2003

Possible contraction and its effect on 2002 major league baseball attendance

Peter W. Newberry
Colby College

Follow this and additional works at: http://digitalcommons.colby.edu/honorstheses

Part of the Economics Commons

Colby College theses are protected by copyright. They may be viewed or downloaded from this site for the purposes of research and scholarship. Reproduction or distribution for commercial purposes is prohibited without written permission of the author.

Recommended Citation
http://digitalcommons.colby.edu/honorstheses/180

This Honors Thesis (Open Access) is brought to you for free and open access by the Student Research at Digital Commons @ Colby. It has been accepted for inclusion in Honors Theses by an authorized administrator of Digital Commons @ Colby. For more information, please contact enhodes@colby.edu.
Possible Contraction and Its Effect on 2002 Major League Baseball Attendance

May 12, 2003

By Peter W. Newberry

Faculty Advisors: Professor Michael Donihue and Professor David Findlay
POSSIBLE CONTRACTION AND ITS EFFECT ON 2002 MAJOR LEAGUE BASEBALL ATTENDANCE

This paper creates three demand equations for baseball attendance to determine the factors that affect spring training and regular season attendance. It also examines what effect the threat of contraction had on the Montreal Expos’ and Minnesota Twins’ attendance. Cross sectional models using data from the 2002 spring training and regular seasons are estimated and analyzed. A third panel data set model is also formed and examined. Results from the spring training and regular season models indicate that the two teams threatened to be eliminated witnessed lower attendance compared to the rest of the league. The panel data model’s results reveal that although attendance was lower than the rest of the league in 2002, both teams witnessed an increase in spectators in comparison with the previous two years. While possible elimination and a successful season by Minnesota may have actually increased attendance relative to previous seasons, both teams still witnessed lower attendance than the rest of the league during 2002.
Introduction

After the 2001 World Series, Major League Baseball's (MLB) commissioner Bud Selig announced that two teams would be eliminated to remedy the financially struggling business of baseball. The two teams proposed for elimination were the Montreal Expos and the Minnesota Twins, due to their perennial financial difficulties and the small markets in which they compete. While the threat of contraction was removed on February 5th, 2002 (nine days before spring training began) fans of both Minnesota and Montreal were discouraged by the threat, causing a reduction in off-season ticket sales and driving many long-time fans away from baseball.

Simultaneously, the proposed contraction may have heightened the interest for some fans that wanted to show MLB that their team deserved to continue playing. A novelty effect, where fans are in attendance to see a team possibly for the last time, may have also caused an increase in fan interest. The question raised by this paper is: what effect did the threat of contraction have on game-day attendance? This question is answered using three different approaches. The first approach involves creating a cross sectional model for spring training game day attendance. This model is then analyzes impact of the threat of contraction on attendance using data from the 2002 spring training season. The second approach uses a similar model, but applies it to 2002 regular season data. The third approach uses panel data from three seasons to compare Minnesota and Montreal's attendance in pre-contraction-threat seasons (2000-2001) to their attendance in the 2002 season. The models developed for this paper are similar to those found in many journal articles dealing with the issue of game day attendance, but are primarily based on Bruggink and Eaton's (1996) model of 1993 Major League Baseball regular season game day attendance and Welki and Zlatoper's (1999) model of 1986-87 National Football League attendance. Not only

---

1 See Appendix I for contraction timeline.
will the models attempt to answer the question raised above, but they will also verify the other
factors that are important in determining spring training and regular season game-day attendance.

The data set for the spring training approach of this examination includes home games of all 20
teams that held their camp in Florida in 2002\(^3\). The teams that held spring training in Arizona are
excluded to create easier comparisons and lessen location bias. The data set for the regular season
approach includes the home games of all 30 teams competing in the 2002 regular season, while the
panel data set includes home games for the Minnesota Twins and the Montreal Expos from the
2000 to the 2002 regular seasons.

Using Ordinary Least Squares (OLS), a number of equations are estimated for each approach
and the results are analyzed to determine the factors that affect attendance and the magnitude of
their effect. The organization of this paper is as follows: an analysis of the literature reviewed for
this paper, an explanation of the model, discussion of the data, and an analysis of the results for
each of the three approaches, and finally, conclusions.

**Literature Review**

Analytical techniques have been used to model demand for professional sports attendance in
numerous studies, and baseball has been a focus of many of these papers. While cross sectional
and panel data models for regular season individual game attendance have been reported in
literature, a model for spring training game day attendance was not found. This is probably due to
the fact that the variables that affect demand for spring training baseball are considerably different
from that of the regular season, and that some of these variables are not available to the researcher.
While the studies reviewed below do not specifically model spring training attendance, they lend
insight into which variables are important and provide a general model upon which a demand for

\(^3\) Arizona and Kansas City moved to Arizona after the 2002 spring training season.
spring training baseball attendance can be built. These studies also serve as the basis for the regular season models presented later in this paper.

In their study, Bruggink and Eaton used data from the 1993 season to create a model that would describe game day attendance for professional baseball games. Their study was designed to determine what the business of Major League Baseball could do after the 1994 strike-shortened season in order to increase attendance and therefore, increase revenues. Explanatory variables included in their model were placed into four different categories: location, expected quality, time and weather factors, and special factors. The data were split up into the National League and the American League, and four regressions were run for each league due to the problem of multicollinearity in the data. Although there were eight regressions, only the results from equation #1 in the American League will be analyzed for simplification reasons.

The variables that were found to have a statistically significant positive effect on attendance were: population of the host city, average ticket price, and the adjusted fan cost index. Of these variables, population was the only one in which the correct sign of the coefficient was observed. Location variables with a statistically significant negative effect on attendance include: local income per capita, whether there are two teams in the area, the age of the stadium, and the percentage of the local population that is black or Hispanic. With the exception of per capita income, the negative coefficients on these variables were concurrent with their expected sign. The authors of the study attributed the incorrect signs of average price, fan cost index, and income to the problem of multicollinearity. The other specifications run in this paper leave out location variables such as these to lessen the effect of multicollinearity.

---

1 Bruggink and Eaton, pp. 18.
2 A measure of the price of complimentary goods such as food and beverages.
Expected quality of the game variables found to be significant and increase fan attendance were: home team wins in the last ten games, visiting pitcher's career record, the number of league leaders that are competing in the game, and the number of All-Stars on the visiting team's roster. All these variables received their expected sign. Statistically significant variables that had a negative effect on attendance were: the total number of games back\(^6\) the two teams competing finished last season, the total number of division championships won in the last three years by the teams competing, the number of All-Stars on the home team’s roster, and current number of games back for the home and visiting teams. Of these variables, the number of division championships won and the number of All-Stars on the home team were the only two where the observed sign on the variable was different from the expected sign. A notable result was that both a successful starting pitcher for the visiting team and All-Stars on the visiting team drew more fans while home team All-Stars actually decreased attendance. The amount of runs scored in the last 10 games by the home team, the ethnicity of the starting pitchers, and the current records of the starting pitchers were found to be insignificant. This was a surprise to the authors because in previous studies the race of the starting pitcher was found to have a significant effect on attendance\(^7\).

Time and weather factors that had a significant positive coefficient include: whether the game is played on a weekend and whether the game is a day game. The expected sign on the day game variable was negative. The authors attribute this to the fact that day games were rare in the American league so observations were limited. Variables with a significant negative coefficient were: if the game is played in April or May, if the game is played in temperatures under 55 degrees.

\(^6\) "Games Back" is calculated by taking the number wins of a given team and subtracting them from the wins of the team in first place. An uneven number of games played by the two teams will affect the given team's "games back" number, causing the standing to change by halves.

\(^7\) See Noll (1974).
Fahrenheit, and if there is a rain delay. Each of the observed coefficients on these variables was consistent with its negative predicted sign.

The final category of variables was special factors. Special factors with a significant positive effect on attendance were: if the game is shown on premium cable television, if the game is shown on local "free" television, if the game is broadcast on ESPN or CBS, if there is a major or minor promotion, if there is a double-header, if the competing teams are in the same division, if the game occurs in Cleveland's last home stand, and if Nolan Ryan is the starting pitcher. There was an expected negative sign on the following variable: if the game is televised on local or premium cable television.

Overall, this study of baseball game day attendance was marked with many coefficients with the opposite sign as expected and variables that were found insignificant. This can be attributed mostly to multicollinearity because many of the location variables were correlated with each other.

Welki and Zlatoper (1999) followed a similar format to that of Bruggink and Eaton, but they focused on professional football. Their analysis modeled individual game day attendance for the National Football League (NFL) with a data set that included 392 regular season games during the 1986-87 NFL seasons. Using the Tobit estimation technique, Welki and Zlatoper attempted to explain variations in attendance with variables that fell into the following categories: economic variables, demographic variables, quality of game variables, and variables reflecting influences not captured by the other categories. An interesting approach taken by Welki and Zlatoper was to include variables that demonstrate the expected competitiveness of the game, or how close the score will be.

---

8 1993 was the final season played at Cleveland's old stadium.
9 1993 was Nolan Ryan's last year playing before he retired.
10 Attendance is measured as proportion of tickets sold, so the Tobit method of estimation must be used.
Variables found to be significant and positively correlated with attendance were: home team’s record, the point spread reported one week before the game, an interaction term between whether it is raining and the temperature, whether the competing teams are division rivals, and whether the game occurs on a day other than Sunday. The coefficients on the point spread and the division rival variables were reported to have an undetermined expected sign, while home team’s record had its expected positive coefficient. It is interesting to note that as the home team moved away from being a heavy favorite, more fans were in attendance, so fans preferred a closer game to one where they think the home team will have a lopsided victory. Variables found to have a significant negative effect on attendance included: average price of a ticket, week of the season in which the game occurs, whether it is raining, and whether the game is blacked out by local television. The coefficient on week of the season was an interesting result because some think that as the season goes on, more people become interested because playoff positions and division championships are at stake. This result demonstrates that as the season progressed, the number of people who lost interest because their team was out of the playoff race outweighed the number of people who gained interest. The coefficient observed on week of the season may also indicate that this variable is correlated with the temperature variable, as it gets colder further along in the season and fewer fans will attend games. The coefficient on blackout was also a notable result, as it revealed that a blacked out game had less attendance than one televised. This implied that while the goal of blacking out a game is to increase attendance, the opposite actually occurred.

While this model was successful in finding some variables that significantly affected attendance with the correct sign, the overall use and effectiveness of this model was minimal. The adjusted r-squared was .481, meaning that these variables only explained 48% of the variation in attendance, where most of the other studies of game day attendance witness r-squared values
between 0.60 and 0.80. Also, a small number the explanatory variables were found to have a significant relationship with attendance.

Like the two previous studies, Medoff (1986) intended to model game day attendance, but his primary goal was to examine the effect of discrimination on attendance. He modeled game day baseball attendance for the 1980 season using the following variables: average price of a ticket, the number of other professional sports teams that play in the area, local per capita income, local population, if the stadium was built after 1970, if the team is in the NL, the team's division standing in 1980, and the percentage of players on the team that are black. He then replaced the final variable with the percentage of pitchers on the team that are black and, subsequently, the percentage of the local population that is black. Medoff's model used data from all the major league teams that played in the U.S. in the 1980 season.

In this study, the variables found to be significant and reduce attendance in every model were: how many professional sports teams play in the area, the team's division standing in the 1980 season, and the per capita income of the local city. The first two were consistent with the theory, as other teams playing in the area are seen as a substitute and as a team has less success, attendance will decrease. It was interesting to see that as income became higher, fewer people attended baseball games. This may have been due to the fact that cities with higher incomes usually have more entertainment substitutes on which to spend money. The only variable that positively affected attendance was population. It was interesting to note that average price of a ticket, the age of the stadium and the percentage of players who are black did not have any significant effect on attendance.

Although this model had a relatively high adjusted r-squared (.66) the results seen here do not necessarily aid in creating the three models examining 2002 attendance. This is due to the fact that
Medoff uses very few explanatory variables, while Bruggink and Eaton showed that there are many significant variables that affect attendance. Another problem is that this study was done on the 1980 season. Expansion and many other advances since 1980 have changed the dynamics of the league dramatically, so a model of the 1980 season may not be an accurate one for more recent seasons.

Knowles, Sherony, and Haupert (1992) also sought to model game day MLB attendance over a complete season, although their principle concern was how the home team's chances to win affected attendance. The explanatory variables used to explain variations in attendance in the 1988 MLB season were: the sum of the games back for the two teams competing, if the game is a weekend game, if the game is an evening game, population of the home city, unemployment rate of the home city, income per capita of the home city, distance to the visiting city, and probability of the home team winning.

Variables in this model found to have a significant positive effect were: if the game is a weekend game, if the game is a night game, population, and probability of the home team winning. Variables with a significant negative effect on attendance included: combined games back, and distance to the visiting team's city. The r-squared value for this model was found to be .36, which might be explained by the small number of explanatory variables included. The overall usefulness of this study is minimal, but the explanatory variable dealing with the distance traveled for the visiting team is something to consider for the models presented below.

Butler (2002) modeled game day attendance for the 1999 season in order to determine the effect of interleague play on attendance. The explanatory variables he used to explain game day attendance fell into four categories: quality of the game variables, time and weather variables, special factor variables, and interleague play variables. An interesting approach taken by Butler is

\footnote{For calculation of this probability see Knowles, Sherony and Haupert (1992), pp. 75-76.}
to not include any economic or demographic variables such as population, income or price of a ticket. Instead, he estimated an intercept term for each of the 30 teams\textsuperscript{12}. Although this is something to consider, as multicollinearity in these variables has been a problem in many of the papers analyzed, the importance of these variables, as well as the desire to see their separate effects leads them to be included in the models of 2002 attendance (both spring training and regular season). While three regressions were run, each with a different interleague play variable, only the first regression will be analyzed for simplification reasons.

Variables found to have a significant positive effect were: the visiting starting pitcher's career record; whether the game is on a Thursday, Friday, Saturday, or Sunday (relative to Monday); whether the game occurs in May, June, July, August, September, or October (relative to April); whether the game is a divisional game; whether the game is the home opener; the log of visitor payroll; whether the home team is on a winning streak; and whether the game is an interleague game. Butler determined that interleague games had 6% more attendance than a regular intraleague game. The only variable found to significantly reduce attendance was home team's current games back. The high r-squared value (.77) indicates that this model explains a large percentage of the variation in attendance, and the fact that it is from a recent year (1999) makes it more applicable to a study of 2002 attendance.

Finally, Noll (1974) used data from the 1970 and 1971 baseball seasons to create a model for game day attendance. His primary goal was to determine if an increase in Black population affects attendance positively or negatively. Variables found to negatively affect attendance were: per capita income, other professional sports playing in the area, and the local Black population. It was interesting to contrast this study to the previous one because Knowles, Sherony, and Haupert found the percentage of black population to be insignificant. This may be due to the fact that Noll's

\textsuperscript{12} Butler (2002), p. 322.
study is from 1970, when discrimination in sports was still much stronger than it was in the '80s. Variables found to positively affect attendance were: population, stadium age, number of star players and recent pennant wins. One interesting note is that as stadium age increased, attendance increased. This is contrary to many modern thoughts about how new, state-of-the-art stadiums affect attendance.

Although none of these papers dealt with the specific topic of spring training baseball attendance or had the same structure of the 3 year panel data model presented below, they aided in developing a general model for attendance upon which all three models (spring training, 2002 regular season, 3 year panel) can be based. Together, these articles gave a good idea of what factors generally affect attendance for sports games, be it spring training or regular season. The following section discusses the theory of attendance for spring training baseball and subsequent sections include a discussion of the data and analysis of the results for the spring training model.

**Model of Spring Training Attendance**

The objective of this study is to examine the impact of possible contraction on 2002 game day attendance for the Minnesota Twins and Montreal Expos. The first approach taken for this analysis is to examine the impact of the threat of contraction on spring training attendance.

A number of studies have examined the determinants of attendance in Major League Baseball, along with other sports and the models for each study are similar in nature. The common variables included in these models fall into several categories: location/stadium issues, time and weather factors, expected quality of the game issues, and special factors. In a model of Major League Baseball spring training attendance, most of the variables fall into these categories. Location issues are variables describing the demographics and economics of the city where the games are

---

13 See Hanssen and Andersen (1999).
played, competing entertainment options for the fan base, as well as the make-up of the stadium. Weather and time factors include any weather or time condition that may affect fan attendance to an individual game, positively or negatively. It is important to note that all spring training stadiums are outdoor stadiums and therefore are subjected to weather conditions. Time variables explain how the specific time/day/week that the game takes place affects the number of fans in attendance. Expected quality-of-the-game variables measure fans’ pre-game notions of whether the game will be a “good game”. Examples of this include past performance of the teams competing, quality of starting pitchers, and quality of the players competing in the game. These variables are thought to be not as important for spring training baseball, relative to regular season, because fans are in attendance to just see a baseball game and do not care as much whether it is a good game or not. Special factors include all the variables that are not included in the above categories, such as proximity of the visiting team’s city and a contraction variable. The following is an introduction to the explanatory variables and their predicted relationship with the dependent variable, attendance.

Location and stadium variables thought to positively affect attendance are: income, population, and the number of service workers (workers in hotel, restaurant, entertainment) in the home team’s spring training camp location, and capacity of the home team’s stadium. Although income and population of the local city may not be an important determinant of spring training attendance, they will be included in this model because they have been found to be significant in previous attendance studies and because this model can serve as an analysis of the hypothesis that local demographics aren’t significant. The service worker variable is a proxy for tourist traffic in the home team’s city, as more tourists mean more employees in these industries. This variable is unique to a model of spring training attendance because, unlike the regular season, most of the fan
base of spring training games is made up of tourists. Location and stadium variables thought to negatively affect attendance are: age of the stadium, average price of a ticket, and a dummy variable equaling 1 if there is another team playing in the area of the home team and 0 otherwise.

Quality of the game variables thought to have a positive relationship with attendance are: the number of division championships won in the last three years for the home and visiting teams, the number of All-Star players on the home and visiting teams’ rosters, and the total number of players competing in the game that finished last season in the top three in their league for home runs, runs batted in, or batting average. Quality of game variables with a predicted negative relationship with attendance are: the number of games back the home and visiting team finished from the division champion last season and dummy variables equaling 1 if the home or visiting team have only half of their roster (split squad). The division championships and games back variables serve to analyze how the quality of the teams competing in the game affect attendance while the All-Star and league leader variables are included to analyze the draw of star players.

Weather and time factors with a predicted positive relationship with attendance are a night game dummy variable and a weekend game dummy variable. Weather and time factors thought to have a negative relationship with attendance are: a weather dummy equal to 1 if it rained any time throughout the day and 0 otherwise, a dummy variable equal to 1 if the game is played in cold weather and 0 otherwise, and a series of four ‘week’ dummies that analyze what happens to spring training attendance as it gets closer to regular season opening day. Included in this model is a hot dummy variable, but the relationship for this is undetermined. Although some people may not like to sit and watch a baseball game when it is over 90 degrees outside, as they could be doing cooler vacation activities, other people may actually prefer the hot temperature if they are visiting from northern destinations.
Special factors that have a predicted positive relationship with attendance are whether the two teams competing are regular season division rivals and whether there is a stadium promotion. Although these games have no bearing on the pennant race within the division, the rivalries still exist and are strong in spring training. Games such as the Red Sox versus the Yankees are likely to draw more fans than ones where there is no division rivalry. While promotions have been found to be an important determinant of attendance, the information on special stadium promotions was not available and therefore will not be included in this model. Distance the visiting team must travel is thought to be negatively correlated with attendance as the further away a fan must travel, assuming the fan is in the city where his/her team is located, to see his/her team play at another stadium, the less likely that fan is to make that journey.

As stated, the purpose of this study is to determine whether a team pegged for contraction will have higher or lower attendance than a team not pegged. This model intends to see whether a contract dummy variable has a positive or negative relationship with attendance. A positive relationship would indicate more fans came to see teams that were on the brink of elimination. This would be a novelty effect, as people would attend a game because it might be one of the last times they see that team play. A pride factor would be involved as well, as fans would want to show MLB that their team deserves to exist. A negative relationship would show that fans were frustrated and did not want to cheer for a team that will not be in MLB next year and that off season ticket sales were hurt by the contraction threat. The predicted relationship is negative. As the Minnesota Twins and the Montreal Expos were the two teams that the MLB targeted to contract, there will also be Twins and Expos dummies included in one regression to analyze the separate effects of contraction on these teams.
The general model for spring training attendance is:

\[ \text{ATTENDANCE} = f(JNC, \text{POP, SERVWORK, CAP, STADAGE, AVEPRICE, 2TEAM, GBHOME, GBVIS, DIVCHMPhOME, DIVCHMPVIS, HOMESTAR, VISTAR, LEADERS, HOMESS, VISSS, RAIN, COLD, HOT, NIGHT, WEEKEND, WEEK1, WEEK2, WEEK3, WEEK4, RIV, PROM, VISPROX, CONTRACT}) \]

A definition of each variable can be found in table 1.

**Data and Methodology**

The task of collecting data for spring training games is more difficult than for regular season games. This is due to the fact that records are not as available for spring training because the games are considered to be of less importance. A lot of information can be found in the box score, but other factors, such as weather, have to be collected from a separate source. The following is a description of the data collected (see appendix two for a summary of the data) for this model and the sources used to find the data.

Attendance data are collected from the box scores that appear on www.espn.com. These figures are ticket sales for the 2002 home spring training games of the twenty Major League Baseball teams that held their spring training in Florida. The 10 teams that had their camps in Arizona are excluded. The average attendance for these 20 teams was 5,106 fans per game. Minnesota's
**TABLE 1**

**Dependent Variable**

ATTENDANCE*  
The number of fans that attended a game as reported in the box score.

**Explanatory Variables**

**Location Variables**

INC*  
The per capita income of the home city as reported in the 2000 census.

POP*  
The population of the home city as reported in the 2000 census.

SERVWORK*  
The number of employees in the restaurant, hotel, and entertainment industries as reported in the 2000 census.

CAP*  
The capacity of the home stadium.

STADAGE  
The age of the home stadium.

AVEPRICE*  
The average ticket price to attend a game.

2TEAM  
A dummy variable equaling one if there are two teams playing in the same spring training site.

**Quality of the game variables**

GBHOME  
The number of games back the home team finished in 2001.

GBVIS  
The number of games back the visiting team finished in 2001.

TOTGB  
The sum of the previous two variables.

DIVCHAMPHM  
The number of division championships the home team won in the past three seasons.

DIVCHAMPVIS  
The number of division championships the visiting team won in the past three seasons.

TOTDIVCHAMP  
The sum of the previous two variables.

HMSTAR  
The number of players on the home team’s roster that competed in the 2001 All-Star game.

VISSTAR  
The number of players on the visiting teams roster that competed in the 2001 All-Star game.

TOTSTAR  
The sum of the previous two variables.

LEADER  
The number of players playing the game that finished in the top three in RBI, home runs, and or batting average in 2001.

HMSS  
A dummy equaling 1 if the home team has a split squad.

VISSS  
A dummy equaling 1 if the visiting team has a split squad.

**Time and Weather Variables**

RAIN  
A dummy equaling 1 if it rained any time from 9 a.m. to game time.

COLD  
A dummy equaling 1 if the game time temperature is under 65 degrees.

HOT  
A dummy equaling 1 if the game time temperature is over 85 degrees.

NIGHT  
A dummy equaling 1 if the game is played after 5 p.m. local time.

WEEKEND  
A dummy equaling 1 if the game is played on a Friday, Saturday, or Sunday.

WEEK1  
A dummy equaling 1 if the game was played from February 24 to March 4.

WEEK2  
A dummy equaling 1 if the game was played from March 5 to March 11.

WEEK3  
A dummy equaling 1 if the game was played from March 12 to March 18.

WEEK4  
A dummy equaling 1 if the game was played from March 19 to March 25.

**Special Factors**

VISPROX*  
The distance in miles from the visiting team’s stadium to the home team’s.

DIVRIV  
A dummy equaling 1 if the teams playing are in the same division.

CONTRACT  
A dummy variable equaling 1 if the home team is the Twins or the Expos.

TWINS  
A dummy variable equaling 1 if the home team is the Twins.

EXPOS  
A dummy variable equaling 1 if the home team is the Expos.

*indicates that the log of the variable is taken.
average was 6,092 fans per game, while the Expos' was 3,473. The Yankees had the highest average with 9,954 fans per game, while Texas had the lowest with 3,211.

Income and population are collected from the 2000 United States Census at www.census.gov. They are the income per capita and population of the home city of each of the teams. The league average for these two variables was $26,367 and 76,745 people. The income per capita and population of Fort Meyers, where the Twins play, was $28,514 and 48,208 people, while the Expos' spring training site, Jupiter, had an income per capita of $54,945 and a population of 39,328 people.

The service worker variable is the total number of people employed in the arts, entertainment, recreation, accommodation, and food service industries in each city. This is also collected from the 2000 United States census. The league average number of service workers was 4,381 employees, while Fort Meyers and Jupiter had 2,657 and 2,117 employees in these industries, respectively.

Stadium age, capacity, and average price of a ticket are all collected from www.springtrainingmagazine.com. The average age of a stadium was about 19 years while the average capacity is 7,458. Average price is calculated as a simple average, as the number of tickets at each price was not available. The average price for all of the teams was about $10.67. Atlanta had the most expensive tickets with an average price of $15.75, while Pittsburgh's $7.66 was the lowest.

Games back for the home and visiting teams are collected from www.espn.com. Both of these variables are the number of games back the team finished from first place in their division in 2001. These variables are included to explain the effect of last season's success on spring training attendance this season. The league average of games back the home team finished last year is

about 14.5, while the Twins finished 6 games back and the Expos finished 20 games back. The Twins had a relatively easy spring schedule playing teams with an average of 19 games back, while the Expos played teams with an average of 12.5 games back. The team with the toughest schedule was the Dodgers (7 Games Back), while the Yankees had the easiest (21 games back).

Division championships are collected by looking at who won their division in the 1999-2001 seasons. The Twins' opponents had an average of 0.68 division titles the last three years, while opponents of the Expos had 0.19 division titles. The league average for opponents was 0.64.

Home and visiting stars are collected by looking at the rosters of the 2001 All-Star game and noting how many All-Stars are on each of the major league teams. The average number of All-Stars on each team was about 1.9. Twins played teams with an average of 1.8 All-Stars on their roster, and the Expos' opponents had an average of 1.6 All Stars.

League leaders is collected by determining how many players competing in each game finished in the top three in RBI, home runs, and batting average in the 2001 season. The average number of leaders that played each game was about 0.7, while the Twins' games featured about 0.6 leaders per game. The average number of league leaders was the highest for Cleveland (3.1). This is due to the fact that 3 players on Cleveland were league leaders in 2001.

Home and visiting split squad variables are collected to determine if the games that were played with split squads had fewer fans in attendance. The Twins had four split squad home games, while the Expos had two. About 8% of the games for spring training were home split squad games, and 14% were visiting split squad games. A quarter of Twins opponents played with a split squad, and 19% of the Expos' games were competed with a visiting split squad.
The week of the season dummy variable is included to determine if more fans attend games the further along spring training is. There were six days in week 1 and week 5, while there were seven days in the weeks in between. There were two to four home games for each team in each week.

The rain, hot, and cold dummy variables are collected from www.weatherunderground.com. Only 3% of the games played in 2002 spring training had rain sometime throughout the day, while 20% of the games were played under hot or cold temperatures. Neither Montreal nor Minnesota played any games when it had rained, but 38% of the Twins games were played in extreme temperatures.

The night and weekend dummy variables are found by looking at the game day and time. Most spring training games occurred during the day as only 11% of the games played in 2002 were night games. The Twins and Expos each had two night games. 37% of the spring training games were weekend games.

The visitor proximity variable is calculated by taking the distance, in miles, between the home and the visiting team’s cities. The distances were calculated using Yahoo!maps on www.yahoo.com. The average distance the visiting team has to travel is 70 miles. Both the Expos’ and the Twins’ opponents had to travel more than the average. The maximum distance is 250 miles. This is the distance between Clearwater, where the Philadelphia Phillies play, and Jupiter.

The contract dummy variable is simply 1 if the home team is Minnesota or Montreal. 32 of the 302 games were home games for the Twins and the Expos; therefore 32 of the games take the value 1 for the contract variable.
Specifications

Although the above is the general model for spring training attendance, four different variations of this model are examined due to the problem of multicollinearity in the location/stadium variables and to determine an accurate specification. Like Bruggink and Eaton (1996), problems with multicollinearity in variables such as population, service workers, capacity, stadium age, average price, home division championships and home All-Stars are found by running auxiliary regressions. These problems are attempted to be solved by running multiple regressions, with and without some of the variables listed above. Heteroskedasticity is also present, so the equations are modified using White-Heteroskedasticity-Consistent Standard Errors.

Specification 1 includes all of the variables presented in the general model above. This specification is the most general one and is most likely to be a victim of multicollinearity. It is also the only equation where the division championship variables are included.

Specification 2 attempts to reduce multicollinearity by removing some of the location and home team specific variables. Service workers, average price, and other teams playing in the same city are now the only location variables included. These are thought to be the most important location variables in a model of spring training attendance. The separate home and visiting games back variables are now combined into a total variable to allow for analysis of the effect of overall quality of the teams on attendance. This is also done for the division championship and All-Star variables.

Specification 3 is thought to be the most accurate model for spring training attendance. Population is removed because theory suggests that it should not have a significant effect on spring training attendance. The two team variable is not present because it is often the case that when one

15 See Appendix 3 for Variance Inflation Factors.
The final specification is included to analyze the separate effects of contraction on the Twins and the Expos. A semi-log functional form is used for all specifications to realize constant price elasticity and to simplify interpretation on the coefficients, as attendance will change in percentage terms.

**Results and Analysis of the Spring Training Approach**

Results of the spring training model can be seen in appendix 3.

The income variable is used in specification 1, 3, and 4 and it is found to be significant in the latter two. The interpretation on the coefficient in specification 3 is as follows: an increase in 1% of per capita income will increase attendance by 0.31%. The positive sign on the coefficients in specifications 3 and 4 is consistent with the theory presented in the model section of this paper. The negative, insignificant coefficient in the first equation is likely caused by multicollinearity, as all the location and home team specific variables are included.

The population variable is included only in the first equation. It is found to have a significant positive effect on attendance, ceteris paribus. Although this is consistent with theory, the
significance is somewhat surprising because it is thought that most of the fan base for spring
training is tourists, and not the people who live in the home city. Population may be picking up
some of the effects of the service worker variable, as these two are closely related.

The service worker variable is used in all the models and is found to be in significant three of
them. The positive signs on the coefficients in specifications 2, 3, and 4 are consistent with theory,
while in specification 1, the negative insignificant coefficient is, again, likely caused by
 multicollinearity. The coefficients on the service worker variables are near 0.1, meaning that a 1%
increase in the number of employees in the service industry will lead to a 0.1% increase in fan
attendance. If it can be said that a 1% increase in tourists would cause a 1% increase in service
workers, then it can be concluded that this 1% increase in tourists will lead to a 0.1% increase in
attendance. This makes sense, as not all incoming tourists will necessarily attend spring training
baseball games.

The capacity variable is significant in the three specifications that it appears in, and the positive
sign on the coefficients are consistent with the theory presented earlier in this paper. The
regressions indicate that a 1% increase in capacity of the home team’s stadium will cause a 1.16%
increase in attendance. This variable is likely not significant for a regular season model, but the
small sizes of the spring training stadiums lead to games being sold out easier. With a sell-out,
there are fans that wanted to come to the game but couldn’t, meaning that if the stadium had more
seats, these fans would be able to watch the game and attendance would increase. Another
interesting note that may cause this result is that some of the spring training stadiums allow for
fans to sit in the grass in the outfield. These seats do not appear in the capacity number, but the
fans are counted in the attendance.
The stadium age variable appears in three equations and is included to analyze the much-argued hypothesis that newer stadiums draw more fans. The mixed results on this variable indicate that multicollinearity is likely, and that the age of the stadium is not a significant factor in determining attendance.

The average price variable is statistically significant in two models in which it appears, and the negative sign on the coefficient in these models is consistent with the theory of demand. The coefficient in equation 3 implies that the price elasticity of demand for spring training baseball is -0.26. The results in the other two equations lead to questions about the soundness of this result. The incorrect sign and insignificance in equations 2 and 4 may indicate that demand for spring training baseball games is price inelastic. This theory is justifiable because people paying for plane tickets and hotel rooms may not be very price sensitive.

The coefficients on the two team in the area variable in models 1 and 2 can be interpreted as follows: if the home team is playing in a city where another team has their spring training, they have about 50% more attendance than if they didn’t, ceteris paribus. Although these variables are found significant, the positive sign is not consistent with the theory that another team acts as a substitute. This result could be because two games are rarely played in the same city on the same day, and therefore the fans are not drawn away to another game.

Games back for the home and visiting teams are included in specification 1, 3, and 4 to examine the effects of last season’s success for the home and visiting teams. Games back home is found to be statistically significant in all three regressions, while games back visitor is only found significant in one. The negative signs on all the coefficients are the same as the predicted signs. The coefficient seen on the home games back variable in specification 4 indicates that if the home team finished the previous season one game further back, then attendance decreases by 0.5%. The
coefficient on visiting games back means that if the visiting team finished one game further back, then attendance decreases by .3%. The larger games back home coefficient is consistent with theory, as the majority of people come to games because they are fans of the home team; a more successful home team will attract more fans relative to a more successful visiting team.

Games back total is included in one specification to determine the effect of overall team quality on attendance. It is statistically significant for the model in which it is included and the value of the coefficient is around -.002 for this specification. Interpretation on this coefficient is: if one of the teams competing in the game finishes the last season one game further back, then attendance decreases by .2%.

Division championships for the home and visiting teams are similar to the games back variables as they are included to study the effect of past team performance on attendance. But while the games back variables examine the effects of the previous season, these two variables include stats from the last three seasons. These variables are only included in specification 1 due to the problem of multicollinearity discussed above and because of theory. They are both found statistically significant and the correct sign on the coefficients are observed. The interpretation of the result is: for every additional division championship won by the home team, attendance would increase by 14%. For every additional division championship won by the visiting team, attendance would increase by 9%.

Total division championships, similar to total games back, is a proxy for overall game quality. It is found statistically significant, with the correct sign on the coefficient in specification 2. Although games back was thought to be a better predictor of how attendance would be affected by team success, it appears that these results indicate otherwise. Fans are more interested in seeing a
team that has had consistent success, and more success (division championship) than a team that finished relatively well in the standings last year.

Home and visiting stars are included in specifications 1, 3 and 4 to determine the effects of All-Star players of each team on attendance. These variables are the drawing power of All-Star players on attendance. The results of regressions 3 and 4 are similar to the predicted results, as the drawing power of home and visiting All-Stars are significantly positive. The results show that if the home team has one more All-Star on their roster, then attendance would increase by 5%, while an increase of one visiting All-Star leads to a 7% increase. The fact that the coefficient on visiting stars has a larger magnitude implies that fans are drawn more by the visiting team's star players than the home team's. This makes sense, as home team fans are able to see their hometown stars multiple times during spring training, while they might only see the visiting team's stars once or twice.

The total stars variable is used in one equation and is found to be a significant determinant of attendance. These results indicate that the total number of All-Stars competing in the game affected attendance positively in 2002 spring training. This is understandable, as both home and visiting All Stars individually affected attendance.

The league leaders variable is not found to be significant in either of the two equations in which it appears. This suggests that the number of league leaders from last season competing in the spring training game had no effect on attendance. It is not uncommon for league leaders to be found on teams that are not successful. In 2002, both the Texas Rangers and the Colorado Rockies had three players in the top three in the batting categories, but their teams finished last in their respective divisions. These results imply that team success is a more important determinant of attendance than individual player success.
The home and visiting team split squad variables are included in all equations. The home team split squad variable is found to be a significant determinant of attendance in each of these specifications. However, the visiting split squad variable is not significant in any of the specifications. These results indicate that fans prefer to attend a game with their home team having a full roster, while a reduced visiting team’s roster has no effect. The value of the coefficient on home split squad variable in the majority of the equations is near -.18, which means that if the home team has a split squad roster, attendance is reduced by 18% for that game.

The rain and hot weather variables are included in two specifications while the cold variable is included in all of them. The cold dummy variable is the only variable that is found to be significant. On days when the temperature is below 65 degrees, about 20% fewer fans attend. The hot variable’s insignificance may be due to the fact that many people don’t care whether it is 80 degrees or 90 degrees and that the fans may specifically be in Florida to enjoy that hot weather. Very few games occurred in 2002 spring training when it had rained earlier in the day, which probably is the cause of finding the rain variable insignificant.

The night game variable is included in the first two specifications while the weekend game variable is included in all of them. The weekend variable is found to be significant in all specifications while the night variable is only significant in one. Most spring training games occurred during the day, causing very few occurrences of the night variable. This could be the reason it was found insignificant. Also, since most of the fans are tourists, the fact that the game is being played during the workday is not a factor. The coefficients on the weekend variable in equations 3 and 4 signify that if the game occurred on a weekend, attendance increased by 12%. This result can be explained by the fact that it is more likely that people will travel from far away for a weekend to see games because they do not have to work.
The week of the season dummy variables are included in all specifications. The coefficients on these variables appear as comparisons to games that occurred in week 5, as the 5th week variable is not included. All four of the week dummy variables are found to be significant; meaning attendance in the first four weeks of spring training was significantly different from that of the 5th week dummy variable. As spring training progresses, teams continue to make cuts and their roster begins to look like the one that will be in place on opening day. The results above state that in the early stages of spring training, when rosters are full of players that will not be on the opening day roster, fewer fans are in attendance compared to the final week of spring training.

The coefficients on 1st week and 2nd week variables are near .30, meaning that attendance in these weeks was about 30% less than attendance in the final week, ceteris paribus. The 3rd week had about 10% less attendance than the final week, while 4th week had about 15% less.

The visitor proximity variable is only statistically significant in two of the equations in which it is included. The coefficient's positive sign is different from the predicted sign. This indicates that the distance the fan must travel to see his/her team at another stadium is not a significant predictor of attendance. This could be explained by the fact that a lot of fans attend spring training to see a baseball game and not a specific team.

A division rival dummy variable is only included in the first two regressions, and it is found to be insignificant. The fact that the games being played for spring training have no bearing on a pennant race makes a division rivalry in spring training less important. Also, teams such as Texas and Los Angeles have no teams in their division that hold their spring training in Florida. These facts put together can explain why the division rival variable is found insignificant.

The purpose of this section of the paper is to determine whether teams that were threatened to be contracted witnessed significantly less spring training attendance than those teams not
threatened. The contract dummy variable is included and found significant in three specifications. The negative coefficient on the variable indicates that the Expos and the Twins had less attendance than the other teams. The coefficient on the contraction variable is found to be between -.23 and -.47 meaning that these two teams had 23% to 48% less attendance than the other teams, depending on which specification is examined. In the fourth specification, Twins and Expos dummy variables are included to examine the effects of contraction on each individual team. While Twins dummy variable is found insignificant, Expos dummy is found to significantly reduce attendance. The coefficient implies that the Expos had 70% less attendance than the other teams, ceteris paribus. This leads to a question of what is driving the combined contraction dummy variable result. While it is shown that the other major league baseball teams had significantly more attendance when a contract dummy variable is included, the fact that the Twins are not significantly different indicates that the Expos may be driving this result. Although, it is important to note that when a contract dummy variable is included, it is significant, showing that, alone, the Twins attendance data may not show effects, but together with the Expos, it does.

These results indicate that a team pegged for contraction had significantly less attendance in 2002 spring training, than the other teams. They also suggest that the Montreal Expos had significantly less attendance than the other teams, while the Minnesota Twins did not. A question that is raised is whether the threat of contraction is the reason for the lower attendance or if the Twins and the Expos (or just the Expos) witness lower attendance on an annual basis. The panel data approach will attempt to answer this. If the threat of contraction is the cause of the lower attendance, it is believed that fans were discouraged and no longer wanted to pay to attend a baseball game where the home team was one of the teams that was going to be eliminated. Although, by the time spring training began, the Twins and Expos were saved, it is thought that the
fans were still very upset that their teams were even considered to be eliminated and were fed up with Major League Baseball. If the opposite result occurred and attendance increased, it is thought that the commissioner may have chosen the wrong teams to be eliminated, as they have such a strong fan base which fought for their team’s survival. But these results indicate that the commissioner’s choice of the Montreal Expos was the correct one as their spring training attendance decreased with the threat. The regular season and panel data approaches will attempt to answer the question of whether the commissioner’s choice of the Twins was correct as well.

**Model of Regular Season Game Day Attendance**

Like the spring training approach, this model attempts to explain game day attendance for Major League Baseball in order to determine the effects of the threat of contraction on attendance. The model of regular season attendance is similar to spring training, although there are a few adjustments. This model is mostly based on Bruggink and Eaton’s (1996) model of 1993 attendance. The variables, once again, fall into these categories: location/stadium issues, time and weather factors, expected quality of the game issues, and special factors. Expected quality of the game variables are thought to be more important in a model of regular season attendance when compared to that of spring training because fans are more interested in a well-contested baseball game, as opposed to just seeing a game. The following is an introduction to the explanatory variables and their predicted relationship with the dependent variable: attendance.

The location and stadium variables thought to positively affect attendance are: population, income, and capacity. Population and income are thought to be more of a determinant of attendance in this model than in the spring training model because the fans attending the game are more likely to live in the area where the game is occurring. Capacity may not be significant
because MLB stadiums tend to be large and are rarely ever filled to capacity\textsuperscript{16}, but it is included to analyze this hypothesis. Location and stadium variables thought to have a negative relationship with attendance are: price, fan cost index, stadium age, whether the game is played indoors, whether there are two baseball teams playing in the area, black population, and Hispanic population. The result on the stadium age and indoor variables are of interest because it will test the hypothesis that building a new, retractable roof stadium will increase attendance and revive financially struggling teams. This is a current issue that MLB is dealing with.

Expected quality of the game variables that have a predicted positive relationship with attendance are: combined division championships in the past three years of the two teams competing, home team wins and runs in the last 10 games, the home and visiting starting pitchers' career records, the home and visiting starting pitchers' current records, home and visiting All-Stars playing in the game, and league leaders playing in the game. Expected quality of the game variables thought to be negatively correlated with attendance are: combined games back the two teams playing finished last year, and the home and visiting team's current games back. The division championship and games back variables are included to analyze the effect of team quality on attendance, while the pitcher's record, All-Stars, and leader variables are included to investigate the impact of individual player talent on attendance.

The only time and weather variable with a predicted positive relationship with attendance is whether the game is played on a weekend. The variables that are thought to have a negative impact on attendance are: if the game is played during the day, if it is played early in the season, if a rain delay occurs, and if it is extremely hot or extremely cold.

\textsuperscript{16} 127 (5\%) games were filled to capacity in 2002.
Interleague games, special stadium promotions\textsuperscript{17}, and games within the division have a predicted positive relationship with attendance. Games of a double header and the distance to the visiting team's city are thought to reduce the number of fans. Although the sign and significance on the contraction variable are difficult to predict, the results of the spring training model leads to a predicted negative relationship with attendance. It is thought that disgruntled fans and loss of off-season tickets sales outweighed any novelty or pride effect. A Twins and Expos dummy will again be included in one regression to analyze the separate effects. It is predicted that the Expos will have significantly less attendance while the Twins may not.

\[ \text{ATTENDANCE} = f (\text{POP}, \text{INC}, \text{PRICE}, \text{FANCOST}, \text{STADAGE}, \text{CAP}, \text{INDOORS}, \text{BLACKPOP}, \text{HISPPOP}, \text{2TEAM}, \text{GBLAST}, \text{DIVCHAMP}, \text{GBHOME}, \text{GBVIS}, \text{WINS10}, \text{RUNS10}, \text{HPITCHCAR}, \text{VPITCHCAR}, \text{HPITCHCUR}, \text{VPITCHCUR}, \text{LEADERS}, \text{HOMESTAR}, \text{VISTAR}, \text{WEEKEND}, \text{DAY}, \text{APRILMAY}, \text{COLD}, \text{HOT}, \text{RAIN}, \text{INTERLEAGUE}, \text{DBLHDR}, \text{RIV}, \text{PROM}, \text{VISPROX}, \text{CONTRACT}) \]

A definition of each variable can be found in table 2.

**Data and Methodology for the Regular Season Model**

Data for regular season games are more accessible than spring training data. Most information is found in the box score, but some variables, such as pitchers' records and demographic factors, must be found from another source. The following is a description of the data collected (see Appendix 2 for a summary of the data) for this model and the sources used to find the data.

\textsuperscript{17} Game day promotions are not included because they were unavailable.
## Table 2

**Dependent Variable**

**ATTENDANCE***
The number of fans in attendance according to the box score.

**Independent Variables**

### Location variables

**POP***
The population of the home city according to the 2000 census.

**INC***
Per capita income of the home city according to the 2000 census.

**PRICE***
The average price of a ticket for the home team.

**FANCOST***
A measure of the cost for a fan to attend the game. Includes parking concessions.

**STADAGE***
The age of the stadium the game is being played in.

**CAP***
Capacity of the stadium the game is being played in.

**INDOORS***
A dummy equaling 1 if the game is played indoors.

**BLACKPOP***
The black population of the home city according to the 2000 census.

**HISPPOP***
The Hispanic population of the home city according to the 2000 census.

**2TEAM***
A dummy equaling 1 if there is another team playing in the same city as where the game is being played.

### Quality of the game variables

**GBLAST***
The sum of games back the two teams playing finished last season.

**DIVCHAMP***
The number of division championships won by both teams in the last three seasons.

**GBHOME***
The current number of games back for the home team.

**GBVIS***
The current number of games back for the visiting team.

**WINSLAST10***
The number of wins in the last 10 games for the home team.

**RUNSLAST10***
The number of runs in the last 10 games for the home team.

**HPTITCHCAR***
The home team’s starting pitcher’s career wins minus loses.

**VPTITCHCAR***
The visiting team’s starting pitcher’s career wins minus loses.

**HPTITCHCUR***
The home team’s starting pitcher’s in-season wins minus loses.

**VPTITCHCUR***
The visiting team’s starting pitcher’s in-season wins minus loses.

**LEADERS***
The number of players playing in the game that finished in the top three for RBI, home runs, or batting average in 2001 or 2002.

**HOMESTAR***
The number of players on the home team’s roster that played in either the 2001 or 2002 All-Star game.

**VISSTAR***
The number of players on the visiting team’s roster that played in either the 2001 or 2002 All-Star game.

### Time and Weather Factors

**WEEKEND***
A dummy equaling 1 if the game is played on Friday, Saturday, or Sunday.

**DAY***
A dummy equaling 1 if the game is played before 5 p.m. local time.

**APRIL/MAY***
A dummy equaling 1 if the game is played in April or May.

**COLD***
A dummy equaling 1 if the game time temperature is 55 degrees or lower.

**HOT***
A dummy equaling 1 if the game time temperature is 90 degrees or higher.

**RAIN***
A dummy equaling 1 if there is a rain delay during the game.

### Special Factors

**INTERLEAGUE***
A dummy equaling 1 if the teams playing are not in the same league.

**DBLHDR***
A dummy equaling 1 if the game is part of a double header.

**DIVRIV***
A dummy equaling 1 if the team playing are in the same division.

**VISPROX***
The distance, in miles, between the home and visiting city.

**CONTRACT***
A dummy equaling 1 when the home team is either the Twins or the Expos.

*indicates that the log of the variable is taken
Attendance figures are collected from the box scores that appear on www.espn.com. These figures are ticket sales for the 2002 home regular season games of the thirty Major League Baseball teams. The data set includes a total of 2,426 games. The league average for attendance was 28,169 fans per game. Minnesota's average was 23,579 fans per game, while the Expos' was 10,025. It is interesting to note that the Twins average attendance was less than the league average, despite a very successful season. The Yankees had the highest average with 43,740 fans per game, while Montreal's was the lowest.

Population and income are demographic variables included in most models of sports attendance, as it is thought that they both positively affect the number of people at games. These are collected from the 2000 census found at www.census.gov. Population and income are measured for the entire metropolitan area for the city where the game is being played. This is thought to be a better measure than using the city limits because fans are likely to be drawn from the entire metropolitan area for games. Income is measured in terms of per capita income. The average population of the cities where MLB teams are based was about 5.5 million; the average per capita income was $24,288. Both Montreal and Minnesota had populations below the average, and the income in Minneapolis is above that of the league average.

The price variable is taken from a student user's website for Rodney Fort's textbook *Sports Economics*. It is an average ticket price that was collected from www.teammarketing.com. The league average price of a single ticket in 2002 was $18.29, while it was $11.78 and $9.00 for the Twins and Expos, respectively. The Red Sox had the most expensive tickets with an average of $38.68, and the Expos had the least expensive. The fan cost index (FCI) is also collected for all thirty teams from Fort's website. This is a measure of how much money a family will spend at a baseball game. This includes ticket, concession, parking, and souvenir prices. The Red Sox again

18 Data cover consolidated metropolitan statistical areas (CMSAs).
had the highest FCI, and Montreal had the lowest. The average cost of a family to attend a baseball game in 2002 was $142.18.

Stadium age is how many years it has been since the stadium was built (in 2002), and does not take into account any renovations to the stadiums. The average age of MLB stadiums was 23 years old; Minnesota's stadium was 20. Boston's Fenway Park was the oldest stadium (90 years), while there were a few stadiums that opened more recently, such as Milwaukee's Miller Park, which opened in 2001. The average capacity of a MLB stadium was about 47,000. Both Montreal's and Minnesota's stadiums are near this figure. Boston's park has the smallest capacity, while San Diego's has the largest. 20% of the games in the 2002 regular season were played indoors. Minnesota, Montreal, Toronto, and Tampa Bay had all of their home games in domes, while Seattle, Arizona, Houston, and Milwaukee used their retractable roofs in 53% of their home games.

The Black and Hispanic population, like total population and income, are collected from www.census.gov. This is the number of people reported as being Black or Hispanic in the metropolitan area of the city where the game is being played. Minneapolis and Montreal had below the league average in both of these categories.

Combined games back that the home and visiting teams finished last season and combined division championships in the last three seasons examine the effect of overall team quality on attendance. The average number of games back that the home and visiting teams finished last season was 28, while the Twins' games involved teams that finished an average 25 games back. The Expos and their opponents finished an average 28 games back. There were an average 1.2 division championships won in the last three seasons per game in 2002. Current games back for each team competing in the game are a measure of current season performance. While the league
average for the home team was about 9 games back, the average amount of games back for the Twins was under 1, as they were in the division lead most of the season. The Twins had an easier schedule than both the Expos and the rest of the MLB (average), playing teams that were on average 10 games back.

Runs in the last ten games and wins in the last ten games are included to investigate the effects of recent game success on attendance. The Twins had an average of 5.7 wins in the last ten games, while the Expos had the same amount of wins as the league average, 4.7. The same pattern can be seen for runs in the last 10 games.

The quality of the players competing in the game is measured by the starting pitchers' career and current records, All-Stars on each team's roster, and league leaders competing in the game. The home team's starting pitcher had an average 10 more wins than losses during his career, while the visiting pitcher had about 9.5 more wins than losses. Twins' starting pitchers in 2002 had about 3 more wins than losses over their career; Expos' pitchers had 2 more losses than wins.

Both these teams played against starting pitchers that had a current record better than the league average. The league leaders variable is measured as the number of players playing in the game that finished in the top three in RBI, home runs, and batting average last season or the current season. The league average was about 2.5 leaders per game, while the Twins' and the Expos' home games had 1.4 and 1.9 leaders per game, respectively. The number of All-Stars is the number of players on each team that competed in either the 2001 or 2002 All-Star game. It is believed that since the All-Star game is in the middle of the season, the number of players that competed in just the 2002 All Star is not a good measure of star players. The average amount of star players on each team in 2002 was 4.1. Both the Twins and the Expos played opponents with about 3.6 stars on their roster.
The time and day dummy variables in the model are weekend games, day games, and games that occur in April/May. A weekend game is considered a game that occurred on Friday, Saturday, or Sunday. 48% of the games in 2002 occurred on the weekend, and the Twins and the Expos have this same percentage. Day games are games that began sometime before five o'clock, local time. The Chicago Cubs traditionally have the most day games, and this holds true for 2002 (74% of their home games). While 26% of the Twins games were day games, the entire league played 34% of its games during the day. About one-third of the games occurred in April or May, which makes sense because the season is six months long.

Weather variables are collected from the box scores of each game. It was considered cold if the game time temperature was 55 degrees or below and hot if it was 90 degrees or above. If there was a rain delay reported in the box score, the rain dummy took the value 1. 6% of the games occurred in cold weather, 7% were played in hot weather, and 2% had a rain delay sometime during the game. The Chicago Cubs had 20% of their home games (16 games) in temperatures below 55 degrees, while the Baltimore Orioles competed 32% of their home games (26 games) in temperatures above 90 degrees. 14% of the Pittsburgh Pirates' home games (11 games) had a rain delay in them. The average game time temperature for a 2002 regular season game was 74 degrees. The Florida Marlins and the San Francisco Giants had the highest and lowest average game time temperatures: 83 and 63 degrees, respectively. The Twins and the Expos both play in domes so these variables did not affect their attendance.

Special factors are any variables that are not included in the other categories. Interleague play has been an issue as it was put into place recently to increase attendance. About 10% of the league's games were interleague games. Most teams nine interleague games, except for the NL Central, where each team played six. This is due to the fact that the NL Central has more teams in
it than any other division and interleague games are usually scheduled so a division from the AL plays an entire division from the NL. 2% of the games were double headers, and the Cubs had the most double headers with four of them. Recently, MLB went to an unbalanced schedule to offer more interdivisional match-ups during the season. 43% of the games in 2002 were contested between teams in the same division. Distance the visiting team had to travel was calculated using www.yahoomaps.com. The visiting team had to travel an average 1,278 miles to each game.

162 games were competed in either Minneapolis or Montreal, so approximately 7% of the games take on the value 1 for the contraction dummy variable.

**Specifications**

The variables discussed above are run in four different regressions because of the possible multicollinearity in the location and home team specific variables (discussed in the spring training section)\(^\text{19}\). Running multiple regressions is an approach used to realize the best model.

Specification 1 is based on the first regression run in Bruggink and Eaton's (1996) study of 1993 game day attendance, while 2, 3, and 4 are unique to this model of 2002 attendance.

Specification 1 includes all the variables in the model presented above and, for this reason, is likely the victim of multicollinearity. Specification 2 attempts to lessen the effects of this multicollinearity by removing all but three location variables. Specification 2 is also more of a basic model of demand for baseball games as it includes income, price, population, one set of team quality variables, one set of player quality variables, time and weather variables, and the contract dummy variable. Specification 3 is thought to be the most accurate model for game day attendance, while specification 4 includes the same variables, except for the inclusion of a Twins

---

\(^{19}\) See Appendix 3 for Variance Inflation Factors.
and Expos dummy variable to examine the separate effects of contraction on these two teams.
Again, a semi log functional form is used to simplify interpretation on the coefficients, and heteroskedasticity is corrected for using White-Heteroskedasticity-Consistent Standard Errors.

Analysis of the Results of the Regular Season Approach

The results for the regular season model can be seen in Appendix 3.

Population is found to have a significant positive effect on attendance in three of the four specifications. The interpretation of the coefficient in Specification 3 is: if population increases by 1%, then attendance will increase by .15%. This is consistent with theory and is similar to results found in other papers. MLB has argued recently that small market teams cannot survive, and while the results do not show the correlation between population and financial stability of a team, they do demonstrate that the size of the market does have a positive effect on attendance, which in turn affects money generated from ticket sales. Income was found to be statistically significant in all four specifications. The positive coefficient indicates that as income rises, more fans will attend games. This is consistent with the theory, but is different from the results that Bruggink and Eaton found.

The average price of a ticket is only included in specification 1 because it is thought that the fan index is a better measure of cost to the consumer. It does, however, have an unpredicted positive coefficient. The Fan Cost Index variable is significant in all of the specifications, and has positive coefficients. The results in specification 3 imply that the price elasticity of a baseball game is about 1.18. The positive coefficients could mean that baseball games are price inelastic and ballparks can charge a higher price for the same number of fans in attendance. This is a possible explanation for the extremely high prices in stadiums in Boston, Seattle and New York.
Stadium age is found to be statistically significant with negative coefficients in equations 1, 3 and 4. This result reinforces the claim that newer stadiums draw more fans. The coefficient in equation 3 means that a stadium that is one year older will witness 0.2% lower attendance. This implies that if the Twins built a new stadium before the 2002 season, they would have had 4% higher game day attendance. Capacity of the stadium is statistically significant in these three equations, and the predicted positive coefficients are witnessed. An increase in 1% of capacity of the home stadium leads to a .64% increase in attendance for equation 3, ceteris paribus. Other studies leave this variable out because it is thought that an increase in seats does not necessarily lead to an increase in people in those seats, but here it is found significant. These two results show that MLB teams could increase their attendance by renovating their existing stadiums or building stadiums that have more seats. Although this is expensive, they could analyze how much revenue can be gained from these stadium improvements and see if it outweighs the costs of renovating.

The final stadium variable included in these specifications is whether the game is played indoors. Teams such as Montreal and Minnesota play all their home games indoors, while teams such as Houston and Arizona play some of their home games indoors with their retractable roof. In equation 3, indoor games witness 6.7% less attendance. According to this model, if the Twins played outdoors they would have about 1,500 more fans at each game. This is simplified, however, because when the game is being played indoors, all the other weather variables are constant. The results in the other two equations are similar indicating that fans prefer an outdoor game.

The black population for the metropolitan area is used in three regressions, while Hispanic population is only used in one. The positive significant coefficients found on the Hispanic variable are not consistent with theory and results from other papers. This result could be biased.
because of population numbers. Cities such as New York and Los Angeles have higher attendance than smaller cities, and these cities are also more likely to have a high minority population. However, it is interesting that the black population variable is found to be negatively significant in all of the equations it is included in. One would think that the larger city effect mentioned above should be the same across minority groups. The coefficients on the variable are near 0.15, meaning that a 1% increase in the black population leads to a 0.15% decrease in game day attendance. It is thought that these results mean that black population is a significant determinant of attendance, while Hispanic population may not be.

The final location variable included in the first specification analyzes the substitution effect of another baseball team in the area. It is found to be significant, but with an unpredicted positive sign. The population numbers could also be the cause of this surprising result, as cities with two teams, such as New York and Los Angeles, have a larger population than cities that don't. There are enough people in the area to attend both teams’ home games, so there is no substitution effect.

The combined games back variable is found insignificant in the only regression in which it is included, meaning that fans were not especially drawn to games by teams that did well last season. However, the combined division championships variable is found significant in specifications 3 and 4. The coefficients are near 0.04, meaning that if one of the two teams competing in the game had won one more division championship in the last three years, attendance would be about 4% higher. The fact that division championships is found significant and last year's games back is not indicates that fans prefer more notable success over a longer period of time to the modest success of the previous year.

The home and visiting teams' current games back are thought to be a very important determinant of attendance as they measure how well a team is playing over the course of the
season. It is a measure of competitive balance between the teams playing and what effect that has on attendance. Both of these variables are found to have a significant negative effect on attendance. The coefficients on home team's games back is near 0.015 for every specification, meaning that if the home team is one game further back from the division leader, attendance will be 1.5% less, while a 0.2% drop in attendance will be witnessed if the visiting team is one game further back. The fact that the magnitude of the home team variable is larger is consistent with theory, as it is thought that success of the home team is more of a factor in determining attendance than success of the visiting team. This is because the fans of the home city want to see their team doing well and are not as drawn by a visiting team doing well.

The next two variables analyze how the success of the home team in the last ten games affects attendance. Wins in the last ten games and runs in the last ten games are included in Specification 1, and found to be significant. The positive coefficient on wins in the last ten games implies that one more win in those games increases attendance by about 0.7%. The runs variable received an unpredicted negative coefficient. These two variables are likely highly correlated, and that could be the cause of this result. It is thought that these variables are not as important as the games back variables because they do not take into account whether the team is in the pennant race, which likely draws many fans.

The next seven variables are included to study the impact of individual player success on attendance. Career records for the home and visiting starting pitchers are included in three of the regressions, while current records are only included in one. It is thought that career records are a more important determinant of attendance because people come to see pitchers such as Roger Clemens and Pedro Martinez, who have been good their whole career, and not pitchers who are having a single good year. Career wins minus losses for the home team pitcher is found to be
significant, while the visiting pitcher's record is not. If the home team pitcher has one more career
win, or one less loss, then attendance will be about 0.1\% higher. It is interesting that the visiting
pitcher's record has no significant effect because fans will only get to see that pitcher a few times
in a season, while they can see their home team starting pitcher about every five home games.

Starting pitchers' current records follow this logic, as the visiting pitcher's record is found to have
a significant positive effect on attendance, and the home pitcher's record has no significant effect.
The number of combined league leaders from last season and this season competing in the game is
found to have a significant positive impact on attendance. It was only included in one of the
specifications because it is thought that All-Stars on each team is a better measure for star players
competing in the game. Therefore, home and visiting All-Stars on each team are included in all of
the specifications. The results show that an addition of one All-Star on the home team will
increase attendance by about 1\%, and an addition of one All-Star on the visiting team will increase
attendance by about 1.5\%. The fact that visiting All-Stars draw more fans than home All-Stars can
be explained by the same logic used with visiting and home pitchers: fans can see their home all-
stars 81 times a year, while they can only see the visiting All-Stars on a few occasions.

Time variables included in all four models are whether the game occurs on the weekend,
whether the game happens during the day, and whether the game happens in the first two months
of the season. The results for the weekend variable are consistent with theory, as a weekend game
brings in about 23\% more fans. The day game variable is found to significantly increase
attendance, while its predicted relationship was negative. This can be explained by the fact that
many day games feature special promotions, and that there is nostalgia attached to watching
baseball outside during the day. If a game was played during the month of April or May, about
20\% fewer fans were in attendance. This result is consistent with theory.
The cold, hot and rain variables are found to be statistically significant in all the specifications. The results on these dummy variables indicate that: if there is a rain delay, there are about 13% fewer fans; if it is under 55 degrees, about 8% fewer fans attend; and if it is over 90 degrees, attendance increases by about 12%. The signs on the coefficients are consistent with theory for the rain and cold dummy variables, while the coefficient on the hot variable is surprising. It was thought that hot temperatures might stop fans from coming to the game, but it seems that fans are actually attracted by hot temperatures. This result shows that people prefer to go to a game in hot weather compared to one that is being played in average temperatures.

Special factors found to significantly reduce attendance in any of the specifications are an interleague dummy, a division rivalry dummy, and a visitor proximity variable. The interleague dummy is found to significantly reduce attendance in two of the specifications. Specification 1’s results imply that an interleague game has 8% less attendance than an intraleague game. This contrasts Butler’s results, but he did indicate that these results might not be showing what is really going on. The negative sign on the division rival dummy variable is inconsistent with the theory that teams playing each other for the pennant are likely to draw more fans. The sign could be a result of the fact that there is now an unbalanced schedule in baseball and teams play other teams in their division many times. This could lessen the drawing power because fans have a chance to see the visiting teams from the home team’s division many times throughout the season. The visitor proximity variable is found to be significant in all three of the regressions in which it is included. The coefficient on the variable in regression equation 3 indicates that a 1% increase in the distance the visiting team travels leads to a 0.05% decrease in attendance. This makes sense, as fans are more likely to travel to a visiting stadium to see their team the closer they are to that

---

stadium. The doubleheader variable is not found to be significant. This is probably due to the fact that fewer than 2% of the games in 2002 were doubleheaders.

The contraction dummy variable is included in the first three regressions and found to be significant in all of them. The negative coefficient implies that the two teams that were pegged for contraction in 2002 realized less attendance than the rest of the league. The coefficient in specification 3 is about -0.35, meaning that the Twins and the Expos had about 35% less attendance compared to the other teams in baseball. This result is similar to the one seen in the spring training results. In the final specification, separate Twins and Expos dummy variables are included to analyze the separate effects of contraction on attendance. Both these dummy variables are found to significantly reduce attendance. The coefficients on the Twins' and Expos' dummy variables indicate that Minnesota had about 18% less attendance than the rest of the league while the Expos' had about 58% less. This is consistent with theory as Minnesota's success in 2001 and 2002 season likely drew more fans, while the Expos remained a losing team. While it cannot be said for sure that contraction is the sole cause of the lower attendance, as these two teams may just witness lower attendance perennially, it is certainly viable to say that contraction caused some of the fans to stay away from these teams. The theory behind these results is that off season ticket sales where hurt by the possibility of the team not being there and that fans of these two teams became disgruntled with baseball. The panel model below will attempt to answer the question of whether the Expos and the Twins witnessed lower game day attendance in the contraction year compared to previous seasons.
Model of Panel Data Attendance Over 3 Years

This model is very similar to the 2002 regular season model presented above, but it is a panel data set modeled over three seasons. There are two different models: one for the Minnesota Twins and one for the Montreal Expos. The data set includes all of the home games for the Minnesota Twins and Montreal Expos from the 2000 season to 2002 season in order to compare the attendance of these two teams in the contraction year to their attendance in the previous two seasons.

These models include all the same variables as the regular season model presented above except for the location variables (population, inc, price) and the weather variables. The location variables are not included because each of the two models has only a single location, Minnesota or Montreal. Therefore, these variables will be the same, or similar, throughout the years studied. The weather variables are not included because all of the home games for both of these teams are in domes, and are not subject to adverse weather conditions. The doubleheader variable is also excluded because neither of these teams had any doubleheaders in the 2000-2002 seasons. This is probably due to the fact that most of the doubleheaders that occur are due to rainouts, and Montreal and Minnesota play in domes, and therefore have no home games rained out.

Although these models are similar to the regular season one, it is probable that the results will not be the same due to the fact that the sample size is somewhat smaller. The general model presented for team i is:
ATTENDANCE, = f (GBLAST, DIVCHAMP, GBHOME, GBVIS, WINSLAST10, RUNSLAST10, HPITCHCAR, VPITCHCAR, HPITCHCUR, VPITCHCUR, LEADERS, HOMESTAR, VISTAR, WEEKEND, DAY, APRILMAY, INTERLEAGUE, RIV, PROM, VISPROX, 2000, 2001)

A definition of all the variables can be seen in table 3.

**Data and Methodology for 3 Year Panel Model**

The collection of data (see Appendix 2 for a summary of the data) for the panel set is very similar to that of the regular season approach, as many of the variables are the same. Attendance figures for the Twins and the Expos are collected for the 2000-2002 regular seasons, 243 homes games for each team. The Twins averaged 20,000 fans per game over these three seasons and managed to increase their attendance each year, while the Expos drew about 10,000 fans per game and witnessed their highest average attendance in the year 2000.

The current games back variable is collected in the same fashion as in the regular season model. The Twins, with success in the last two seasons, averaged about five fewer games back than the Expos did, while the Expos' opponents averaged 7 games back and the Twins' averaged 8.5. This could imply that the Expos had a harder schedule over the past three seasons, but may also be a result of their division.

The player quality variables in this model are similar to the variables used in the previous models. The league leaders variable is defined as the sum of players in the game that were league leaders in either the current season or the previous season. The Expos' games
### Table 3

**Dependant Variable**

**ATTENDANCE**

The number of fans in attendance according to the box score.

**Independent Variables**

**Quality of the game variables**

- **GBLAST**: The total number of games back the two teams playing finished the previous season.
- **DIVCHAMP**: The total number of division championships won by the teams playing in the previous three seasons.
- **GBHOME**: The current number of games back for the home team.
- **GBVIS**: The current number of games back for the visiting team.
- **WINSLAST10**: The number of wins in the last 10 games for the home team.
- **RUNSLAST10**: The number of runs scored in the last 10 games for the home team.
- **HPITCHCAR**: The home team's starting pitcher's career wins minus loses.
- **VPITCHCAR**: The visiting team's starting pitcher's career wins minus loses.
- **HPITCHCUR**: The home team's starting pitcher's in-season wins minus loses.
- **VPITCHCUR**: The visiting team's starting pitcher's in-season wins minus loses.
- **LEADERS**: The number of players playing in the game that finished in the top three in RB1, home runs, or batting average in the previous or current season.
- **HOMESTAR**: The number of players on the home team's roster that played in either the 2001 or 2002 All-Star game.
- **VISSTAR**: The number of players on the visiting team's roster that played in either the 2001 or 2002 All-Star game.

**Time and Weather Factors**

- **WEEKEND**: A dummy equaling 1 if the game is played on Friday, Saturday, or Sunday.
- **DAY**: A dummy equaling 1 if the game is played before 5 p.m. local time.
- **APRIL/MAY**: A dummy equaling 1 if the game is played in April or May.

**Special Factors**

- **INTERLEAGUE**: A dummy equaling 1 if the teams playing are not in the same league.
- **DIVRIV**: A dummy equaling 1 if the teams playing are in the same division.
- **VISPROX**: The distance, in miles, between the home and visiting city.
- **2000**: A dummy equaling 1 when the game occurs in the 2000 season.
- **2001**: A dummy equaling 1 when the game occurs in the 2001 season.

*indicates that the log of the variable is taken.
averaged more league leaders than the Twins' games did over the seasons studied, while both
teams played opponents with about four all stars on their roster. Both teams had home starting
pitchers averaging a losing career record and opponents starting pitchers averaging about nine
more career wins than losses. However, the Twins starting pitchers averaged a winning current,
in-season record.

The Twins averaged more runs and wins in the last ten games than the Expos, and were able to
increase these two figures in each of the three seasons. Finally, similar to the regular season study,
about one-third of the games were played in April or May, one-half of them were on the weekend.
one-quarter of them were day games and one-ninth of them were interleague games. This is
somewhat consistent between both teams.

2000 and 2001 dummy variables are included to compare these years attendance figures to that
of 2002, the contraction year. Below is a description of the specifications run to analyze the
impact of contraction on attendance.

**Specifications**

Three different specifications for each team are run to realize the best panel data model for
describing attendance. Many variables must be left out of the equations because of perfect
multicollinearity. In Specification 1, the simplest equation, the games back variables are the
measure of current team success while the leaders variable is included to study the impact of star
players on attendance. The three time variables are included along with the 2000 and 2001 dummy
variables.

In Specification 2, the pitchers' career records are included to see if the prior success of the
starting pitchers affects attendance. The interleague dummy and the visitor proximity variables are
included because these were found to be a significant determinant of attendance in the regular season, and they are consistent with the theory of demand for baseball games.

Specification 3 includes all of the variables that can be included, taking into account the problem of perfect multicollinearity. A double-log functional form is used to simplify interpretation of the coefficients and heteroskedasticity is corrected for in each of the specifications using White- Heteroskedasticity -Consistent-Standard Errors.

**Analysis of Results of the 3 Year Panel Data Set**

Results for the 3 year panel data set can be seen in appendix 3.

Twins games back is found to be significant in all three of the regressions. The negative coefficient can be interpreted as follows: if the Twins were one game further back from the division leader at the time of the game, then attendance would have decreased by 2%. The Expos games back variable is also a significant determinant of attendance, and attendance decreased by 1% with the Expos being one game further back. The fact that the Twins coefficient has a larger negative magnitude could be explained by their recent success. The Expos have been perennially at the bottom of the NL East while the Twins had recent chances at the pennant. This could cause the Expos fans to not be as responsive to a reduction in standing, as they are used to being at the bottom of the standings. The visitor games back variables are also found to have a significant negative effect on attendance and have a similar magnitude in both data sets. The coefficient is about .01, meaning that the visiting team being one game further back led to a 1% decrease in attendance. It is interesting to note that both the Expos and the visiting team moving one game further back affects attendance equally, meaning that fans are equally responsive to home and visiting teams current success.
The league leaders variable is included in all three regressions for both data sets. It is found to have a significant positive effect on attendance in all three of the Twins equations, with the coefficient being between 0.03-0.05. In the Expos equations, however, it is only found significant in regression 3. These results imply that Twins fans are more likely to come to a game where big time players are competing, while the Expos fans are less receptive to these players.

Visitor All-Stars and home and visiting starting pitchers' current records are included in the final regression. Each additional All-Star on the visiting team's roster added 4% attendance for the Twins, while visiting team All-Stars had no significant effect on the Expos' attendance. This again reinforces the idea of player draw in Minnesota versus Montreal presented above. Home starting pitcher's current record is found to be significant for both teams with coefficients being .03 and .02 for Montreal and Minnesota, respectively. These results indicate that the fans in Minnesota and Montreal are not really drawn by the current success of a visiting pitcher, but rather the prospect of winning the game because their team has a winner on the mound. Starting pitchers' career records are only included in the final regression and the results show that these are not significant determinants of attendance. The average career wins minus losses for starting pitchers for both these teams are negative, meaning that their pitchers have not had much success in their career. This fact is probably the reason for the result presented above, as fans are not drawn to pitchers that have losing or, close to losing, career records.

Wins in the last ten games for the home team is only found to be significant in Twins regression 3. The fact that the Expos did not have very much success over the three seasons studied could explain why Twins fans reacted to a streak, and Expos fans did not. Similar to the results of the regular season model, runs in the last ten games is not found significant.
Games that occurred on the weekend have near 40% more attendance for the Twins and 35% for the Expos. This dummy variable is found significant in all six regressions run. The April/May dummy variable is found to significantly reduce attendance in all of the Twins equations and two of the Expos equations, while the day game variable is significant in all the Expos equations and no Twins equations. This is an interesting result that may need further investigation to explain. It could be that, in order to attract more fans, the Expos have more promotions for their day games than do the Twins.

The interleague game dummy is found to be significant in two equations in the Twins data set and one in the Expos. The coefficients on the variables imply that an interleague game will increase attendance for these two teams by about 20%. This contrasts the full league results as it implies that Montreal and Minnesota fans are drawn by opponents in the other league. Visitor proximity is found to be significant in the Twins regressions with a coefficient near -.12. This means that a 1% increase in the amount of miles traveled from the visiting team's city, leads to a .12% decrease in attendance. This variable in the Expos regressions is found to be insignificant. The theory behind this variable and this result suggests that fans from visiting cities are less likely to travel to Montreal than Minneapolis to see their team play. This could also be explained by the fact that Montreal has not had success recently, so fans will not be drawn from distances to see a game. The fact that they have to cross the border into Canada may also deter fans from making the drive/flight.

The 2000 season dummy variable is found significant in the three regressions run with the Twins data. The negative coefficient suggests that the Twins had 32%-44% less game day attendance in 2000 compared to the contraction year of 2002. The 2001 dummy variable is also found significant in all three Twins equations and has a significant coefficient of around -.90.
meaning that attendance in 2001 was about 90% less than the 2002 season. These results imply that the Twins witnessed higher attendance in 2002, the contraction year, than two previous seasons. While it may not be solely because of the contraction possibility, it is likely that that did have an effect. The success of the Twins in 2002 is likely a big reason why these results are witnessed, but the fact that the Twins were on the brink of elimination may have caused fan interest to increase. The results on the Expos dummy variables are similar but not the same. The 2000 dummy variable is found significant and positive in two of the equations. The coefficients of .10 to .20 mean that the Expos had about 10-20% more attendance in 2000 compared to 2002. However, the 2001 dummy variables are found significant and to have a negative effect on attendance. This means that while attendance from the 2001 to the 2002 season increased, the 2002 Expos had less game day attendance than they did in the 2000 season. The low r-squared value in the Expos regressions leads to questions about the validity of this result. While the Twins results show that attendance increased in the contraction year, the results from the Expos equations are likely not reliable.

Conclusion

These three models were built to examine the effects of possible contraction on 2002 attendance, and while these models accomplished this, there were many other results that are worth noting.

The results seen in the spring training model are particularly appealing because no other such model exists; these are results of something never examined before. One major difference between a study of spring training attendance and a study of regular season attendance is that there must be a variable that accounts for tourist traffic in the area. An interesting note about the results of this
paper is that the variable that accounts for tourist traffic, SERVWORK, is found to be positively significant, meaning that the more workers in the service sector there are, the more fans there will be at spring training baseball games. Another variable unique to spring training found to be significant is the split squad variable. This result implies that at a home split squad game, fewer fans are in attendance. The week dummy variables also lend insight to a unique spring training model because the coefficients on the variables show that as the spring training season goes on, more fans will be in attendance. One surprising number is that having two teams playing in the same spring training area does not negatively affect attendance. In fact, it received a positive coefficient.

While the spring training model revealed some groundbreaking results, the regular season model reinforced some previous finding by other authors. Like Bruggink and Eaton, significant positive coefficients were found on population, ticket prices, wins in the last 10 games, league leaders, visiting All-Stars, the weekend game dummy, and the game day dummy. The positive coefficient on ticket price is somewhat surprising, as according to the model of demand: as price rises, quantity demanded falls. This supports the claim that ticket prices are inelastic for baseball games. It seems as though the owners can raise the ticket price without seeing attendance drop. Stadium age, current games back and a few of the weather variables were also found significant in both studies, but they received negative coefficients. Unlike Bruggink and Eaton, division rivalries were not found to significantly increase attendance. This could be explained by the unbalanced schedule now played in baseball. The home team plays the teams in its division many more times then they used to, possibly making each game with the division rival less interesting. Some interesting findings not seen in Bruggink and Eaton's paper are effects of indoor and interleague games on attendance. The result on the indoor game variable confirms the often-
presumed fact that fans prefer outdoor baseball, while the interleague play having a negative effect on attendance contrasts previous notions. The visitor proximity variable is also unique to this model and is found to have significant negative effects on attendance. This variable could be a measure of two different things. Teams that are relatively close in distance are more likely to be rivals, such as the Red Sox and Yankees: when these rivalries are played, more people are likely to attend the game. Also, it could simply be measuring the fans that are willing to drive or fly to a visiting team’s stadium to watch a game: the closer they are in distance, the more likely they are to make that drive/flight.

The three-year panel data set is somewhat similar to the regular season, so the results should be consistent. It is interesting to note, however, the differences between the two teams examined, and the differences with the entire league. Leaders and visiting All-Stars are found significant in the Twins and the rest of the league’s regressions, while the Expos’ fans respond somewhat less to star players. A noteworthy result is that individually, both these teams drew more fans when their home pitcher had a better in-season record, while their pitchers’ career record did not have any effect. The final interesting result is the coefficient found on the day game variables for the Expos and the rest of the league versus the Twins. Day games seem to draw more fans for the Expos, while the Twins’ fans do not respond to games played during the day.

The overall purpose of this paper was to determine the effects, if any, of possible contraction on game day attendance for the Minnesota Twins and the Montreal Expos. Three approaches were used to examine these effects and mixed results were realized. The spring training approach found that the Twins and the Expos, as one contraction unit, had less game day attendance for the 2002 spring training season. Although it is not certain that the fewer fans in attendance were caused by the threat of contraction, it is possible that fans were not interested in seeing two teams
because they were on the brink of elimination. If the threat of contraction was the cause of the fewer fans, it is thought that off season ticket sales were harmed by the uncertain future, and that fans of these teams became disgruntled toward baseball because of the contraction threat. However, when separate Twins and Expos dummy variables were included, it was found that, while the Expos remained to have significantly less attendance, the Twins variables dropped out as being significant. This could mean either the Twins did have less attendance, but not significantly less, or that there is some correlation issue involved. If the Twins did have less attendance, but not enough to be deemed significant, it is thought that the Expos dominated the combined result above. This outcome is possibly caused by the fact that Twins fans were fighting more to show MLB that their team deserved to exist. Multiple court hearings kept the Twins from being contracted, and they were finally taken permanently off the contraction list in May of 2002, while there is still talk of moving or contracting the Expos today. The split result could signify that Twins fans believed that their team was still going to be around, while the Expos' fans had little to cheer or hope for.

Like the spring results, the contracted teams had less game day attendance than the rest of the league. The magnitudes are similar to the spring results, showing that the Twins and the Expos had about 20-40% less attendance than the rest of the league. Also similar to the spring training model, it cannot be said for sure that contraction is the cause of this lower attendance; the Expos and Twins may have had lower attendance without the contraction threat. If the threat was the cause of the lower attendance, it is thought that off season season-ticket sales were hurt by the threat, as fans would not buy tickets for games that may not actually happen. It is also possible that fans were fed up with baseball and upset that their team was even considered for contraction. The same separate dummy variables were included in the regular season regression and it was found that both the Twins and the Expos individually had less attendance than the rest of the
league. This contrasts the findings in the spring training model. It is, however, consistent with the combined contraction dummy variable results in both models. The magnitude on the Expos variable was much larger than the Twins variable which can be explained by the recent success of the Twins compared to the Expos. The reason why the Twins had significantly less attendance in the regular season and not in spring training is tough to decipher, but it is thought that off season ticket sales were the primary reason. Spring training games do not sell out as much as regular season games, so there is not as much of a need to buy season tickets or advanced tickets; it is likely that most spring training tickets sales happen at the gate. So while the threat of contraction was looming in the off season and many sales were lost for the regular season, spring training ticket sales were less affected because they were sold at the gate the day of the game, after the threat of contraction was removed.

Finally, the three-year panel data set approach was used to compare the Twins and the Expos game day attendance in 2002 to that of the previous two seasons. It is thought that comparing a team to itself is the best way to determine how contraction affected attendance. The Twins results indicated that in 2002, they had more fans per game than the previous two seasons. The Twins had a highly successful season in 2002, which could be skewing this result. But if, in fact, contraction was the cause of this higher attendance, this contrasts results presented earlier, and Twins fans actually came out to support for a team that was almost eliminated before the season. This story can be told consistently, however. The higher attendance compared to previous seasons means that more fans came to games in 2002, probably because of the success of the team, and possibly because they wanted to show positive support their team that was almost eliminated. But while this higher attendance was realized, the Twins still had fewer fans in attendance in the 2002 season compared to the rest of the league. This could prove that the Twins may still be a good candidate
for contraction, as a successful season with increasing attendance still does not compare with the
other teams in the league. The Expos results in this model are questionable due to low r-squares
and insignificant variables, but it too indicates that Montreal improved on attendance in 2002
compared to the previous seasons. This can be explained by the same reasoning as above except
for the fact that the Expos did not have a successful season in 2002. So the increase in attendance
may have been solely caused by support of fans who didn't want to see their team go.

The results in these models are mixed, but it can be said that the Twins had more attendance in
2002 than in previous seasons, that both the Twins and the Expos had less attendance in 2002 than
the other teams in the league, and that the Expos had fewer spring training fans in 2002 than the
other teams in Florida. While it cannot be said for sure that the contraction threat was the primary
cause for any of these effects, it is probable that it did have some effect on the Twins and the
Expos attendance, whether positive or negative. A further study, possibly involving a panel data
set for all the MLB teams, could perhaps explain the actual effects of the contraction threat on
game day attendance.
Oct. 24: The Windsor Star of Ontario reports that Major League Baseball will buy out the Montreal Expos and Florida Marlins after the World Series and fold them. Selig says that "no definitive decisions" have been made on contraction.

Oct. 26: The Miami Herald reports that the Minnesota Twins are a more likely candidate for contraction than the Marlins.

Oct. 29: Selig says contraction is possible by the start of 2002.

Oct. 31: A Minnesota state legislator says a top baseball official informed him the Twins and Expos are the two teams under consideration. Selig declined to name the teams. The Minnesota attorney general's office begins exploring legal strategies to fight elimination of the Twins.

Nov. 4: Luis Gonzalez's bloop single caps a two-run rally in the bottom of the ninth inning to win Game 7 of the World Series for the Diamondbacks.

Nov. 5: The Metropolitan Sports Facilities Commission, the board that operates the Metrodome, says it expects the Twins to fulfill their lease commitment for 2002 and threatens legal action if the lease is not fulfilled.

Nov. 6: At the owners' quarterly meeting in Chicago, owners vote 28-2 to eliminate two unnamed teams.

Nov. 7: The Players' Association files a grievance that the owners' plan to eliminate two teams violates their labor contract.

Nov. 10: Jim Pohlad, son of the Twins' owner, sends a letter to Twins employees that reads, "Our willingness to go along with contraction, if the commissioner so decides, has come from a feeling of hopelessness," he wrote. "Within the context of baseball's commitment, when we are posed the question, 'Why should the Minnesota Twins not be contracted?' we are unable to find a plausible answer."

Nov. 13: Florida attorney general Bob Butterworth subpoenas documents from Selig and the two teams in his state to find out if Marlins and Devil Rays are candidates for contraction.


Nov. 15: Selig tells the Star Tribune that Minnesotans critical of baseball's contraction plan should "look themselves in the mirror." Selig points out the Twins have made no progress toward building a new ballpark, and defended team owner Carl Pohlad, who has offered to fold the team in the contraction plan.

Nov. 16: A Minnesota judge grants an injunction requested by the Metropolitan Sports Facilities Commission to force the Twins to play their 2002 home schedule in the Metrodome. The Twins and MLB appeal the decision.

Nov. 24: Alabama businessman Donald Watkins says he's interested in buying the Twins and working to build a new stadium in the Twin Cities.

Dec. 4: Baseball arbitrator Shyam Das begins hearing testimony in the Players' Association grievance to block contraction.

Dec. 10: Officials from both the players and the owners indicate a deal is close that would delay contraction until 2003.
Dec. 11: A Florida judge refuses baseball’s plea to temporarily block a subpoena from Florida attorney general Butterworth to solicit information from MLB about possible contraction of the Marlins and Devil Rays.

Dec. 13: Talks between owners and players regarding contraction collapse.

Dec. 18: A federal judge blocks Butterworth’s investigation, saying baseball’s antitrust exemption prevents the inquiry.

The grievance by the Players’ Association to block contraction is recessed until Jan. 3.

Dec. 21: The commissioner’s office finally gives approval for teams to release their 2002 schedules and start selling tickets.

Dec. 23: Selig also says contraction could come as late as February.

Dec. 27: The Minnesota Court of Appeals begins hearing testimony on MLB’s appeal to lift an earlier decision that the Twins must fulfill their 2002 lease to play in the Metrodome.

Jan. 4: The Twins finally name coach Ron Gardenhire as their manager to replace Tom Kelly, who had retired. The Marlins remain without a manager and general manager, unable to make any offseason moves.

Jan. 8: It is learned that Pohlad loaned Selig and the Brewers $3 million in 1995, an apparent violation of baseball’s rules. The loan was made by Tempus Investment Corp., one of Pohlad’s companies. Selig was CEO of the Brewers and acting commissioner at the time.

Jan. 10: Watkins is given the go-ahead by MLB to contact the Twins about making an offer to buy the team.

Jan. 12: The Washington Post reports that contraction is likely to be called off for 2002 and that the Expos could be moved to Washington, D.C., for 2003. Hall of Famer Frank Robinson will likely be Montreal’s manager or GM for 2002, the paper also reports.

Jan. 22: The Minnesota Court of Appeals upholds the injunction that forces the Twins to uphold their 2002 lease on the Metrodome. The Twins and MLB appeal to the Minnesota Supreme Court.

Jan. 24: The union’s grievance hearing over contraction resumes, but was quickly recessed until the week of Feb. 4.

Jan. 28: Sources confirm that management officials told the Players’ Association that the Twins and Expos were the two teams under consideration for contraction.

Jan. 31: Sandy Alderson tells USA Today that contraction is still possible for 2002 up to Opening Day.

Feb. 4: The Minnesota Supreme Court refuses to consider MLB’s appeal of the injunction that forces the Twins to uphold their lease.

Feb. 5: Nine days before the start of spring training, contraction called off for 2002.

May. 30: Twins are permanently taken off the contraction list as a settlement for a lawsuit against MLB.

Source: espn.com
## Appendix 2

### Data for 2002 Spring Training

<table>
<thead>
<tr>
<th></th>
<th>ATTENDANCE</th>
<th>INC</th>
<th>POP</th>
<th>SERVWORKERS</th>
<th>CAP</th>
<th>STADAGE</th>
<th>AVEPRICE</th>
<th>2TEAM</th>
<th>TEAM</th>
<th>GBHOME</th>
<th>GBVIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minnesota Ave</td>
<td>6,092.31</td>
<td>28,514.00</td>
<td>48,208.00</td>
<td>2,657.00</td>
<td>7,500.00</td>
<td>11.00</td>
<td>10.50</td>
<td>1.00</td>
<td>6.00</td>
<td>19.38</td>
<td></td>
</tr>
<tr>
<td>Montreal Ave</td>
<td>3,473.63</td>
<td>54,945.00</td>
<td>39,328.00</td>
<td>2,117.00</td>
<td>7,200.00</td>
<td>5.00</td>
<td>11.75</td>
<td>1.00</td>
<td>20.00</td>
<td>12.69</td>
<td></td>
</tr>
<tr>
<td>League Ave</td>
<td>5,105.66</td>
<td>26,367.13</td>
<td>76,745.24</td>
<td>4,381.39</td>
<td>7,458.11</td>
<td>18.70</td>
<td>10.67</td>
<td>0.20</td>
<td>14.55</td>
<td>14.67</td>
<td></td>
</tr>
<tr>
<td>Max Avg</td>
<td>9,954.06</td>
<td>54,945.00</td>
<td>303,447.00</td>
<td>14,000.00</td>
<td>10,200.00</td>
<td>80.00</td>
<td>15.75</td>
<td>21.09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Min Avg</td>
<td>3,211.08</td>
<td>15,071.00</td>
<td>1,924.00</td>
<td>522.00</td>
<td>4,000.00</td>
<td>7.67</td>
<td></td>
<td>6.78</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>DIVCHAMPH</th>
<th>DIVCHMPVIS</th>
<th>HMSTAR</th>
<th>VISSTAR</th>
<th>LEADER</th>
<th>HMSS</th>
<th>VISSS</th>
<th>RAIN</th>
<th>COLD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minnesota Ave</td>
<td>0.00</td>
<td>0.69</td>
<td>3.00</td>
<td>1.75</td>
<td>0.56</td>
<td>0.19</td>
<td>0.25</td>
<td>0.00</td>
<td>0.13</td>
</tr>
<tr>
<td>Montreal Ave</td>
<td>0.00</td>
<td>0.19</td>
<td>1.00</td>
<td>1.63</td>
<td>0.00</td>
<td>0.13</td>
<td>0.19</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>League Ave</td>
<td>0.60</td>
<td>0.64</td>
<td>1.89</td>
<td>1.92</td>
<td>0.69</td>
<td>0.08</td>
<td>0.14</td>
<td>0.03</td>
<td>0.19</td>
</tr>
<tr>
<td>Max Avg</td>
<td>1.25</td>
<td></td>
<td></td>
<td>2.63</td>
<td>2.93</td>
<td>0.24</td>
<td>0.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Min Avg</td>
<td>0.19</td>
<td></td>
<td></td>
<td>1.29</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>HOT</th>
<th>NIGHT</th>
<th>WEEKEND</th>
<th>WEEK1</th>
<th>WEEK2</th>
<th>WEEK3</th>
<th>WEEK4</th>
<th>VISPROX</th>
<th>RIV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minnesota Ave</td>
<td>0.25</td>
<td>0.13</td>
<td>0.31</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>0.18</td>
<td>81.81</td>
<td>0.06</td>
</tr>
<tr>
<td>Montreal Ave</td>
<td>0.06</td>
<td>0.13</td>
<td>0.38</td>
<td>0.13</td>
<td>0.31</td>
<td>0.19</td>
<td>0.25</td>
<td>85.51</td>
<td>0.38</td>
</tr>
<tr>
<td>League Ave</td>
<td>0.10</td>
<td>0.12</td>
<td>0.37</td>
<td>0.32</td>
<td>0.24</td>
<td>0.24</td>
<td>0.219</td>
<td>70.39</td>
<td>0.19</td>
</tr>
<tr>
<td>Max Avg</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>115.43</td>
<td>0.41</td>
</tr>
<tr>
<td>Min Avg</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>48.38</td>
<td>0.00</td>
</tr>
</tbody>
</table>
### Data for 2002 Regular Season

<table>
<thead>
<tr>
<th>ATTENDANCE</th>
<th>POP</th>
<th>INC</th>
<th>PRICE</th>
<th>FANCOST</th>
<th>STADAGE</th>
<th>CAP</th>
<th>INDOORS</th>
<th>BLACKPOP</th>
<th>HISPOP</th>
<th>GB</th>
<th>DIVCHAMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minnesota Ave</td>
<td>23,758.93</td>
<td>2,888,847.00</td>
<td>26,347.00</td>
<td>11.76</td>
<td>121.13</td>
<td>20.00</td>
<td>48,678.00</td>
<td>1.00</td>
<td>157,695.00</td>
<td>98,337.00</td>
<td>25.05</td>
</tr>
<tr>
<td>Expos Ave</td>
<td>10,025.25</td>
<td>3,548,800.00</td>
<td>22,063.00</td>
<td>9.00</td>
<td>84.89</td>
<td>26.00</td>
<td>46,500.00</td>
<td>1.00</td>
<td>122,320.00</td>
<td>46,705.00</td>
<td>28.89</td>
</tr>
<tr>
<td>League Ave</td>
<td>28,169.32</td>
<td>5,449,103.13</td>
<td>24,228.48</td>
<td>18.26</td>
<td>145.18</td>
<td>23.05</td>
<td>46,737.48</td>
<td>0.20</td>
<td>734,151.05</td>
<td>1,078,306.05</td>
<td>28.15</td>
</tr>
<tr>
<td>Max Ave</td>
<td>43,739.98</td>
<td>16,373,845.00</td>
<td>30,789.00</td>
<td>39.68</td>
<td>228.73</td>
<td>90.00</td>
<td>66,307.00</td>
<td>1.00</td>
<td>2,580,852.00</td>
<td>6,598,488.00</td>
<td>63.08</td>
</tr>
<tr>
<td>Min Ave</td>
<td>10,025.25</td>
<td>1,070,052.00</td>
<td>20,454.00</td>
<td>9.00</td>
<td>84.89</td>
<td>1.00</td>
<td>33,871.00</td>
<td>0.00</td>
<td>119,509.00</td>
<td>17,100.00</td>
<td>8.11</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GBHOME</th>
<th>GBVIS</th>
<th>WINS LAST10</th>
<th>RUNS LAST10</th>
<th>HPITCHCAR</th>
<th>VISPITCHCAR</th>
<th>HPITCHCUR</th>
<th>VISPITCHCUR</th>
<th>LEADERS</th>
<th>HOMESTAR</th>
<th>VISSTAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minnesota Ave</td>
<td>0.31</td>
<td>9.87</td>
<td>5.70</td>
<td>46.20</td>
<td>2.98</td>
<td>11.65</td>
<td>1.65</td>
<td>1.08</td>
<td>1.35</td>
<td>6.00</td>
</tr>
<tr>
<td>Expos Ave</td>
<td>9.88</td>
<td>9.22</td>
<td>4.70</td>
<td>42.56</td>
<td>-2.43</td>
<td>9.27</td>
<td>0.37</td>
<td>1.05</td>
<td>1.93</td>
<td>3.00</td>
</tr>
<tr>
<td>League Ave</td>
<td>8.83</td>
<td>8.80</td>
<td>4.72</td>
<td>43.83</td>
<td>10.10</td>
<td>9.52</td>
<td>0.74</td>
<td>0.78</td>
<td>2.47</td>
<td>4.10</td>
</tr>
<tr>
<td>Max Ave</td>
<td>24.43</td>
<td>11.24</td>
<td>6.07</td>
<td>52.79</td>
<td>59.68</td>
<td>15.16</td>
<td>4.30</td>
<td>1.93</td>
<td>8.93</td>
<td>13.00</td>
</tr>
<tr>
<td>Min Ave</td>
<td>0.27</td>
<td>5.37</td>
<td>2.95</td>
<td>34.80</td>
<td>-6.65</td>
<td>3.12</td>
<td>-2.99</td>
<td>-0.26</td>
<td>0.94</td>
<td>2.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>WEEKEND</th>
<th>DAY</th>
<th>APRIL/MAY</th>
<th>COLD</th>
<th>HOT</th>
<th>RAIN</th>
<th>INTERLEAGUE</th>
<th>DBLHDR</th>
<th>DIVRIV</th>
<th>VISPROX</th>
<th>2TEAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minnesota Ave</td>
<td>0.48</td>
<td>0.26</td>
<td>0.32</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.11</td>
<td>0.00</td>
<td>4.71</td>
<td>1,015.32</td>
</tr>
<tr>
<td>Expos Ave</td>
<td>0.48</td>
<td>0.23</td>
<td>0.33</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.11</td>
<td>0.00</td>
<td>0.51</td>
<td>1,293.96</td>
</tr>
<tr>
<td>League Ave</td>
<td>0.48</td>
<td>0.32</td>
<td>0.33</td>
<td>0.06</td>
<td>0.07</td>
<td>0.02</td>
<td>0.10</td>
<td>0.02</td>
<td>0.43</td>
<td>1,276.19</td>
</tr>
<tr>
<td>Max Ave</td>
<td>0.51</td>
<td>0.74</td>
<td>0.36</td>
<td>0.20</td>
<td>0.32</td>
<td>0.14</td>
<td>0.11</td>
<td>0.10</td>
<td>0.52</td>
<td>2,242.08</td>
</tr>
<tr>
<td>Min Ave</td>
<td>0.48</td>
<td>0.14</td>
<td>0.29</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.07</td>
<td>0.00</td>
<td>0.18</td>
<td>823.94</td>
</tr>
</tbody>
</table>
Data for the 3 Year Panel Data Set

### TWINS

<table>
<thead>
<tr>
<th></th>
<th>ATTENDANCE</th>
<th>GBHOME</th>
<th>GBVIS</th>
<th>LEADERS</th>
<th>VISSTAR</th>
<th>WINSLAST10</th>
<th>RUNSLAST10</th>
<th>HPITCHCAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Ave</td>
<td>19,688.82</td>
<td>5.18</td>
<td>8.67</td>
<td>1.47</td>
<td>4.19</td>
<td>4.75</td>
<td>43.67</td>
<td>-0.72</td>
</tr>
<tr>
<td>00 Ave</td>
<td>13,082.90</td>
<td>13.54</td>
<td>5.30</td>
<td>1.48</td>
<td>4.41</td>
<td>3.93</td>
<td>41.26</td>
<td>-4.48</td>
</tr>
<tr>
<td>01 Ave</td>
<td>22,286.58</td>
<td>1.84</td>
<td>10.84</td>
<td>1.58</td>
<td>4.49</td>
<td>4.63</td>
<td>42.18</td>
<td>-0.63</td>
</tr>
<tr>
<td>02 Ave</td>
<td>23,758.93</td>
<td>0.31</td>
<td>9.67</td>
<td>1.36</td>
<td>3.68</td>
<td>5.70</td>
<td>48.20</td>
<td>2.96</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>VPLAYERCAR</th>
<th>HPITCHCUR</th>
<th>VPITCHCUR</th>
<th>WEEKEND</th>
<th>DAY</th>
<th>APRIL/MAY</th>
<th>INTERLAEGUE</th>
<th>VISPROX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Ave</td>
<td>8.95</td>
<td>1.12</td>
<td>0.77</td>
<td>0.49</td>
<td>0.25</td>
<td>0.34</td>
<td>0.13</td>
<td>1,041.56</td>
</tr>
<tr>
<td>00 Ave</td>
<td>8.41</td>
<td>-0.37</td>
<td>0.79</td>
<td>0.48</td>
<td>0.22</td>
<td>0.35</td>
<td>0.16</td>
<td>1,098.84</td>
</tr>
<tr>
<td>01 Ave</td>
<td>6.79</td>
<td>2.07</td>
<td>0.42</td>
<td>0.45</td>
<td>0.27</td>
<td>0.35</td>
<td>0.11</td>
<td>1,010.70</td>
</tr>
<tr>
<td>02 Ave</td>
<td>11.88</td>
<td>1.88</td>
<td>1.09</td>
<td>0.48</td>
<td>0.26</td>
<td>0.32</td>
<td>0.11</td>
<td>1,015.32</td>
</tr>
</tbody>
</table>

### EXPOS

<table>
<thead>
<tr>
<th></th>
<th>ATTENDANCE</th>
<th>GBHOME</th>
<th>GBVIS</th>
<th>LEADERS</th>
<th>VISSTAR</th>
<th>WINSLAST10</th>
<th>RUNSLAST10</th>
<th>HPITCHCAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Ave</td>
<td>9,734.81</td>
<td>10.78</td>
<td>7.13</td>
<td>1.99</td>
<td>4.10</td>
<td>4.21</td>
<td>39.96</td>
<td>-1.18</td>
</tr>
<tr>
<td>00 Ave</td>
<td>11,434.73</td>
<td>10.87</td>
<td>5.46</td>
<td>2.30</td>
<td>4.48</td>
<td>3.88</td>
<td>39.33</td>
<td>3.60</td>
</tr>
<tr>
<td>01 Ave</td>
<td>7,935.12</td>
<td>11.85</td>
<td>6.66</td>
<td>1.73</td>
<td>4.15</td>
<td>4.05</td>
<td>37.96</td>
<td>-4.74</td>
</tr>
<tr>
<td>02 Ave</td>
<td>9,835.84</td>
<td>9.76</td>
<td>9.11</td>
<td>1.94</td>
<td>3.66</td>
<td>4.70</td>
<td>42.58</td>
<td>-2.45</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>VPLAYERCAR</th>
<th>HPITCHCUR</th>
<th>VPITCHCUR</th>
<th>WEEKEND</th>
<th>DAY</th>
<th>APRIL/MAY</th>
<th>INTERLAEGUE</th>
<th>VISPROX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Ave</td>
<td>9.57</td>
<td>-0.14</td>
<td>0.63</td>
<td>0.48</td>
<td>0.22</td>
<td>0.33</td>
<td>0.11</td>
<td>1,320.74</td>
</tr>
<tr>
<td>00 Ave</td>
<td>10.05</td>
<td>0.23</td>
<td>0.80</td>
<td>0.48</td>
<td>0.22</td>
<td>0.33</td>
<td>0.11</td>
<td>1,390.33</td>
</tr>
<tr>
<td>01 Ave</td>
<td>9.12</td>
<td>-1.01</td>
<td>0.10</td>
<td>0.48</td>
<td>0.22</td>
<td>0.31</td>
<td>0.11</td>
<td>1,271.78</td>
</tr>
<tr>
<td>02 Ave</td>
<td>9.54</td>
<td>0.38</td>
<td>0.99</td>
<td>0.48</td>
<td>0.23</td>
<td>0.34</td>
<td>0.11</td>
<td>1,299.85</td>
</tr>
</tbody>
</table>
## Appendix 3
### Output for the Spring Training Model

<table>
<thead>
<tr>
<th></th>
<th>Specification 1</th>
<th>Specification 2</th>
<th>Specification 3</th>
<th>Specification 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-2.14501*</td>
<td>6.89354***</td>
<td>-5.3009***</td>
<td>-7.24680***</td>
</tr>
<tr>
<td>Location Variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INC</td>
<td>-0.00018</td>
<td>0.3179***</td>
<td>0.47185***</td>
<td></td>
</tr>
<tr>
<td>POP</td>
<td>0.14211**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SERVWORK</td>
<td>-0.05086</td>
<td>0.14160***</td>
<td>0.0993***</td>
<td>0.10441***</td>
</tr>
<tr>
<td>CAP</td>
<td>1.13936***</td>
<td>1.1594***</td>
<td>1.14064***</td>
<td></td>
</tr>
<tr>
<td>STADAGE</td>
<td>0.00296**</td>
<td>-0.0010</td>
<td>-0.00209**</td>
<td></td>
</tr>
<tr>
<td>AVEPRICE</td>
<td>-0.37910**</td>
<td>0.03915</td>
<td>-0.2632*</td>
<td>-0.01373</td>
</tr>
<tr>
<td>2TEAM</td>
<td>0.53625***</td>
<td>0.52992***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality of the Game Variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GBHOME</td>
<td>-0.00962***</td>
<td>-0.0088***</td>
<td>-0.00520***</td>
<td></td>
</tr>
<tr>
<td>GBVIS</td>
<td>-0.00081</td>
<td>-0.0019</td>
<td>-0.00313**</td>
<td></td>
</tr>
<tr>
<td>TOTGB</td>
<td>-0.00222**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIVCHAMPHM</td>
<td>0.13903***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIVCHMPVIS</td>
<td>0.09119***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTDIVCHAMP</td>
<td>0.05893***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HMSTAR</td>
<td>0.00176</td>
<td>0.0541***</td>
<td>0.04295***</td>
<td></td>
</tr>
<tr>
<td>VISSTAR</td>
<td>0.03456**</td>
<td>0.0707***</td>
<td>0.06223***</td>
<td></td>
</tr>
<tr>
<td>TOTSTAR</td>
<td>0.05073***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LEADER</td>
<td>-0.01809</td>
<td>0.01566</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HMSS</td>
<td>-0.1322***</td>
<td>-0.13786*</td>
<td>-0.1807***</td>
<td>-0.16723**</td>
</tr>
<tr>
<td>VISSS</td>
<td>0.05663</td>
<td>0.02102</td>
<td>0.0706</td>
<td>0.07541</td>
</tr>
<tr>
<td>Time and Weather Factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAIN</td>
<td>-0.03797</td>
<td>0.02384</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COLD</td>
<td>-0.20890***</td>
<td>-0.1883**</td>
<td>-0.2166**</td>
<td>-0.22538***</td>
</tr>
<tr>
<td>HOT</td>
<td>-0.05027</td>
<td>-0.00044</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NIGHT</td>
<td>0.06345</td>
<td>0.17047***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WEEKEND</td>
<td>0.14018***</td>
<td>0.15175***</td>
<td>0.1262***</td>
<td>0.12127***</td>
</tr>
<tr>
<td>WEEK1</td>
<td>-0.29460***</td>
<td>-0.28170**</td>
<td>-0.2954**</td>
<td>-0.31350**</td>
</tr>
<tr>
<td>WEEK2</td>
<td>-0.30434***</td>
<td>-0.31116***</td>
<td>-0.2998**</td>
<td>-0.31453***</td>
</tr>
<tr>
<td>WEEK3</td>
<td>-0.09960*</td>
<td>-0.10181*</td>
<td>-0.0868*</td>
<td>-0.10092**</td>
</tr>
<tr>
<td>WEEK4</td>
<td>-0.14455***</td>
<td>-0.14231**</td>
<td>-0.1391**</td>
<td>-0.15168**</td>
</tr>
<tr>
<td>Special Factors</td>
<td>Specification 1</td>
<td>Specification 2</td>
<td>Specification 3</td>
<td>Specification 4</td>
</tr>
<tr>
<td>----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>VISPROX</td>
<td>0.03968***</td>
<td>0.04318*</td>
<td>0.0314*</td>
<td>0.02869</td>
</tr>
<tr>
<td>DIVRIV</td>
<td>-0.00224</td>
<td>0.02360</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CONTRACT</td>
<td>-0.38962***</td>
<td>-0.47269***</td>
<td>-0.2333***</td>
<td></td>
</tr>
<tr>
<td>TWINS</td>
<td></td>
<td></td>
<td>0.06190</td>
<td>-0.70483***</td>
</tr>
<tr>
<td>EXPOS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Adj r-squared | 0.63            | 0.49            | 0.56            | 0.61            |
| F-stat         | 18.33           | 14.15           | 20.20           | 22.90           |

VIF<5

SERVWORK

DIVCHAMPHM

***p value less than .02

**p value less than .1

*p value less than .2

Output for Regular Season Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Specification 1</th>
<th>Specification 2</th>
<th>Specification 3</th>
<th>Specification 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-3.31886***</td>
<td>0.82587</td>
<td>-5.50942***</td>
<td>-4.01083***</td>
</tr>
<tr>
<td>Location Variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POP</td>
<td>-0.00740</td>
<td>0.01743**</td>
<td>0.14963***</td>
<td>0.18277***</td>
</tr>
<tr>
<td>INC</td>
<td>0.19667***</td>
<td>0.49780***</td>
<td>0.29620***</td>
<td>0.20948***</td>
</tr>
<tr>
<td>PRICE</td>
<td>0.47308***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FANCOST</td>
<td>0.60532***</td>
<td>0.78803***</td>
<td>1.17840***</td>
<td>1.08458***</td>
</tr>
<tr>
<td>STADAGE</td>
<td>-0.00296***</td>
<td>-0.00206***</td>
<td>-0.00163***</td>
<td></td>
</tr>
<tr>
<td>CAP</td>
<td>0.84104***</td>
<td>0.64079***</td>
<td>0.59779***</td>
<td></td>
</tr>
<tr>
<td>INDOORS</td>
<td>-0.01435</td>
<td>-0.06711***</td>
<td>-0.07864***</td>
<td></td>
</tr>
<tr>
<td>BLACKPOP</td>
<td>-0.11844***</td>
<td>-0.15239***</td>
<td>-0.17068***</td>
<td></td>
</tr>
<tr>
<td>HISPPPOP</td>
<td>0.01690***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2TEAM</td>
<td>0.21443***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality of the Game Variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GBLAST</td>
<td>0.00037</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIVCHAMP</td>
<td>0.00558</td>
<td>0.03666***</td>
<td>0.04617***</td>
<td></td>
</tr>
<tr>
<td>GBHOME</td>
<td>-0.01546***</td>
<td>-0.01573***</td>
<td>-0.01661***</td>
<td>-0.01477***</td>
</tr>
<tr>
<td>GBVIS</td>
<td>-0.00181***</td>
<td>-0.00340***</td>
<td>-0.00222***</td>
<td>-0.00223***</td>
</tr>
<tr>
<td>WINSLAST10</td>
<td>0.00759</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RUNSLAST10</td>
<td>-0.00222***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPITCHCAR</td>
<td>0.00110***</td>
<td></td>
<td>0.00096***</td>
<td>0.00100***</td>
</tr>
<tr>
<td>VPIITCHCAR</td>
<td>-0.00002</td>
<td></td>
<td>-0.00003</td>
<td>-0.00012</td>
</tr>
<tr>
<td>HPITCHCUR</td>
<td>0.00149</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VPIITCHCUR</td>
<td>0.00424**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LEADERS</td>
<td>0.02936***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOMESTAR</td>
<td>0.01459***</td>
<td>0.03563***</td>
<td>0.01019***</td>
<td>0.00718***</td>
</tr>
<tr>
<td>VISSTAR</td>
<td>0.02121***</td>
<td>0.02252***</td>
<td>0.01562***</td>
<td>0.01382***</td>
</tr>
<tr>
<td>Time and Weather Factors</td>
<td>Specification 1</td>
<td>Specification 2</td>
<td>Specification 3</td>
<td>Specification 4</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>WEEKEND</td>
<td>0.22730***</td>
<td>0.23543***</td>
<td>0.22999***</td>
<td>0.22969***</td>
</tr>
<tr>
<td>DAY</td>
<td>0.07967***</td>
<td>0.07260***</td>
<td>0.08725***</td>
<td>0.08780***</td>
</tr>
<tr>
<td>APRIL/MAY</td>
<td>-0.20463***</td>
<td>-0.20397***</td>
<td>-0.20714***</td>
<td>-0.19722***</td>
</tr>
<tr>
<td>COLD</td>
<td>-0.08702***</td>
<td>-0.11377***</td>
<td>-0.07946***</td>
<td>-0.08818***</td>
</tr>
<tr>
<td>HOT</td>
<td>0.11808***</td>
<td>0.12275***</td>
<td>0.12317***</td>
<td>0.13282***</td>
</tr>
<tr>
<td>RAIN</td>
<td>-0.13206***</td>
<td>-0.15336***</td>
<td>-0.12823***</td>
<td>-0.12785***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Special Factors</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>INTERLEAGUE</td>
<td>-0.08115***</td>
<td>-0.03504*</td>
<td></td>
<td>-0.02781</td>
</tr>
<tr>
<td>DBLHDR</td>
<td>0.05427</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIVRIV</td>
<td>-0.10551***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VISPROX</td>
<td>-0.07280***</td>
<td>-0.04787***</td>
<td>-0.04738***</td>
<td></td>
</tr>
<tr>
<td>CONTRACT</td>
<td>-0.24342***</td>
<td>-0.41868***</td>
<td>-0.35644***</td>
<td>-0.17881***</td>
</tr>
<tr>
<td>TWINS</td>
<td></td>
<td></td>
<td></td>
<td>-0.17881***</td>
</tr>
<tr>
<td>EXPOS</td>
<td></td>
<td>-0.58006***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| R-squared                        | 0.599           | 0.517           | 0.573           | 0.581           |
| F-stat                           | 106.66          | 185.26          | 140.99          | 139.67          |

**VIF > 5**

**POP**

***p value less than .02

**PRICE**

**p value less than .1

**FANCOST**

*p value less than .2

**HISPPOP**

**p value less than .2
### Output for the 3 Year Panel Data Set

#### TWINS

<table>
<thead>
<tr>
<th>Variable</th>
<th>Specification 1</th>
<th>Specification 2</th>
<th>Specification 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>9.88210***</td>
<td>10.77221***</td>
<td>10.64706***</td>
</tr>
<tr>
<td>Quality of the Game Variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GBHOME</td>
<td>-0.02654***</td>
<td>-0.02750***</td>
<td>-0.02283***</td>
</tr>
<tr>
<td>GBVIS</td>
<td>-0.01146***</td>
<td>-0.01187***</td>
<td>-0.01106***</td>
</tr>
<tr>
<td>LEADERS</td>
<td>0.04831***</td>
<td>0.04646***</td>
<td>0.03181***</td>
</tr>
<tr>
<td>VISSTAR</td>
<td></td>
<td></td>
<td>0.04286***</td>
</tr>
<tr>
<td>WNSTLAST10</td>
<td></td>
<td></td>
<td>0.02857***</td>
</tr>
<tr>
<td>RUNSLAST10</td>
<td></td>
<td></td>
<td>-0.00306***</td>
</tr>
<tr>
<td>HPITCHCAR</td>
<td></td>
<td>0.0006</td>
<td>-0.00612***</td>
</tr>
<tr>
<td>VPITCHCAR</td>
<td></td>
<td>0.00063</td>
<td>-0.00061</td>
</tr>
<tr>
<td>HPITCHCUR</td>
<td></td>
<td></td>
<td>0.02294**</td>
</tr>
<tr>
<td>VPITCHCUR</td>
<td></td>
<td></td>
<td>0.00654</td>
</tr>
<tr>
<td>Time and Weather Factors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WEEKEND</td>
<td>0.39398***</td>
<td>0.37232***</td>
<td>0.36904***</td>
</tr>
<tr>
<td>DAY</td>
<td>0.03407</td>
<td>0.02662</td>
<td>0.03311</td>
</tr>
<tr>
<td>APRIL/MAY</td>
<td>-0.03546***</td>
<td>-0.30311***</td>
<td>-0.23992***</td>
</tr>
<tr>
<td>Special Factors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INTERLEAGUE</td>
<td></td>
<td>0.16710**</td>
<td>0.26364***</td>
</tr>
<tr>
<td>VISPROX</td>
<td></td>
<td>-0.12091***</td>
<td>-0.12469***</td>
</tr>
<tr>
<td>2000</td>
<td>-0.34667***</td>
<td>-0.32789***</td>
<td>-0.44880***</td>
</tr>
<tr>
<td>2001</td>
<td>-0.93183***</td>
<td>-0.89865***</td>
<td>-0.98514***</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.67</td>
<td>0.69</td>
<td>0.71</td>
</tr>
<tr>
<td>F-stat</td>
<td>63.9</td>
<td>46.8</td>
<td>36</td>
</tr>
</tbody>
</table>

*** p value less than .02
** p value less than .1
*p value less than .2
## EXPOS

<table>
<thead>
<tr>
<th>Quality of the Game Variables</th>
<th>9.15845***</th>
<th>8.99183***</th>
<th>8.91100***</th>
</tr>
</thead>
<tbody>
<tr>
<td>GBHOME</td>
<td>-0.01734***</td>
<td>-0.01075***</td>
<td>-0.01182***</td>
</tr>
<tr>
<td>GBVIS</td>
<td>-0.01237***</td>
<td>-0.01289***</td>
<td>-0.00642***</td>
</tr>
<tr>
<td>LEADERS</td>
<td>0.01847</td>
<td>0.01544</td>
<td>-0.00299</td>
</tr>
<tr>
<td>VISSTAR</td>
<td></td>
<td></td>
<td>0.08354***</td>
</tr>
<tr>
<td>WINSLAST10</td>
<td></td>
<td></td>
<td>0.00061</td>
</tr>
<tr>
<td>RUNSLAST10</td>
<td></td>
<td></td>
<td>-0.00196</td>
</tr>
<tr>
<td>HPITCHCAR</td>
<td></td>
<td>0.00128</td>
<td>0.00009</td>
</tr>
<tr>
<td>VPITCHCAR</td>
<td></td>
<td>0.00065</td>
<td>-0.00139</td>
</tr>
<tr>
<td>HPITCHCUR</td>
<td></td>
<td></td>
<td>0.02849**</td>
</tr>
<tr>
<td>VPITCHCUR</td>
<td></td>
<td></td>
<td>0.00739</td>
</tr>
<tr>
<td>Time and Weather Factors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WEEKEND</td>
<td>0.35691***</td>
<td>0.34791***</td>
<td>0.37151***</td>
</tr>
<tr>
<td>DAY</td>
<td>0.15293**</td>
<td>0.15311**</td>
<td>0.17843***</td>
</tr>
<tr>
<td>APRIL/MAY</td>
<td>-0.26164***</td>
<td>-0.14172**</td>
<td>-0.08010</td>
</tr>
<tr>
<td>Special Factors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INTERLEAGUE</td>
<td></td>
<td>0.21615**</td>
<td>0.19800**</td>
</tr>
<tr>
<td>VISPROX</td>
<td></td>
<td>0.0072</td>
<td>-0.01759</td>
</tr>
<tr>
<td>2000</td>
<td>0.19749***</td>
<td>0.17883**</td>
<td>0.12655**</td>
</tr>
<tr>
<td>2001</td>
<td>-0.19655***</td>
<td>-0.20568**</td>
<td>-0.21465***</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.3</td>
<td>0.3</td>
<td>0.35</td>
</tr>
<tr>
<td>F-stat</td>
<td>14.2</td>
<td>9.8</td>
<td>8.8</td>
</tr>
</tbody>
</table>

*** p value less than .02
** p value less than .1
* p value less than .2
Bibliography


Data Sources-
www.espn.com
www.census.gov
www.weatherunderground.com
www.usatoday.com
www.yahoo.com
users.pullman.com/rodfort/PHSportsEcon/Students/StudentContent.html