Determining if Factors Affecting NHL Attendance Are Equal Across Geographical Regions

John Provenzano
Faculty: Jenchi Cheng
Department of Economics

Abstract. This paper estimates a nightly demand model for National Hockey League attendance for five teams from different divisions. The magnitudes of the coefficients are compared for each team to determine if factors that influence attendance affect each team the same. All variables except games played in October were found to be statistically different between at least two teams. Games played between division opponents and previous season success had the largest regional variation while “calendar” variables (weekend and games played in March and April) have the lowest regional differences. The analysis shows that a general attendance model of the NHL is inaccurate and can cause misleading interpretations for purposes of marketing and profit maximization.

1. Introduction

Professional sports attendance has been a widely studied issue. The majority of the literature for National Hockey League (NHL) attendance estimates a general model for average nightly attendance demand for the entire league. The only distinction between the teams is US and Canadian. If the factors that affect demand vary from region to region, a general model for the entire league is inaccurate. For this analysis, only US teams were considered. The hypothesis of the experiment is that there are regional differences in factors that affect the demand for NHL attendance. In this research, whether or not the coefficients of the demand models are statistically significant is not the goal. The purpose of this paper is to determine if the magnitude of the coefficients of the individual demand models are statistically different.

2. Results

The NHL divides their divisions by geographic region. One team from five of the six divisions was taken for the sample. Ideally, a model with every team would make the analysis complete, however, if differences are found using only one team from each division, that is enough to conclude there are differences in how demand is affected in different regions. No team was selected from the Northwest division. The only two American teams in that division are the Minnesota Wild and the Colorado Avalanche. The Wild have sold out every home game in their existence and the Avalanche sold out every game from 1997-2007. The same model for teams that always sell out cannot be applied to teams that have fluctuation in attendance.

The teams chosen were the Anaheim Ducks, Chicago Blackhawks, Boston Bruins, Carolina Hurricanes and New Jersey Devils. Boston and Chicago were chosen because they are two of the original six NHL teams. Carolina and Anaheim were chosen because they are non-traditional hockey cities. New Jersey was chosen because they are in a large market, have been very successful and do not sell out regularly.

The Zellner Method for Seemingly Unrelated Regressions (SUR) was used for the model. The teams cannot be treated independently despite being in different cities. A factor that affects demand for the NHL as a whole will affect demand in both Boston and Anaheim. The SUR method takes this into account and pools the errors for better estimation. A Wald statistic was used to determine if the magnitude of each coefficient was the same in the two cities being compared. The example statistic below follows a Chi-squared distribution with one degree of freedom.

\[ \text{Wald} = (\beta_{\text{Anaheim}} - \beta_{\text{Boston}}) / \text{Var}(\beta_{\text{Anaheim}} - \beta_{\text{Boston}}) \sim \chi^2(1) \]

The sample period was for home games of each team between the 1997-98 season and 2008-09 season. Four of the teams had 451 observations. Chicago played one home game away from their home arena and thus had only 450 observations. Each team will be compared against each other team for a total of ten tests.

 Arenas in the NHL have different capacities. Therefore, utilized capacity per game was used as the dependent variable rather than absolute attendance. The data used was reported paid attendance. For example, if Anaheim had
paid attendance of 14,300 out of the 17,100 capacity, the value of the dependent variable in this case would be 0.836. The independent variables in the model were chosen because they are generally significant in the literature models. The variables are: Division game, Goals-for Per Game, Goals-against Per Game, October Game, March or April Game, Weekend Game, Opening Night, Point Total in the Previous Season, Points-per-Game in Current Season, Playoffs Reached in Previous Season and Round Two of the Playoffs Reached in the Previous Season. In order to get the best possible comparison, the same model was used for each team regardless of whether or not the variable significantly affects attendance in a specific model. There were two exceptions. New Jersey reached the playoffs every year and Chicago never made it to round two of the playoffs in any year of the sample. Attempting to run a regression with these models would result in a singular matrix and the analysis could not be done. The model is below where $\beta = (\beta_1, \beta_2, \ldots, \beta_n)$ and $X = (\text{Division}, \text{Goals-per-game}, \text{October game}, \ldots, \text{Opening Night})$.

\[
\text{Capacity}_i = \beta_0 + \beta_i X'_i + e_i \tag{2}
\]

The results of the SUR and Wald tests confirm the hypothesis that significant differences exist between teams with respect to all factors tested except for October games. This is surprising because October is when baseball season is ending. Baseball will likely take attendance away from hockey. Anaheim and Boston have had successful baseball teams, whereas Carolina does not have a baseball team. A city with a baseball team should have hockey attendance affected differently during baseball season than a city without a team but that was not the case. Success in the previous season and whether or not a team was playing a divisional game were significantly different for six out of the ten tests. Goals-against per game was different for five tests. Weekend, March and April games and opening night games were different for only two tests. Current season points per game and goals-for per game were different in three tests. Due to the problem of singular matrices, reaching the playoffs in the previous year and reaching round two in the previous year’s playoffs only had six tests instead of ten. Making it to round two of the playoffs was different for three tests and reaching the playoffs in the previous season differed for only one test.

Finding significant differences in all but one of the twelve factors shows a general attendance model that treats teams as homogenous firms is inaccurate. This has two relevant implications. The first is in the individual team’s front office decision making. If a general manager wants to acquire a high priced player to boost a team’s goal scoring, he has to decide if the expected increase in attendance will be more than the cost of the player. If he uses what another team experienced as a guide, the results may be different than what is expected.

The second implication is for the NHL’s marketing efforts. If a general model of attendance is inaccurate, general marketing techniques for the entire league will also be inaccurate. Once regional differences of these factors are established, better marketing plans can be developed to expand national interest in the game.

3. Conclusion

This research has shown that the traditional general attendance models for the NHL are inaccurate because the magnitudes of several factors that affect nightly attendance vary from region to region. There are relevant profit maximizing implications for individual teams knowing that their attendance demand is not necessarily the same as the average NHL team. The NHL would benefit from this insight as a way to market the game in a unique way to different geographical regions.

4. Acknowledgements

I would like to thank Dr. Jenchi Cheng, Dr. Philip Hersch, Dr. Jodi Pelkowski and Brenda Lehman for their assistance and suggestions for this research.

References
