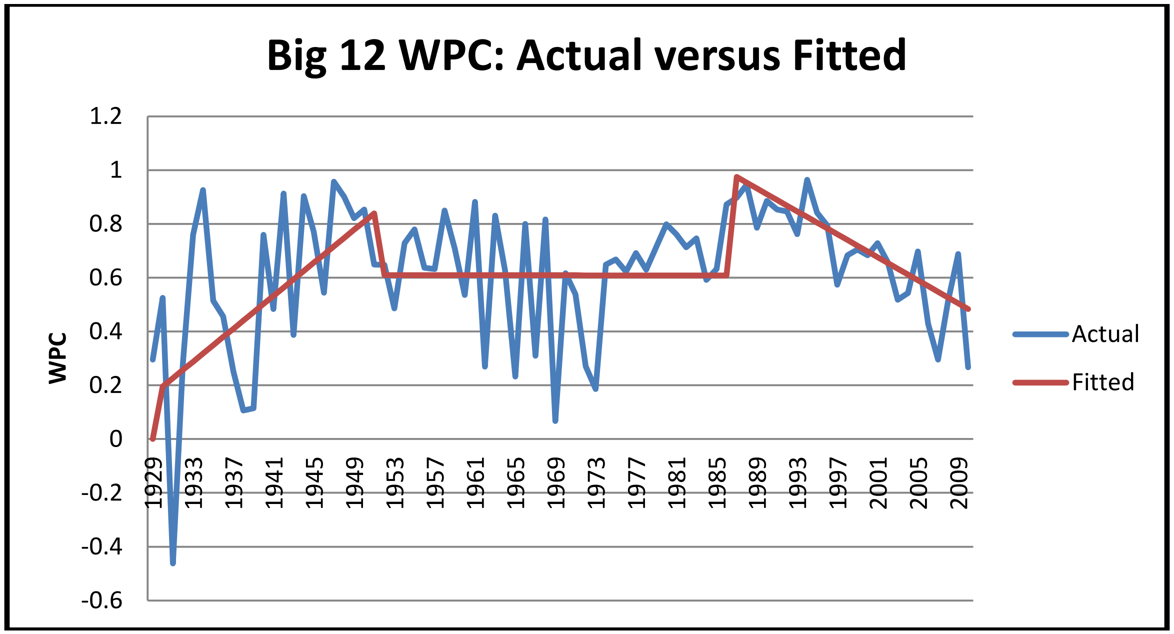


Figure 4.31: Big Ten WPC: Actual versus Fitted (1896-2010)

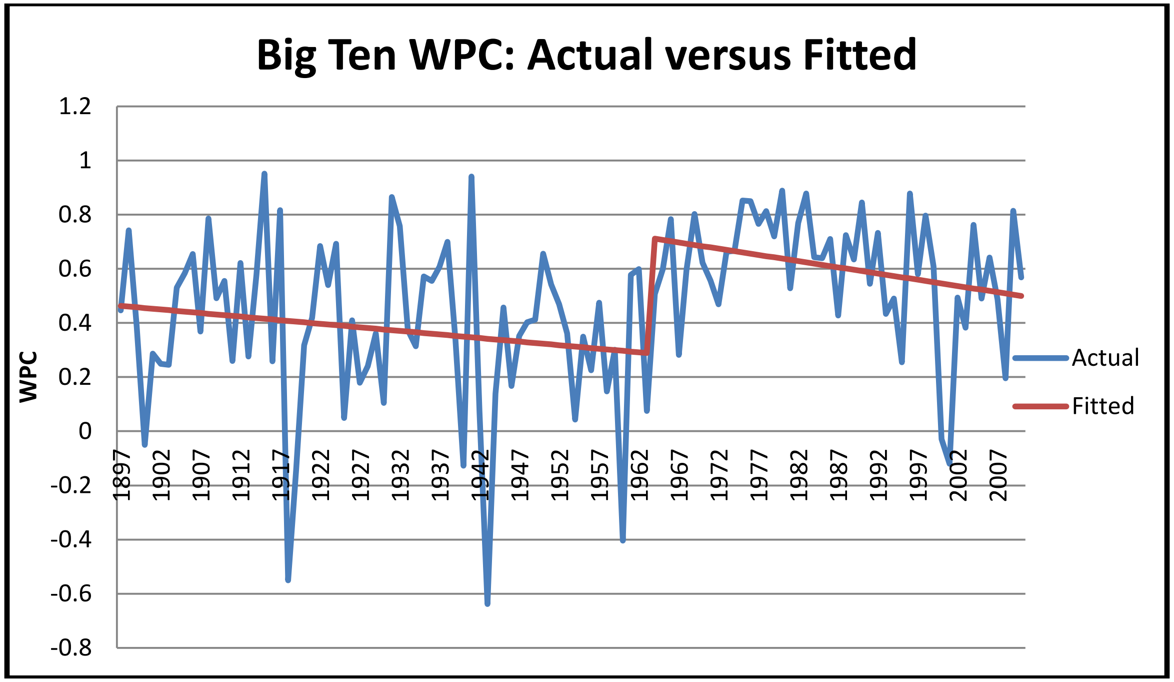
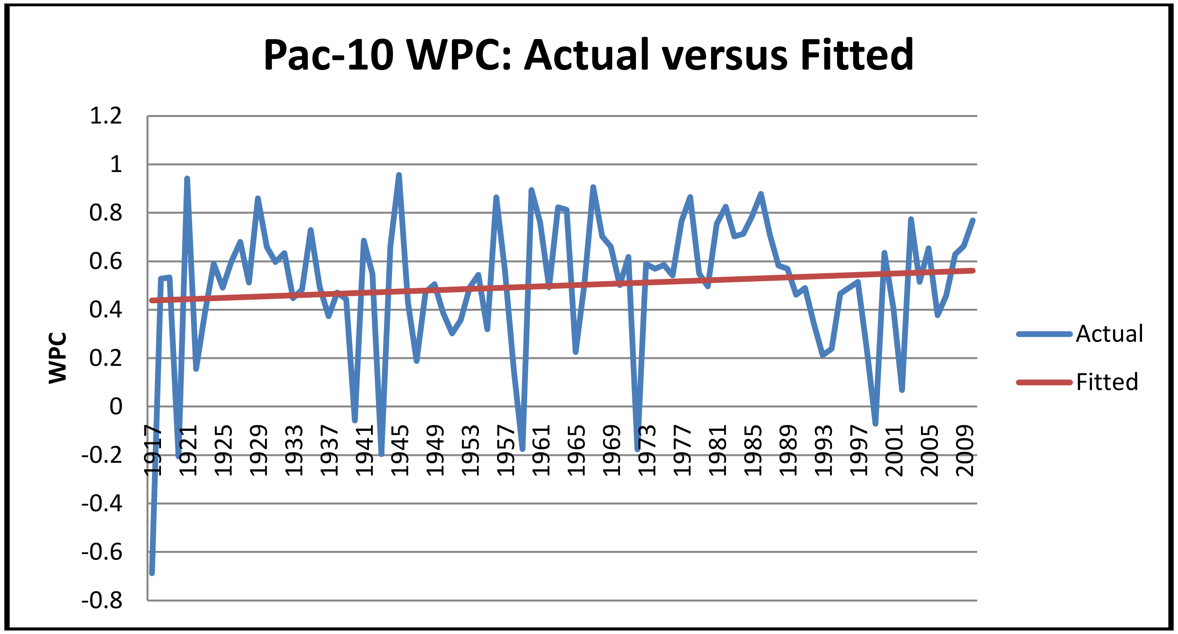
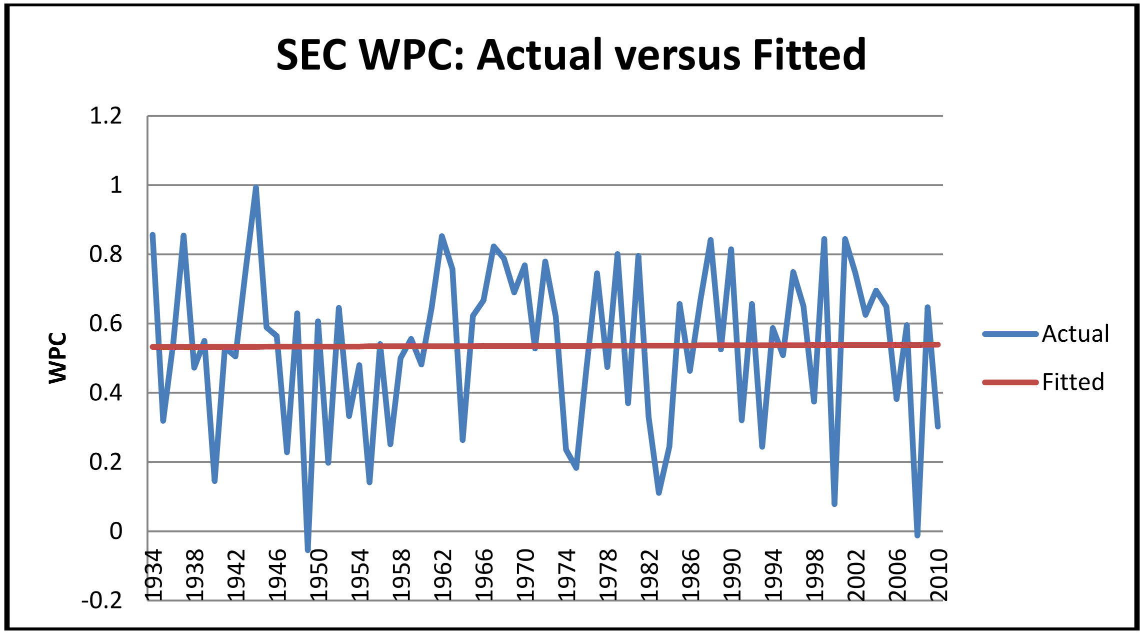


Figure 4.32: Pac-10 WPC: Actual versus Fitted (1916-2010)



Note: Stationary without breaks.

Figure 4.33: SEC WPC: Actual versus Fitted (1933-2010)



Note: Stationary without breaks.

4.

Table 4.1: NCAA Football Broadcasting Revenue Sharing Arrangements (FY 2009-2010)

	% of Television Revenue Shared Equally	% of Television Revenue Shared Based on Number of TV Appearances
ACC[a]	100	0
Big 10[b]	100	0
Big 12[c]	50	50
Big East[d]	100	0
Pac-10[e]	45	55
SEC[f]	100	0

Notes: [a] Wellman (2010); [b] Lesmerises (2010); [c] Ubben (2010); [d] Furfari (2010); [e] Wilner (2010); [f] Withers (2010)

Table 4.2: Playing History by Conference

Conference	Dates
ACC	1953-present
Big 12	1928-present
Big Ten	1896-present
Big East	1991-present
Pac-10	1916-present
SEC	1933-present

Table 4.3: Conference RSD Averages by Decade

	ACC	Big Ten	Big 12	Big East	Pac-10	SEC
1890-1899	-	1.59	-	-	-	-
1900-1909	-	1.92	-	-	-	-
1910-1919	-	1.64	-	-	1.35	-
1920-1929	-	1.52	1.38	-	1.57	-
1930-1939	-	1.61	1.40	-	1.60	1.65
1940-1949	-	1.46	1.47	-	1.66	1.56
1950-1959	1.60	1.43	1.51	-	1.58	1.53
1960-1969	1.31	1.47	1.67	-	1.43	1.64
1970-1979	1.56	1.62	1.52	-	1.51	1.63
1980-1989	1.51	1.63	1.68	-	1.37	1.50
1990-1999	1.68	1.61	1.64	1.69	1.45	1.62
2000-2010	1.41	1.48	1.53	1.49	1.52	1.51
All Years	1.50	1.58	1.55	1.58	1.51	1.58

Note: Initial values for each conference represent averages for truncated decades based on when each conference began play.

Table 4.4: Conference RSD Averages by Time Period

	ACC	Big Ten	Big 12	Big East	Pac-10	SEC
1950 - 2010	1.51	1.54	1.59	-	1.48	1.57
1960 - 2010	1.49	1.56	1.61	-	1.45	1.58
1970 - 2010	1.54	1.59	1.59	-	1.46	1.57
1980 - 2010	1.53	1.57	1.62	-	1.44	1.55
1990 - 2010	1.54	1.55	1.58	1.59	1.48	1.57

Note: The Big East did not begin play until 1991.

Table 4.5: Conference MVR Averages by Decade

	ACC	Big Ten	Big 12	Big East	Pac-10	SEC
1890-1899	-	0.6955	-	-	-	-
1900-1909	-	0.6914	-	-	-	-
1910-1919	-	0.6835	-	-	0.7009	-
1920-1929	-	0.6472	0.6382	-	0.6773	-
1930-1939	-	0.6334	0.6241	-	0.6396	0.6281
1940-1949	-	0.5158	0.5917	-	0.6018	0.5847
1950-1959	0.5436	0.4686	0.5069	-	0.4853	0.5235
1960-1969	0.4200	0.4609	0.5027	-	0.4681	0.5005
1970-1979	0.4303	0.4711	0.4262	-	0.3696	0.4397
1980-1989	0.3367	0.3927	0.4757	-	0.3564	0.3657
1990-1999	0.3894	0.3757	0.4007	0.4198	0.3277	0.3770
2000-2010	0.3260	0.3248	0.3710	0.3588	0.3344	0.3253
All Years	0.3992	0.5196	0.4896	0.3862	0.4815	0.4600

Note: Initial values for each conference represent averages for truncated decades based on when each conference began play.

Table 4.6: Conference MVR Averages by Time Period

	ACC	Big Ten	Big 12	Big East	Pac-10	SEC
1950 - 2010	0.3992	0.4141	0.4459	-	0.3893	0.4203
1960 - 2010	0.3794	0.4035	0.434	-	0.3705	0.4001
1970 - 2010	0.3695	0.3895	0.4172	-	0.3467	0.3757
1980 - 2010	0.3499	0.3631	0.4143	-	0.3393	0.3550
1990 - 2010	0.3562	0.3490	0.3851	0.3862	0.3312	0.3499

Note: The Big East did not begin play until 1991.

Table 4.7: Conference WPC Averages by Decade

	ACC	Big Ten	Big 12	Big East	Pac-10	SEC
1890-1899	-	0.5244	-	-	-	-
1900-1909	-	0.4145	-	-	-	-
1910-1919	-	0.3625	-	-	0.1244	-
1920-1929	-	0.3895	0.2958	-	0.5007	-
1930-1939	-	0.5187	0.3469	-	0.5331	0.5999
1940-1949	-	0.2182	0.7440	-	0.4194	0.4885
1950-1959	0.5899	0.3572	0.6970	-	0.3801	0.4251
1960-1969	0.4435	0.4422	0.5369	-	0.6780	0.6591
1970-1979	0.4312	0.6981	0.5584	-	0.5408	0.5607
1980-1989	0.5845	0.6847	0.7747	-	0.7017	0.5006
1990-1999	0.5324	0.6167	0.7909	0.6486	0.3382	0.5749
2000-2010	0.5635	0.4262	0.5469	0.5392	0.5405	0.5048
All Years	0.5202	0.4674	0.6195	0.5853	0.5025	0.5356

Note: Initial values for each conference represent averages for truncated decades based on when each conference began play.

Table 4.8: Conference WPC Averages by Time Period

	ACC	Big Ten	Big 12	Big East	Pac-10	SEC
1950 - 2010	0.5202	0.5357	0.6491	-	0.5301	0.5370
1960 - 2010	0.5121	0.5707	0.6397	-	0.5595	0.5589
1970 - 2010	0.5288	0.6020	0.6648	-	0.5306	0.5345
1980 - 2010	0.5602	0.5710	0.6991	-	0.5273	0.5260
1990 - 2010	0.5487	0.5169	0.6631	0.5853	0.4442	0.5382

Note: The Big East did not begin play until 1991.

Table 4.9: ACC Conference Championship History

Team	First Year	Most Recent Year	Total Years	Championships	Championship Ties	Total Championships	% Championships/Year	YPC
Boston College	2005	2010	6	0	0	0.00	0.00	***
Clemson	1953	2010	58	12	0	12.00	0.21	4.83
Duke	1953	2010	58	4	4	6.00	0.10	9.67
Florida State	1992	2010	19	10	2	11.00	0.58	1.73
Georgia Tech	1983	2010	28	2	1	2.50	0.09	11.20
Maryland	1953	2010	58	7	2	8.00	0.14	7.25
Miami (FL)	2004	2010	7	0	0	0.00	0.00	***
North Carolina	1953	2010	58	4	1	4.50	0.08	12.89
North Carolina State	1953	2010	58	5	1	5.50	0.09	10.55
South Carolina	1953	1971	19	1	1	1.50	0.08	12.67
Virginia	1954	2010	57	0	2	1.00	0.02	57.00
Virginia Tech	2004	2010	7	4	0	4.00	0.57	1.75
Wake Forest	1953	2010	58	2	0	2.00	0.03	29.00

Note: The Championship Ties column includes two-way championship ties, which causes the Total Championships column to produce fractional values.

Table 4.10: Big 12 Conference Championship History

Team	First Year	Most Recent Year	Total Years	Championships	Championship Ties	Total Championships	% Championships/Year	YPC
Baylor	1996	2010	15	0	0	0.00	0.00	***
Colorado	1948	2010	63	4	2	4.83	0.08	13.04
Iowa State	1928	2010	83	0	0	0.00	0.00	***
Kansas	1928	2010	83	1	3	2.50	0.03	33.20
Kansas State	1928	2010	83	2	0	2.00	0.02	41.50
Missouri	1928	2010	83	5	1	5.50	0.07	15.09
Nebraska	1928	2010	83	26	5	28.50	0.34	2.91
Oklahoma	1928	2010	83	32	7	36.33	0.44	2.28
Oklahoma State	1960	2010	51	0	1	0.33	0.01	154.55
Texas	1996	2010	15	3	0	3.00	0.20	5.00
Texas A&M	1996	2010	15	1	0	0.00	0.00	***
Texas Tech	1996	2010	15	0	0	0.00	0.00	***

Note: The Championship Ties column includes two-way championship ties, which causes the Total Championships column to produce fractional values.

Table 4.11: Big Ten Conference Championship History

Team	First Year	Most Recent Year	Total Years	Championships	Championship Ties	Total Championships	% Championships/Year	YPC
Chicago	1896	1939	44	6	0	6.00	0.14	7.30
Illinois	1896	2010	115	10	4	11.75	0.10	9.79
Indiana	1900	2010	111	1	1	1.33	0.01	83.46
Iowa	1900	2010	111	4	6	6.75	0.06	16.44
Michigan	1896	2010*	105	18	19	27.24	0.26	3.85
Michigan State	1953	2010	58	4	2	5.08	0.09	11.40
Minnesota	1896	2010	115	10	5	12.33	0.11	9.33
Northwestern	1896	2010*	113	2	5	4.16	0.04	27.16
Ohio State	1913	2010	98	18	16	26.16	0.27	3.75
Penn State	1993	2010	18	1	2	2.00	0.11	9.00
Purdue	1896	2010	115	1	6	3.49	0.03	32.95
Wisconsin	1896	2010	115	7	4	8.66	0.08	13.28

Note 1: The Championship Ties column includes two-way championship ties, which causes the Total Championships column to produce fractional values.
Note 2: * Michigan did not play in the Big Ten from 1907 to 1916 and Northwestern did not field a football team during the 1906 and 1907 seasons.

Table 4.12: Big East Conference Championship History

Team	First Year	Most Recent Year	Total Years	Championships	Championship Ties	Total Championships	% Championships/Year	YPC
Boston College	1991	2004	14	0	1	0.25	0.02	56.00
Cincinnati	2005	2010	6	2	0	2.00	0.33	3.00
Connecticut	2004	2010	7	0	1	0.83	0.12	8.43
Louisville	2005	2010	6	1	0	1.00	0.17	6.00
Miami (FL)	1991	2003	13	5	3	6.333	0.49	2.05
Pittsburgh	1991	2010	20	0	1	0.58	0.03	34.48
Rutgers	1991	2010	20	0	0	0.00	0.00	***
South Florida	2005	2010	6	0	0	0.00	0.00	***
Syracuse	1991	2010	20	2	3	3.083	0.15	6.49
Temple	1991	2004	14	0	0	0.00	0.00	***
Virginia Tech	1992	2003	12	2	1	2.333	0.19	5.14
West Virginia	1991	2010	20	2	3	3.58	0.18	5.59

Note: The Championship Ties column includes two-way championship ties, which causes the Total Championships column to produce fractional values.

Table 4.13: Pac-10 Conference Championship History

Team	First Year	Most Recent Year	Total Years	Championships	Championship Ties	Total Championships	% Championships/Year	YPC
Arizona	1978	2010	33	0	1	0.33	0.01	***
Arizona State	1978	2010	33	2	1	2.50	0.08	16.50
California	1916	2010	95	9	4	10.83	0.11	10.56
Idaho	1922	1958	37	0	0	0.00	0.00	***
Montana	1924	1949	26	0	0	0.00	0.00	***
Oregon	1916	2010*	90	5	4	6.83	0.08	18.00
Oregon State	1916	2010*	90	2	3	3.33	0.04	45.00
Southern California	1922	2010	89	28	10	32.66	0.37	3.18
Stanford	1918	2010	93	8	4	9.83	0.11	11.63
UCLA	1928	2010	83	11	6	13.49	0.16	7.55
Washington	1916	2010	95	10	5	12.16	0.13	9.50
Washington State	1917	2010*	90	2	2	3.00	0.03	45.00

Note 1: The Championship Ties column includes two-way championship ties, which causes the Total Championships column to produce fractional values.
Note 2: * Oregon, Oregon State and Washington State were not members of the conference over the entire examination period.

Table 4.14: Southeastern Conference Championship History

Team	First Year	Most Recent Year	Total Years	Championships	Championship Ties	Total Championships	% Championships/Year	YPC
Alabama	1933	2010	78	16	6	18.83	0.24	4.88
Arkansas	1992	2010	19	0	0	0.00	0.00	***
Auburn	1933	2010	78	5	2	5.83	0.07	15.60
Florida	1933	2010	78	8	0	8.00	0.10	9.75
Georgia	1933	2010	78	8	4	10.00	0.13	9.75
Georgia Tech	1933	1963	31	3	2	3.83	0.12	10.33
Kentucky	1933	2010	78	1	2	2.00	0.03	78.00
Louisiana State	1933	2010	78	9	2	10.00	0.13	8.67
Mississippi	1933	2010	78	6	0	6.00	0.08	13.00
Mississippi State	1933	2010	78	1	0	1.00	0.01	78.00
Sewanee	1933	1939	7	0	0	0.00	0.00	***
South Carolina	1992	2010	19	0	0	0.00	0.00	***
Tennessee	1933	2010	78	9	4	10.66	0.14	8.67
Tulane	1933	1965	33	1	2	1.83	0.06	33.00
Vanderbilt	1933	2010	78	0	0	0.00	0.00	***

Note: The Championship Ties column includes two-way championship ties, which causes the Total Championships column to produce fractional values.

Table 4.15: Conference ADF and PP Unit-Root Tests

Conference	*Metric*	T (seasons)	*ADF (p)* *Constant*	*ADF (p)* *Trend*	*PP (l)* *Constant*	*PP (l)* *Trend*
ACC	RSD	58	-3.048 (2)**	-3.091 (2)	-5.677 (3)***	-5.624 (3)***
Big Ten		115	-6.130 (1)***	-6.430 (1)***	-7.208 (4)***	-7.394 (4)***
Big 12		83	-4.767 (1)***	-5.151 (1)***	-7.916 (3)***	-8.539 (3)***
Big East		20	-	-	-	-
Pac-10		95	-4.578 (2)***	-4.684 (2)***	-9.653 (3)***	-9.706 (3)***
SEC		78	-6.339 (1)***	-6.354 (1)***	-9.473 (3)***	-9.585 (3)***
ACC	MVR	58	-2.778 (1)*	-4.133 (1)***	-4.708 (3)***	-6.458 (3)***
Big Ten		115	-2.360 (1)	-7.393 (1)***	-3.995 (4)***	-11.108 (4)***
Big 12		83	-3.252 (1)**	-6.252 (1)***	-4.618 (3)***	-7.424 (3)***
Big East		20	-	-	-	-
Pac-10		95	-2.118 (1)	-5.204 (1)***	-2.778 (3)*	-7.937 (3)***
SEC		78	-2.076 (1)	-5.809 (1)***	-2.411 (3)	-7.452 (3)***
ACC	WPC	57	-5.961 (1)***	-6.329 (1)***	-7.753 (3)***	-8.015 (3)***
Big Ten		114	-5.824 (1)***	-6.230 (1)***	-8.476 (4)***	-8.750 (4)***
Big 12		82	-4.520 (2)***	-4.251 (2)***	-6.839 (3)***	-6.965 (3)***
Big East		19	-	-	-	-
Pac-10		94	-4.141 (2)***	-4.125 (2)***	-10.138 (3)***	-10.085 (3)***
SEC		77	-4.541 (1)***	-4.510 (1)***	-9.320 (3)***	-9.276 (3)***

p: the number of lags

l: lag truncation.

***, **, * = significant at 99%, 95%, and 90% critical levels, respectively.

Table 4.16: Conference RSD Two-Break LM Unit-Root Tests

Conference	$\hat{k}$	$\hat{T}_b$	$\hat{t}_{\gamma j}$	*Test Statistic*	*Critical Value Break Points*
ACC	0	1971, 2001	2.833***, -5.230***	-7.052**	λ = (0.21, 0.84)
Big Ten	0	1962, 1976	2.925***, -3.592***	-8.370**	λ = (0.58, 0.70)
Big 12	0	1961, 1984	-0.403, -4.738***	-10.066**	λ = (0.41, 0.69)
Big East	-	-	-	-	-
Pac-10	0	1926, 1985	7.259***, -3.572***	-10.558**	λ = (0.12, 0.74)
SEC	0	1981, 1998	-6.089***, 1.242	-10.116**	λ = (0.63, 0.85)

$\hat{k}$ is the optimal number of lagged first-difference terms included in the unit root test to correct for serial correlation.
$\hat{T}_b$ denotes the estimated break points. $\hat{t}_{\gamma j}$ is the value of DT_{jt} for j = 1,2. See Lee and Strazicich (2003) Table 2 for critical values.
***, ** = significant at 99% and 95% critical levels, respectively.

Table 4.17: Conference MVR Two-Break LM Unit-Root Tests

Conference	$\hat{k}$	$\hat{T}_b$	$\hat{t}_{\gamma j}$	*Test Statistic*	*Critical Value Break Points*
ACC	0	1971, 1990	1.633*, 3.249***	-8.104***	λ = (0.33, 0.66)
Big Ten	0	1940, 1980	-8.277***, 2.704***	-11.577**	λ = (0.39, 0.74)
Big 12	0	1940, 1952	0.359, 3.899***	-8.073***	λ = (0.16, 0.30)
Big East	-	-	-	-	-
Pac-10	0	1961, 1966	-4.423***, 1.288	-7.713***	λ = (0.48, 0.54)
SEC	0	1952, 1964	3.180***, -5.224***	-8.212***	λ = (0.26, 0.41)

$\hat{k}$ is the optimal number of lagged first-difference terms included in the unit root test to correct for serial correlation.
$\hat{T}_b$ denotes the estimated break points. $\hat{t}_{\gamma j}$ is the value of DT_{jt} for j = 1,2. See Lee and Strazicich (2003) Table 2 for critical values.
***, ** = significant at 99% and 95% critical levels, respectively.

Table 4.18: Conference WPC Two-Break LM Unit-Root Tests

Conference	$\hat{k}$	$\hat{T}_b$	$\hat{t}_{\gamma j}$	*Test Statistic*	*Critical Value Break Points*
ACC	0	1981, 1985	2.974***, 1.687**	-8.263***	λ = (0.50, 0.57)
Big Ten	0	1953, 1979	2.220**, -3.867***	-9.852***	λ = (0.50, 0.73)
Big 12	0	1939, 1985	-2.375**, 2.771***	-8.633***	λ = (0.13, 0.67)
Big East	-	-	-	-	-
Pac-10	2	1927, 1940	-0.718, -3.198***	-4.588	λ = (0.12, 0.25)
SEC	3	1948, 1973	5.319***, -0.438	-6.252**	λ = (0.19, 0.52)

$\hat{k}$ is the optimal number of lagged first-difference terms included in the unit root test to correct for serial correlation.
$\hat{T}_b$ denotes the estimated break points. $\hat{t}_{\gamma j}$ is the value of DT_{jt} for j = 1,2. See Lee and Strazicich (2003) Table 2 for critical values.
***, ** = significant at 99% and 95% critical levels, respectively.

Table 4.19: Conference RSD One-Break LM Unit-Root Tests

Conference	$\hat{k}$	$\hat{T}_b$	$\hat{t}_{\gamma j}$	*Test Statistic*	*Critical Value Break Points*
ACC	0	1996	-0.631	-6.564***	λ = 0.76
Big Ten	0	1963	-0.801	-8.290***	λ = 0.59
Big 12	0	1985	1.695**	-9.547***	λ = 0.70
Big East	-	-	-	-	-
Pac-10	0	1973	3.081***	-9.713***	λ = 0.61
SEC	0	1955	-4.638***	-9.981***	λ = 0.29

$\hat{k}$ is the optimal number of lagged first-difference terms included in the unit root test to correct for serial correlation.
$\hat{T}_b$ denotes the estimated break points. $\hat{t}_{\gamma j}$ is the value of DT_{jt} for j = 1,2. See Lee and Strazicich (2003) Table 2 for critical values.
***, ** = significant at 99% and 95% critical levels, respectively.

Table 4.20: Conference MVR One-Break LM Unit-Root Tests

Conference	$\hat{k}$	$\hat{T}_b$	$\hat{t}_{\gamma j}$	*Test Statistic*	*Critical Value Break Points*
ACC	0	1993	4.483***	-7.333***	λ = 0.71
Big Ten	0	1919	-7.120***	-11.064***	λ = 0.21
Big 12	0	1979	2.920***	-7.780***	λ = 0.63
Big East	-	-	-	-	-
Pac-10	0	1971	-4.523***	-4.933**	λ = 0.59
SEC	0	1967	-1.049	-7.997***	λ = 0.45

$\hat{k}$ is the optimal number of lagged first-difference terms included in the unit root test to correct for serial correlation.
$\hat{T}_b$ denotes the estimated break points. $\hat{t}_{\gamma j}$ is the value of DT_{jt} for j = 1,2. See Lee and Strazicich (2003) Table 2 for critical values.
***, ** = significant at 99% and 95% critical levels, respectively.

Table 4.21: Conference WPC One-Break LM Unit-Root Tests

Conference	$\hat{k}$	$\hat{T}_b$	$\hat{t}_{\gamma j}$	*Test Statistic*	*Critical Value Break Points*
ACC	0	1985	6.033***	-7.148***	λ = 0.56
Big Ten	0	1967	0.116	-9.503***	λ = 0.62
Big 12	8	1942	1.712***	-3.899	λ = 0.17
Big East	-	-	-	-	-
Pac-10	2	1928	-3.177***	-4.076	λ = 0.13
SEC	3	1959	4.992***	-5.765***	λ = 0.34

$\hat{k}$ is the optimal number of lagged first-difference terms included in the unit root test to correct for serial correlation.
$\hat{T}_b$ denotes the estimated break points. $\hat{t}_{\gamma j}$ is the value of DT_{jt} for j = 1,2. See Lee and Strazicich (2003) Table 2 for critical values.
***, ** = significant at 99% and 95% critical levels, respectively.

Table 4.22: Conference RSD Sequential Break Point Test Results

Conference	$SupF_t(1)$	$SupF_t(2)$	$SupF_t(3)$	$SupF_t(4)$	$SupF_t(5)$	*UDmax*	*WDmax*
ACC	12.493**	13.685***	11.213***	9.685**	7.740***	13.685**	17.311***
Big Ten	23.809***	14.269***	11.176***	9.114***	7.580***	23.809***	23.809***
Big 12	11.453*	10.671**	8.583**	7.511**	5.897**	11.453*	12.130*
Big East	-	-	-	-	-	-	-
Pac-10	8.058	8.015	7.035	6.086	5.261	8.058	10.315
SEC	4.980	5.673	5.626	4.565	3.816	5.673	7.719

Conference	*SupF(2/1)*	*SupF(3/2)*	*SupF(4/3)*	*SupF(5/4)*	*Breaks*
ACC	18.604***	7.643	0.562		2
Big Ten	4.004	3.896	1.358	1.417	1
Big 12	8.331	3.795	3.147		0
Big East	-	-	-	-	-
Pac-10	8.892	3.586	3.647	3.482	0
SEC	5.360	4.062	1.056	0.932	0

***Significant at the 99% critical level

**Significant at the 95% critical level

*Significant at the 90% critical level

Table 4.23: Conference MVR Sequential Break Point Test Results

Conference	*$SupF_t(1)$*	*$SupF_t(2)$*	*$SupF_t(3)$*	*$SupF_t(4)$*	*$SupF_t(5)$*	*UDmax*	*WDmax*
ACC	13.802**	16.574***	13.017***	10.154***	8.171***	16.574***	20.966***
Big Ten	8.786	6.134	4.862	4.175	2.769	8.786	8.786
Big 12	6.818	8.690**	8.135*	8.8200***	7.944***	8.8200	17.443***
Big East	-	-	-	-	-	-	-
Pac-10	14.302**	11.911**	10.668***	8.757**	6.851**	14.302**	14.636**
SEC	6.139	7.544	6.993	6.183	5.5271*	7.544	10.837

Conference	*SupF(2/1)*	*SupF(3/2)*	*SupF(4/3)*	*SupF(5/4)*	*Breaks*
ACC	14.140***	6.279	1.326	0.940	2
Big Ten	7.797	5.252	1.054		0
Big 12	10.254*	7.226	3.210		0
Big East	-	-	-	-	-
Pac-10	10.356*	7.789	1.905	0.058	1
SEC	5.965	7.253	5.409		0

***Significant at the 99% critical level

**Significant at the 95% critical level

*Significant at the 90% critical level

Table 4.24: Conference WPC Sequential Break Point Test Results

Conference	$SupF_t(1)$	$SupF_t(2)$	$SupF_t(3)$	$SupF_t(4)$	$SupF_t(5)$	*UDmax*	*WDmax*
ACC	5.634	4.853	4.183	3.276	2.577	5.634	5.739
Big Ten	16.596***	11.706**	11.478***	9.567***	7.922***	16.596***	17.393***
Big 12	14.217**	15.949***	14.958***	11.582***	9.092***	15.949***	22.386***
Big East	-	-	-	-	-	-	-
Pac-10	10.397*	10.567**	9.211**	8.116**	7.019**	10.567*	13.762**
SEC	6.401	5.716	5.095	4.262	3.560	6.401	7.058

Conference	*SupF(2/1)*	*SupF(3/2)*	*SupF(4/3)*	*SupF(5/4)*	*Breaks*
ACC	3.249	2.288	2.030		0
Big Ten	6.2141	7.837	2.535		1
Big 12	23.734***	11.156*	1.738	0.918	2
Big East	-	-	-	-	-
Pac-10	11.585	6.310	6.3223	2.248	0
SEC	3.507	3.174	1.693	1.148	0

***Significant at the 99% critical level

**Significant at the 95% critical level

*Significant at the 90% critical level

Table 4.25: Conference Break Test Results

Conference	T_1	T_2
RSD		
ACC	1966 [65, 81]	1993 [87, 94]
Big Ten	1964 [62, 77]	-
Big 12	-	-
Big East	-	-
Pac-10	-	-
SEC	-	-
MVR		
ACC	1970 [69, 72]	1988 [87, 91]
Big Ten	-	-
Big 12	-	-
Big East	-	-
Pac-10	1947 [44, 52]	-
SEC	-	-
WPC		
ACC	-	-
Big Ten	1963 [61, 75]	-
Big 12	1950 [49, 67]	1985 [83, 86]
Big East	-	-
Pac-10	-	-
SEC	-	-

Note: 90% confidence intervals are in [ ].

Table 4.26: Conference Breakpoint Regression Results

Conference	α_1	β_1	α_2	β_2	α_3	β_3	$\bar{R}^2\,(R^2)$
RSD							
ACC	-0.051	1.814	0.003	1.453	-0.038	3.434	0.282
	(-3.526)***	(14.815)***	(0.518)	(9.329)***	(-3.564)***	(6.366)***	(0.345)
Big Ten	-0.007	1.827	-0.002	1.811	-	-	0.195
	(-5.439)***	(34.562)***	(-1.011)	(8.034)***	-	-	(0.216)
MVR							
ACC	-0.013	0.591	-0.011	0.693	-0.006	0.624	0.620
	(-5.413)***	(22.020)***	(-4.510)***	(9.980)***	(-3.053)***	(7.093)***	(0.653)
Pac-10	-0.003	0.707	-0.003	0.604	-	-	0.761
	(-2.135)**	(25.845)***	(-6.306)***	(17.332)***	-	-	(0.769)
WPC							
Big Ten	-0.003	0.466	-0.005	1.024	-	-	0.153
	(-1.508)	(6.663)***	(-1.544)	(3.699)***	-	-	(0.175)
Big 12	0.031	0.165	-0.001	0.610	-0.021	2.215	0.282
	(4.256)***	(1.633)	(-0.008)	(4.028)***	(-3.597)***	(5.221)***	(0.327)

***Significant at the 99% critical level
**Significant at the 95% critical level
*Significant at the 90% critical level
α_M *and* β_M refer to the slope and intercept coefficients for regime M, respectively.
Note: Any series not listed here is stationary without breaks.

Table 4.27: RSD and the G.I. Bill: 1932-1941/1946-1955

Conference	Pre-G.I. Bill Average	Post-G.I. Bill Average
ACC	N/A	N/A
Big Ten	1.57	1.38
Big 12	1.43	1.55
Big East	N/A	N/A
Pac-10	1.56	1.69*
SEC	1.62	1.54**

Source: Data collected from Quirk and Quirk (2012); RSD values calculated by authors

a) G.I. Bill: Before (1932-1941); After (1946-1955)

b) 1942-1945 is not included because many programs eliminated football or lost players for a year or more because of WWII

c) ACC did not begin play in football until 1953

d) Big East Conference did not begin play in football until 1991

e) SEC began play in football in 1933

f) ***Significant at 0.01; ** Significant at 0.05; * Significant at 0.10

Table 4.28: MVR and the G.I. Bill: 1932-1941/1946-1955

Conference	Pre-G.I. Bill Average	Post-G.I. Bill Average
ACC	N/A	N/A
Big Ten	0.5948	0.4944***
Big 12	0.6415	0.5083***
Big East	N/A	N/A
Pac-10	0.6050	0.5310**
SEC	0.6033	0.5247***

Source: Game level data collected from http://www.sports-reference.com/cfb/

a) G.I. Bill: Before (1932-1941); After (1946-1955)

b) Note: 1942-1945 is not included because many programs eliminated football or lost players for a year or more because of WWII

c) ACC did not begin play in football until 1953

d) Big East Conference did not begin play in football until 1991

e) SEC began play in football in 1933

f) ***Significant at 0.01; ** Significant at 0.05; * Significant at 0.10

Table 4.29: WPC and the G.I. Bill: 1932-1941/1946-1955

Conference	Pre-G.I. Bill Average	Post-G.I. Bill Average
ACC	N/A	N/A
Big Ten	0.5032	0.3756
Big 12	0.4649	0.7367***
Big East	N/A	N/A
Pac-10	0.4704	0.3998
SEC	0.5344	0.3936

Source: Data collected from Quirk and Quirk (2012)

a) G.I. Bill: Before (1932-1941); After (1946-1955)

b) 1942-1945 is not included because many programs eliminated football or lost players for a year or more because of WWII

c) ACC did not begin play in football until 1953

d) Big East Conference did not begin play in football until 1991

e) SEC began play in football in 1933

f) ***Significant at 0.01; ** Significant at 0.05; * Significant at 0.10

Table 4.30: RSD and Athletic Grant-In-Aid: 1947-1956/1957-1966

Conference	Pre-Athletic Grant-In-Aid	Post-Athletic Grant-In-Aid	Alt. Post-Athletic Grant-In-Aid
ACC	N/A	N/A	N/A
Big Ten	1.38	1.41	1.51
Big 12	1.55	1.60	1.65*
Big East	N/A	N/A	N/A
Pac-10	1.69	1.43**	1.40**
SEC	1.52	1.64	1.63*

Source: Data collected from Quirk and Quirk (2012); RSD values calculated by authors

a) Athletic Grant-In-Aid: Before (1947-1956); After (1957-1966)

b) Alternate Post-Grant-In-Aid period: (1961-1970); allows for 4-year adaptation period

c) ACC did not begin play in football until 1953

d) Big East Conference did not begin play in football until 1991

e) SEC began play in football in 1933

f) ***Significant at 0.01; ** Significant at 0.05; * Significant at 0.10

Table 4.31: MVR and Athletic Grant-In-Aid: 1947-1956/1957-1966

Conference	Pre-Athletic Grant-In-Aid	Post-Athletic Grant-In-Aid	Alt. Post-Athletic Grant-In-Aid
ACC	N/A	N/A	N/A
Big Ten	0.4819	0.4714	0.4609
Big 12	0.5019	0.5341	0.4848
Big East	N/A	N/A	N/A
Pac-10	0.5068	0.4692	0.4436
SEC	0.5254	0.5356	0.4990

Source: Game level data collected from http://www.sports-reference.com/cfb/

a) Athletic Grant-In-Aid: Before (1947-1956); After (1957-1966)

b) Alternate Post-Grant-In-Aid period (1961-1970); allows for 4-year adaptation period

c) ACC did not begin play in football until 1953

d) Big East Conference did not begin play in football until 1991

e) SEC began play in football in 1933

f) ***Significant at 0.01; ** Significant at 0.05; * Significant at 0.10

Table 4.32: WPC and Athletic Grant-In-Aid: 1947-1956/1957-1966

Conference	Pre-Athletic Grant-In-Aid	Post-Athletic Grant-In-Aid	Alt. Post-Athletic Grant-In-Aid
ACC	N/A	N/A	N/A
Big Ten	0.3814	0.3665	0.5448
Big 12	0.7460	0.6365	0.5450*
Big East	N/A	N/A	N/A
Pac-10	0.4432	0.5046	0.6387*
SEC	0.3747	0.5596*	0.6878**

Source: Data collected from Quirk and Quirk (2012)

a) Athletic Grant-In-Aid: Before (1947-1956); After (1957-1966)

b) Alternate Post-Grant-In-Aid period: (1961-1970); allows for 4-year adaptation period

c) ACC did not begin play in football until 1953

d) Big East Conference did not begin play in football until 1991

e) SEC began play in football in 1933

f) ***Significant at 0.01; ** Significant at 0.05; * Significant at 0.10

Table 4.33: RSD and the Death of the CFA: 1986-1995/1996-2005

Conference	Pre-CFA Death	Post-CFA Death
ACC	1.57	1.56
Big Ten	1.55	1.56
Big 12	1.71	1.62
Big East	1.79	1.61
Pac-10	1.37	1.56***
SEC	1.53	1.61

Source: Data collected from Quirk and Quirk (2012); RSD values calculated by authors

a) Death of CFA: Before (1986-1995); After (1996-2005)

b) Big East Conference did not begin play in football until 1991, so 5-year pre- and post- periods are used for Big East analysis

c) ***Significant at 0.01; ** Significant at 0.05; * Significant at 0.10

Table 4.34: MVR and the Death of the CFA: 1986-1995/1996-2005

Conference	Pre-CFA Death	Post-CFA Death
ACC	0.3581	0.3613
Big Ten	0.3687	0.3463
Big 12	0.4512	0.3923
Big East	0.4182	0.4266
Pac-10	0.3450	0.2996*
SEC	0.3684	0.3642

Source: Game level data collected from http://www.sports-reference.com/cfb/

a) Death of CFA: Before (1986-1995); After (1996-2005)

b) Big East Conference did not begin play in football until 1991, so 5-year pre- and post- periods are used for Big East analysis

c) ***Significant at 0.01; ** Significant at 0.05; * Significant at 0.10

Table 4.35: WPC and the Death of the CFA: 1986-1995/1996-2005

Conference	Pre-G.I. Bill Average	Post-G.I. Bill Average
ACC	N/A	N/A
Big Ten	0.5032	0.3756
Big 12	0.4649	0.7367***
Big East	N/A	N/A
Pac-10	0.4704	0.3998
SEC	0.5344	0.3936

Source: Data collected from Quirk and Quirk (2012)

a) G.I. Bill: Before (1932-1941); After (1946-1955)

b) 1942-1945 is not included because many programs eliminated football or lost players for a year or more because of WWII

c) ACC did not begin play in football until 1953

d) Big East Conference did not begin play in football until 1991

e) SEC began play in football in 1933

f) ***Significant at 0.01; ** Significant at 0.05; * Significant at 0.10

CHAPTER 5

Conclusion

Conclusions specific to each individual chapter are contained within those respective chapters. This final conclusion section will briefly describe the plan for extraction of individual papers from the larger dissertation chapters.

Specific to chapter one, a paper examining demand for live NFL attendance utilizing secondary market PSL and STR sale prices is currently under review. Because these data are unique additional opportunities exist to advance the research on demand for NFL football. One possibility is to examine the relationship between the economic and corporate structure of NFL cities and how these factors are related to the level of demand in each city. Specifically, NFL cities vary based not only on simple demographics, but also in the degree to which they serve as a business centers. For example, Chicago varies from Jacksonville in the number of firms located in each city, as well as the size, scope and specialty of those firms. Variation in these economic activity variables could play a part in explaining market differences in NFL demand.

The plan for the second chapter is to extract two separate papers from the work focused on the MLB Draft. The first paper will be a descriptive piece which highlights selection into the labor market and labor market outcomes by player type. This is beneficial because there is no existing academic work which provides a long-run examination of historical outcomes from the MLB Draft. This paper is suitable for a general sport management journal. The second paper from this chapter will focus on historical labor market outcomes through the use of the econometric methods specified. Again, there is no existing literature in the context of professional baseball which examines the relationship between training and employment outcomes.

It is expected that three papers will be extracted from the chapter on competitive balance in college football. The first will be a descriptive piece outlining the historical behavior of competitive balance in college football. The second will focus on the time series techniques and how they inform researchers on the behavior of balance while also identifying structural breaks in the data which must be identified before level analysis can be performed on a series. The final paper will focus on the invariance proposition in college football. It will highlight the time series techniques utilized and evaluate how balance has behaved in response to the three key institutional changes identified. The overall lack of attention paid to competitive balance in college sports allows for three separate papers to be extracted from this chapter.