5-1-2019

The Gretzky Externality: An Analysis of the Superstar Effect and Age Curves in the National Hockey League

Jed P. Buchholz
Macalester College

The Gretzky Externality:

An Analysis of the Superstar Effect and Age Curves in the National Hockey League

Jed Patrick Buchholz

Honors Thesis

Macalester College

Advisor: Prof. J. Peter Ferderer

Economics

May 1st, 2019

Introduction

In his seminal paper, Rosen (1981) explained the phenomenon of "superstars"—the "small numbers of people who earn enormous amounts of money and dominate the activities in which they engage" (p. 845). Rosen cites comedians, musicians, textbook writers, and so on as examples and identifies two elements of markets where superstar effects emerge: (1) imperfect substitutability between the sellers where, for example, seeing several mediocre performers does not sum to the performance of one outstanding performer, and (2) joint consumption technology which allows the superstar to expend the same effort whether 10 or 10,000 people are in attendance. The presence of these elements causes the net revenue earned by the superstar to be convex in talent so that small differences in talent are magnified into large earning differences at the high-end of the talent continuum.

Adler (1985) offers a competing explanation of the superstar effect where differences in talent play a minor role. Adler argues that stardom occurs because the utility derived from consumption of the service depends on knowledge shared between consumers and, as a consequence, they are more inclined to patronize stars that everyone else is familiar with. For example, a person will purchase the music of a specific artist, not necessarily because they are more talented than other performers, but because many others consume this performer's music and widespread prior knowledge about them makes discussion (knowledge sharing) easier and is a source of utility. This perpetuates a virtuous cycle, or a positive feedback loop, where small (or nonexistent) initial differences in ability can, over time, lead to large differences in demand for particular suppliers as demand begets more demand.

Since the classic work of Rosen and Adler, many empirical analyses have attempted to determine whether superstar effects are present in markets and, if so, which model does a better job of explaining them. One industry that has been explored by several authors is professional sports with the most significant body of research conducted on European soccer and American basketball. Given modern communication technologies (television, the internet, etc.) joint consumption and knowledge sharing are possible in professional sports. Anecdotal evidence suggests that certain players receive a disproportionate amount media attention. Moreover, professional sports have numerous metrics that enable researchers to quantify differences in talent. For these reasons, professional sports labor markets are a natural setting to explore whether top players are “Rosen” superstars or “Adler” superstars.

According to research in soccer and basketball, the superstar effect is observed when a player generates attendance and revenue above and beyond that which they contribute through their impact on team performance. Using this definition, Hausman and Leonard (1997) and Berri and Schmidt (2004) find a quantifiable superstar effect in the NBA consistent with the predictions of the Adler and Rosen models when analyzing player presence on gate revenue. Lucifora and Simmon (2003) similarly find that both Rosen and Adler models of superstardom explain game attendance patterns in Italian professional soccer. This is expanded to German leagues by Brandes, Franck, and Nüesch (2008) who find that particular players have an outsized impact on attendance even after controlling for their influence on team performance. These superstars' effects are consistent with both the Adler or Rosen models. In addition, they provide evidence for both local and national superstars which depends on the player's performance. These will be further described later on.

These studies beg the question: are superstar effects prevalent in other sports? In this thesis I examine the superstar effect in professional hockey. There have been no studies conducted on the presence of superstars in hockey at any level making this analysis novel. As another team sport it should possess similar characteristics to previously studied sports such as basketball and soccer. However, there are distinct differences between hockey and these sports that lead us to believe that superstar effects may be different across sports. These differences include the structure of the game (e.g., the role of the individual versus the team, the presence of violence in hockey, etc.), demographics of fans (e.g., northern locations, a significant influence of Canadian culture), media coverage, etc. The objective of this paper is to test for the superstar effect in professional hockey and compare it to those observed in professional basketball and soccer.

While analyses of labor markets in the NHL have not examined the superstar effect, considerable effort has been made to measure player performance.[1] In general, these studies incorporate performance variables (e.g.,

[1] Jones, Nadeau, and Walsh's (1999) analysis of discrimination and salaries in the NHL and Idson and Kahane's (2000) analysis of player compensation use an independent variable which summed awards and all-star selections for players.

points per game, goals against average, etc.) as explanatory variables in linear models for players' salaries. I use this data to identify potential superstars in professional hockey and then, like previous studies in basketball and soccer, measure the impact of a star player on game attendance after controlling for other determinants of demand, including the influence of the presumed superstar on team performance.

This paper analyzes the presence of a superstar effect in the NHL using game-level data from all NHL seasons between the 1979-1980 and 2017-2018 seasons. In particular, the focus is on the impact a superstar has on game attendance. Is the impact commensurate with quantifiable measures of ability and contribution to team success, consistent with standard neoclassical theory of factor prices, or is there a large component of demand that is not explained by performance? As stated above, previous analyses of superstar effects in professional sports have not investigated professional hockey. These studies have additionally been conducted over a shorter time-span and have only considered career averages. In this analysis, I utilize game attendance to measure demand and various characteristics of each game to control for factors that might affect demand. This is used in conjunction with performance data for different players across different NHL eras to measure ability. This is the largest longitudinal data set ever assembled for this type of analysis and the first of its kind for hockey. In addition, given my dataset it is possible to analyze the effects of "age curves" in hockey. That is, the magnitude of the superstar effect throughout a player's career is examined by estimating their impact on attendance on a season-by-season basis to explore whether the demand for a player's services becomes delinked from their productivity. This type of analysis has never been conducted for an investigation of the superstar effect and allows us to test between competing models If Rosen's model is correct, we should observe that the superstar effect rises and falls over the superstar's career as their ability rises, reaches a peak and then declines as they approach retirement. In contrast, the Adler model implies that the superstar effect will not diminish with age as the utility consumers derive from discussion of the player persists even as talent diminishes.

I provide four main findings. First, there is a superstar effect in hockey identified by a statistically significant influence on game attendance when a superstar is present. For the eight players we consider, only two--Wayne Gretzky and Alexander Ovechkin–have a superstar effect that significantly effects both home and away attendance. Second, the observed superstar effect in hockey is smaller in magnitude than what has been found in professional basketball. Third, there are observed instances of local superstars where there is a significant effect on attendance for home games but no such effect on away games. Finally, using new methods to analyze the determinant of the superstar effect over a player's career (age curves) , the Rosen and Adler models both appear to find support depending on the player examined. That is, Wayne Gretzky's capacity to generate additional fans becomes delinked from performance as his career advances, consistent with the Rosen model. The opposite is true for Alexander Ovechkin who sees close correlation between his effect on attendance and his annual performance. Overall, my results imply that superstar effects are influenced by confounding factors—the unique nature of the sport, type of players (goal scorer vs. all-around player), media coverage--but are also more prevalent than previously thought.

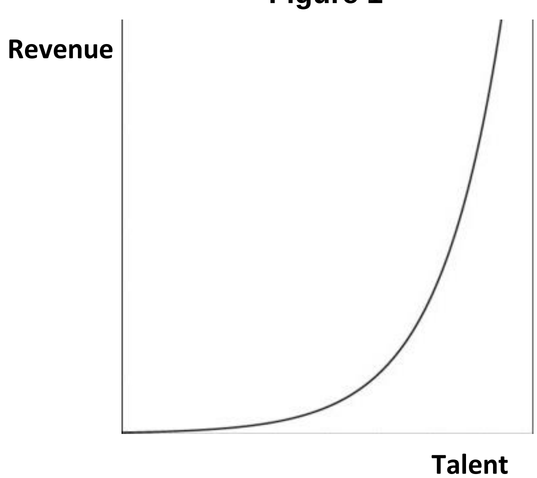
Theoretical Model

Rosen (1981) explains the large differences in earnings by a small number of individuals in a market. He discusses the relationship between suppliers and consumers using talent or quality as the sole determinant of "box office" appeal. The net revenue function of the seller is convex with respect to talent, see Figure 1, which indicates that small differences in ability become more significant as talent increase.

The cause of this convexity, according to Rosen, is the imperfect substitution among sellers. Greater difference in substitution marks more disproportionate earnings. The demand for greater talent increases disproportionately more than the discrepancy in talent. Rosen provides an example of a surgeon who has a 10 percent lower mortality rate than a competing surgeon. Rosen argues that the more talented surgeon will receive a premium for their service greater than 10 percent because buyers care deeply about small improvements in probability of success. Consider another example, specifically in sports, of renowned soccer talent Lionel Messi. Should Messi make a single incredible move to retain possession of the ball, it will provide more utility and is valued more than five mediocre passes by a lesser talent. Rosen goes on to explain that as more sellers with different levels of talent enter the market, the difference between variation in talent and variation in earnings increases, meaning that a small number of sellers with marginally greater talent may dominate their market with respect to revenue and demand for their services. This is best illustrated in Figure 1 which shows the revenue or wage (on the Y-axis) of a seller plotted against their talent (on the X-axis).[2] The convexity of the curve leads to there being a few sellers at the higher end of talent where small differences in talent at this range will have a considerably larger impact on revenue.

Rosen's model contrasts to Adler (1985) who describes markets as requiring knowledge, meaning that fans of these sellers may communicate with others to create further demand via spillover effects. For example, these effects may create significant variance in the size of fan following of two players even if their talent is similar. Because demand is driven by knowledge, individuals may create a snowball effect through shared knowledge and superstars are created without preferences based on ability. Adler's model describes a market with varying or equal talent and a demand function primarily driven by knowledge; therefore, popularity is the driving distinction among superstars which Adler argues is largely determined by luck. An example of this is the music industry. Consider two acts of equal talent who are appearing in their first public

Figure 1



[2] See Rosen, S. (1981). The economics of superstars. The American economic review, 71(5), 845-858 for a more complete description

performances. If band A plays before a crowd that includes a well-known music writer, a positive review of the band will provide widespread knowledge greater than the information shared exclusively by patrons of the band B. Because of this, band A will garner more attention and thus accrue more revenue than band B despite equal talent by way of these information effects. Because of the contrasting hypotheses behind the superstar effect—popularity vs. talent—multiple measures of superstar status will be used that encompass each theory.

In their original theoretical models, neither Rosen nor Adler consider how the superstar effect evolves over a supplier's age. However, MacDonald (1988) utilizes Rosen's model to describe "rising stars" and creates two classes of stars, young and old. MacDonald describes survivorship bias; how younger, less successful performers will leave the market leaving only older, more successful performers and promising young talent. The older performers will play for larger, higher revenue crowds while younger stars will perform for the opposite. This survival component is a cornerstone of MacDonald's model. Survival is based on a performer's talent and only those who are most successful will survive in the market. Because of this, these performers will face a relatively non-competitive market and thus generate revenue at a rate much greater than their relative performance. MacDonald seems to describe how older talent, benefiting from information effects, will be more successful but emphasizes Rosen's findings that talent is the implicit cause of superstar effects.

These competing explanations are a significant part of the sports literature with previous analyses seeking to explain which model explains the presence of the superstar effect in their respective leagues relative to their chosen quantitative measure of demand. While professional sports is a relatively small market--the cardboard box industry generates three times the revenue of the entire MLB (Fort, 2000)--sports is a well reported and well followed market. This is an important characteristic of the market that may help cause a superstar effect. Professional sports is unlike other products as viewing sports is a non-rival good. Performers are able to sell their product to a much larger market and in doing so generate much higher revenues. In contrast, if a company sells a cardboard box to an individual it cannot sell it to another. The cardboard box is a rival good.

With respect to sports economics as a whole, age curves of players are a topic of great importance from both a performance and economic perspective. An age curve is a tool that models a player's performance by some metric as a function of age across their career. These age curves exist in all sports as younger players have greater physical potential; however, the effect differs heavily from sport to sport. Rob Vollman (2016) summarizes the intricacies of estimating such curves in professional hockey. A significant issue in describing the expected value of a player at a given age is survivorship bias in which aging players may be replaced by younger players leaving only high-caliber players after a peak age. Hockey statisticians have designed various methods to accommodate for this in order to accurately measure age curves and have even analyzed player data to establish the peak age of offensive performance for hockey players to be between age 27 and 28 with similar levels of offensive performance between ages 24 and 32 (Brander, Egan, Yeung, 2014). Given these estimates, even superstar players should see diminished performance after a long tenure in the NHL and, per the Rosen model, see a regression in the magnitude of their impact on demand. If this is not observed and the demand remains relatively constant regardless of diminishing talent, then popularity and Adler's likely explains the effect.

Previous Research

The superstar effect influences the demand side of the market. Borland and Macdonald (2003) provide a thorough analysis of what could possibly effect the demand for viewing a sport giving numerous control variables to consider. This will be explained more closely when discussing data, yet one important finding from the meta-analysis is that demand is based on five main factors: consumer preference, economic factors, quality of viewing, contest characteristics, and stadium capacity with the superstar effect encompassed by contest characteristics. A significant point for this expansion of the superstar effect literature is Hansen and Gauthier's (1989) meta-analysis across several professional North American leagues which found that factors of demand varied by sport and league. Numerous determinants of attendance were considered for several North American sports leagues. Their findings illustrated a number of differences among professional leagues determinants of demand including scheduling, team roster quality, and ticket price.

With this, having no previous similar analysis to consider, there may be new, important variables specific to hockey to consider. Paul (2003) analyzed several specific factors affecting the demand for NHL games measured by attendance with findings including positive effects for violence--fighting being a unique component of the sport not present in the other major American professional league sports--and rivalries. The superstar effect is mentioned, but never investigated.

Empirical analyses of the superstar effect, as mentioned, have primarily dealt with soccer and basketball. These two sports are similar to hockey making their findings significant to this study for different reasons. Basketball is played on a similar sized surface with similar numbers of players, whereas soccer utilizes a goaltending system and has more dedicated positions to offense and defense. Previous relevant empirical research begins with Hausman and Leonard's (1997) analysis of the effect of 25 players receiving the most all-star votes on television ratings and revenues in the NBA. They found that there was a clear effect of star players' participation' that increased both ratings and revenues. They also found that these effects were present at both home and away games, especially by marquee players such as Michael Jordan and Larry Bird.

Berri and Schmidt (2004) further the NBA analysis by examining star players effects on gate revenues. They use fan all-star votes to identify star players as well as dummy variables for the presence of star players at games similar to Hausman and Leonard (1997). Berri and Schmidt (2004) find that while superstars significantly affect gate revenue, the ability of a team to generate wins is more important. The system of fan all-star votes as a way to encapsulate popularity is frequently used; however, it may be an imperfect measure of popularity as fan voting may be driven by popularity or talent or both in a given season. One would assume that if a player performs at a high level, they will receive acclaim. Because of this, popularity and talent become interwoven when considering all-star voting. Berri and Schmidt only utilize a model using fan all-star votes and not a specific performance based metric that would represent player talent representing the Rosen model. This ignores the propositions of Rosen to include talent despite Berri and Schmidt's claim that on court performance is the most significant determinant of revenue.

Empirical research regarding the superstar effect in sports is expanded in a subsequent paper by Berri and Schmidt (2006), which analyzes the effect of star players on road NBA games. Their findings suggest that a superstar externality is present in the NBA and affects road attendance when a star player is visiting providing additional revenue to opposing teams. They find that the average impact of the top 25 all-star recipients was 4,353 additional fans at away games compared to 9,846 fans generated by the player's win production calculated by the player's win shares statistic multiplied by the coefficient for team wins when estimating game attendance. This shows that talent is a significant determinant of attendance. The effect on both home and away attendance was analyzed by Jane (2016) who considered game-level data for the 2010-2011 season and 2011-2012 season of the NBA to find that individual superstars did contribute to both home and away attendance and such effects were the result of popularity, measured by dummy variables for the presence of all-star players, rather than talent, measured directly by salary and performance metrics.

This research was expanded to Italian professional soccer by Lucifora and Simmon's (2003). Their analysis excluded goalkeepers as they are measured and observed in different ways from other types of players. Their use of dichotomous variables for the presence of high scoring, top-tier players finds a clear presence of a superstar effect. The analysis uses a performance based metric to select players, but cites players' popularity as a factor for attendance. Thus, it is unclear which model works best to explain the effect in the Italian league.

Brandes, Franck, and Nüesch (2008) conduct a further investigation of German professional leagues to determine how national and local stars affect attendance.[3] They find that national stars exhibit superstar externalities, represented by greater attendance and team revenue, at both home and away games while local stars only effect home attendance. They additionally consider the Rosen and Adler models, considering the impact of performance metrics (goals and assists) and popularity variables (citations in German press) and find that popularity at generated revenue for local stars while performance is the driving factor for national superstars. Franck and Nüesch (2012) conduct a second analysis of German soccer to determine the cause of the superstar effect in the league. They compare performance metrics and media mentions to both

[3] Brandes, Franck, and Nüesch explain "superstars as "players whose market values are in the top 2% quantile of the league's distribution of market values and a 'local hero' as the most valued player of a particular team that has no superstars." (267)

market value using an OLS model. They find that each model, Rosen and Adler, is useful in explaining various star's effects as player performance and media popularity in this quantile has statistically significant effects on market value.

The most similar paper to mine is Johnson and Humphreys' (2017) analysis of the superstar effect in the NBA. Their paper utilizes 33 seasons of game-level data and an analysis of five superstars using various controls for attendance. Using dichotomous variables for the presence of star players at both home and away games as well as control variables, Johnson and Humphreys find a quantifiable superstar effect for each player on attendance, on average, over the course of their careers. The paper investigates the Rosen and Adler debate by comparing the significance of advanced performance metrics and fan all-star votes to find evidence of both effects. Their model to assess the Rosen and Adler models simply replaces the indicator variable for a player being present at a game with values for their VORP (victories over replacement player) and number of fan all-star votes received in a given season. Their results suggest that Larry Bird derives his superstar effect from performance while Michael Jordan and LeBron James are driven by their popularity when analyzing all-star voting. Such fan based popularity metrics did not exist in professional hockey until very recently requiring a different measurement approach. The time period of Johnson and Humphreys' analysis is similar in length to that used in my paper and as such permits the same potential for novel contribution; however, rather than estimating only career averages, season-by-season analysis will be used to determine the longevity of a superstar's observed externality relative to popularity or talent.

Empirical Model

The model in this analysis is consistent with previous papers (Berri and Schmidt 2006, Lewis and Yoon 2016, Jane 2016, Johnson and Humphreys 2017). A reduced form empirical model is used to explain differences in game attendance. This model has been used previously in the NBA and MLB implying applicability across sports. The model, borrowed from Johnson and Humphreys (2017), describes attendance as a function of various factors of demand following the general form

$$A_{ijst} = f(T_{ijst}, G_{ijst}, STAR_{ijst}, e_{ijst})$$

where A_{ijst} is total attendance for a game played by team i against team j on date t in season s.

T_{ijst} is a vector of variables that describes individual teams involved in each game including characteristics of team quality and team performance prior to the game taking place. Variables measured in the T vector are home and away team record prior to a game and a dummy variable indicating if the home or away team is the defending Stanley Cup champion. G_{ijst} is a vector of variables that describes the characteristics each game including metro population and if the game was played on a weekend. Variables measured in the G vector are metro area population, a dummy variable indicating a rivalry game, and a dummy variable indicating a weekend game. $STAR_{ijst}$ is a vector of variables that describes the presence of superstars or all-stars on team i or j. This is represented by three sets of variables encompassing the three models. Model 1 uses the number of all-star votes received by all players on a team in a given season for both home and away teams. Model 2 uses the number of all-star award winners in a season for both home and away teams. Model 3 uses dummy variables representing the presence of an individual superstar player on either the home or away team. The presence of multiple superstars present at a single game is controlled for using a dummy variable. The selection and measures of superstars will be discussed later.

The error term e_{ijst} is assumed to be a mean zero variable that is random accounting for all other characteristics of demand that are not included in the model that may vary by team. Johnson and Humphreys (2017) describe the model as “an approximation to the optimality conditions that emerge from a general economic model of attendance at sporting events that includes profit maximizing teams and utility maximizing customers/fans.” Given this, the model serves as a general model of fan attendance for regular NHL games.

Causality of an observable superstar effect is visible in the $STAR_{ijst}$ vector. As described, a superstar player is one who generates attendance and revenue at a rate greater than their on ice performance would predict, in this case contributions to wins. By controlling for other factors of demand including team quality and player performance, the mere presence of a superstar, captured by a dichotomous variable, should present a quantifiable effect on attendance if they are a superstar. The goal of this paper is not to distinguish at what point a player becomes a superstar or begins exhibiting a superstar effect, but rather to accomplish the same goal of previous papers. That is, to test whether that a superstar effect is observed in the NHL by analyzing the effect of the most successful and popular players who are most likely to exhibit such an effect. Should an effect be observed econometrically, then the superstar effect may be analyzed further to determine if it is better explained by the Rosen or Adler model.

An econometric issue faced in the model is the censoring of the dependent variable. Attendance is limited (censored) by stadium capacity and thus true effects of demand cannot be observed past a certain threshold which could affect the observed impact of a superstar. In order to account for this, previous studies have used a Tobit maximum likelihood estimator (Jane 2016). A Tobit estimator allows censored observations to be estimated beyond their limiting threshold. An issue with the Tobit estimator is that all observations of the dependent variable are required to be censored to the same level, i.e. the same stadium capacity. Instead, the censored normal estimator (Amemiya, 1973) used in Johnson and Humphreys (2017) will be employed. The merit of Amemiya's estimator is that its values may be censored at varied level across observations better predicting the effect of demand factors on attendance per game. This is conducted as an interval regression with an unlimited upper threshold permitting adjustments to censored observations.

An estimation of the model should provide quantifiable effects of these various controls for demand. The most important variables to observe are the star variables for presence of individual superstar players at home and away games. The coefficient will represent the average number of fans a superstar is responsible for bringing into the stadium based on their presence after all other determinants of demand. This coefficient represents a quantifiable superstar effect if observed.

Superstar Selection

The identification of superstars is inherently subjective; however, the measurements used to qualify a player as a potential superstar are consistent with previous studies. In particular, a method similar to Johnson and Humphreys (2017) will be used where superstar players must be considered a principal talent, maintain superstar status throughout their career having never regressed to a point of being considered an average player, and have careers represented in the data sample.

Players who best embody these characteristics are Wayne Gretzky, Mario Lemieux, Jaromir Jagr, Sidney Crosby, Alexander Ovechkin, Joe Sakic, Brett Hull, and Mark Messier. Other players considered but ultimately left out include Bobby Orr and Gordie Howe whose careers are well before the dataset, Steve Yzerman who has significant portions of his career data missing from databases, Mike Bossy whose career began several years prior to the sample, and Mats Sundin who was a first overall pick and captain but never won an MVP award. Both Crosby and Ovechkin continued playing after the 2017-2018 season which is the last season in this sample.

Table 1 provides an overview of each player’s statistics and qualifications for superstar status. The NHL has two awards for most valuable player: the Hart

Trophy, being awarded by a vote of the Professional Hockey Writers' Association to a player deemed most valuable to his team, and the Lindsay Award (formerly known as the Pearson Award) awarded to the most outstanding player in the regular season by a vote of the members of the NHL Players Association. Each player being considered has won one or several of these awards and has additionally played on Stanley Cup champion teams. All-star awards, as mentioned, are awarded to outstanding players at each position at the conclusion of the season. All players have received the award throughout their careers. The final statistic refers to point shares, an advanced statistic in which standing points (explained below) are attributed to the performance of each individual player. For example, Wayne Gretzky singlehandedly contributed 12.55 points to his team on average each season he played, equivalent to more than 6 wins.

Table 1: Superstar Qualifications

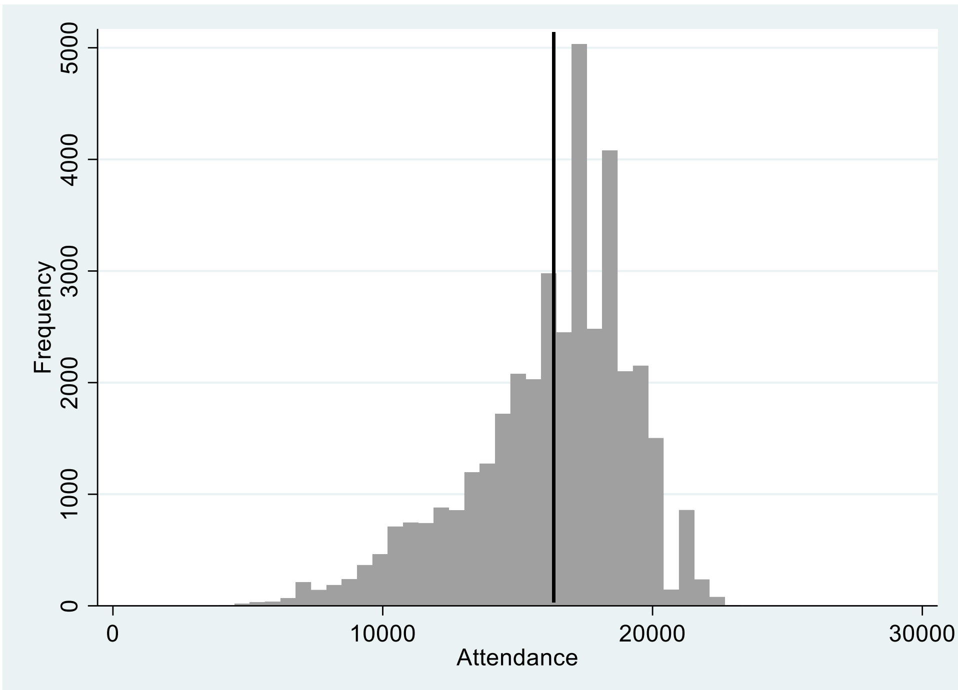
	Years Played	Hart Trophies	Pearson/Lindsay Trophies	All-Star Awards	Stanley Cups Won	Average Point Shares per Season
Wayne Gretzky	20	9	5	15	4	12.55
Mario Lemieux	17	3	4	9	2	10.5
Sidney Crosby	13 (active)	2	3	7	3	10.61
Alex Ovechkin	13 (active)	3	3	11	1	11.68
Jaromir Jagr	24	1	3	8	2	9.05
Brett Hull	17	1	1	3	2	10
Joe Sakic	20	1	1	3	2	9.9
Mark Messier	25	2	2	5	6	7.27

Data

The data is a collection of information for over 39,500 NHL games beginning in the 1979-1980 season until the 2017-2018 season. The 1979-1980 season is selected as the beginning of the sample due to that season marking the final merger between professional leagues. The NHL and its rival league, the WHA, merged prior to the season adding four new teams to the existing seventeen NHL teams. This merger established the NHL as the premier professional hockey league in North America. All subsequent expansions were in the form of the addition of expansion teams initiated by the NHL. Attendance data were collected from archived box scores from hockeydb.com from the 1979-1980 season to the end of the 1980-1981 season and from the 1993-1994 season to the end of the sample. Attendance values from this source were unavailable from the 1981-1982 season until the 1992-1993 season. These values were obtained from historical records of the *New York Times*. Sellout games occur in 13,927 observations, accounting for 36.5% of all games. There are 1318 missing attendance values that could not be found accounting for 3% of observations. Data missing from online databases were largely collected from the *New York Times*. As a result the majority of missing games are those played at night on the west coast, but the teams effected are largely random. Population data were collected from American and Canadian census bureaus while the remainder of hockey data were collected from hockeyreference.com. Figure 1 shows the distribution of attendance for all games used in the sample with the average shown by the black line.

There are a number of similar variables between this study and previous studies. Berri and Schmidt (2006), Jane (2014), and Johnson and Humphreys (2017) all use win percentages of home and away teams and population of the home team's city as controls for attendance. Stanley Cup championships won in the previous season, weekend games, and rivalries are introduced in Jane (2014). My analysis utilizes similar variables but adjusts them to the NHL's differing systems where applicable. Rather than using the mid-season all-star

Figure 1: Attendance Histogram



game, end-of-season award voting is used. Additionally, all-star voting is not conducted by fans but instead by a panel of professional writers. Johnson and Humphreys (2017) seek to test between the Rosen vs. Adler models by comparing the impact of fan all-star voting and performance statistics on game attendance to determine which is a better explanatory variable. As mentioned, my analysis will use a system based on age curves compared to observed superstar effects rather than using an empirical model like Johnson and Humphreys that compares data used to reflect player popularity and performance to game attendance. This is because similar measures are unavailable for the NHL.

Table 2 contains summary statistics for continuous variables used in the model. Game attendance reflects the number of fans present at an NHL game typically played in a dedicated arena. The NHL has had games scheduled in neutral cities like Halifax, Nova Scotia or Orlando, Florida. There are additionally a series of special regular season games played at large outdoor stadiums that account for the 25 largest crowds in the dataset. These observations have been dropped from the data as they are special circumstances which do not reflect general regular season conditions. These games are not included in Figure 1 either.

The win percentage variable reflects the overall performance of the teams prior to the game analyzed. This control is perhaps the most significant as previous studies have concluded that attendance is driven heavily by a team's capacity to produce wins. The NHL utilizes a point system to determine standings that has evolved over time. By modern rules, two points are awarded for any win, one point is awarded for an overtime or shootout loss, and no points are awarded for a regulation time loss. The changes in awarding points have revolved around overtime rules. Prior to the 1983-1984 season, games that concluded in regular time with the same score would be recorded as ties with each team receiving one point. Between the 1983-1984 season and the 1999-2000 season, an additional five minutes would be played in overtime where a game ending in a tie would again award a single point to each team, and two points were earned by the team winning in overtime with no point to the loser. Between the 1999-2000 and

Table 2: Summary of Continuous Variables

	Mean	Std Dev	Min	Max
Game attendance	16246	3038	1664	27227
Home team record prior to game	.493	.147	0	1
Away team record prior to game	.496	.147	0	1
Metro Area Population (millions)	4.416	4.589	.256	20.351
Number of All-Star votes received by home team	182	274.69	0	1924
Number of All-Star votes received by away team	181	274	0	1924
Stadium capacity	17836	1817.4	7424	27227
Observations	38149			

2005-2006 seasons, each team would be guaranteed an awarded point if the game went to overtime with an overtime win awarding two points. From the 2005-2006 season to the time of this paper, ties are no longer possible. Instead an overtime period concluding in a tie will go to a shootout with the same modern point awarding system listed above. By awarding overtime and shootout loss points, there exists point inflation between years. To accommodate for this, the 5-point system instituted by Ralph Slate of Hockeydb.com is used awarding 5 points per regulation win, 4 points per overtime win, 3 points per shootout win, 2 points per shootout loss, 1 point per overtime loss, 0 points per regulation loss, and 2.5 points per tie which is then divided by the maximum possible point value (5 points times number of games played). This should accurately reflect the teams' capacity to win and standardize performance across eras. The average is skewed due to the missing attendance observations giving a value less than .500 for both home and away teams.

For both home and away teams, all-star data was collected as a measure of team performance to be tested against the performance of individual superstar players. All-star data refers to voting at the conclusion of the season for players considered exceptional. All-stars are voted on by members of the Professional Hockey Writers Association. All-stars are divided into first and second teams with a single player for each offensive position (left wing, right wing, and center), two defensive players, and a goaltender. In total there are twelve all-stars each season considered the best in their position. I measure the number of all-stars on each team for each year. The number of all-star votes received by home and away teams is an additional qualifier of talent. Rather than the presence of high-level players, the overall quality of the team is reflected as there are recipients who do not make either all-star team.

Dichotomous variables are summarized in Table 3. These variables account for game characteristics, team performance, and the presence of superstar players. Standard deviations are not reported. Defending Stanley Cup champions are present in 7.6% of all game observations and are included to account for possible fan preference for watching defending champions play. 15.6% of games are between rival teams. These rivalries may exist for historical or geographical reasons such as Original Six teams, regional rivals, or division rivals. Notable examples of rivalries include the Battle of Ontario between the Toronto Maple Leafs and Ottawa Senators as well as former rivalries such as the Battle of Quebec between the Montreal Canadiens and the Quebec Nordiques. There appears to be a propensity for the NHL to schedule games on Friday,

Table 3: Summary of Dichotomous Variables

	Mean
Home team defending Stanley Cup champion	.038
Away team defending Stanley Cup champion	.038
Rivalry	.156
Weekend	.464
Multiple superstars present	.023
Wayne Gretzky playing for home team	.017
Wayne Gretzky playing for away team	.018
Mario Lemieux playing for home team	.012
Mario Lemieux playing for away team	.010
Sidney Crosby playing for home team	.011
Sidney Crosby playing for away team	.011
Alex Ovechkin playing for home team	.013
Alex Ovechkin playing for away team	.013
Jaromir Jagr playing for home team	.023
Jaromir Jagr playing for away team	.022
Joe Sakic playing for home team	.018
Joe Sakic playing for away team	.017
Brett Hull playing for home team	.015
Brett Hull playing for away team	.016
Mark Messier playing for home team	.018
Mark Messier playing for away team	.016
Observations	38260

Saturday, or Sunday accounting for almost half of all games. Superstar players appear in relatively few games, none exceeding 2.3% of the sample. This is exploited in order to analyze the effect of superstar players' presence on game attendance. There are relatively few instances of multiple superstars in a single game—890 with the majority a result of Wayne Gretzky and Mark Messier who played several seasons as teammates—but this anomaly is controlled for nonetheless. These superstar variables were collected using game level performance data from hockeyreference.com. Superstars were considered present in all instances where they played in a game regardless of ice time. Game time decisions and single game absences are not counted as the effect of their absence is likely not observed by fans since they often do not know of the absence until just before a game and thus will anticipate the player's participation. For significant and extended absences such as injury, presence is not included. This includes injuries like Wayne Gretzky's herniated disk at the start of the 1992-1993 season and absences such as Mario Lemieux's 1993 cancer treatment.

Results

Regression results are illustrated in table 4. The use of a censored normal estimator in the regression prevented the use of fixed effects. Instead fixed effects for season, home team, and away team were added manually using dummy variables. These fixed effects should control for the various issues by standardizing attendance given larger stadiums or growth in popularity of the sport. Their coefficients and results are omitted from the listed results. All sold-out games were given an indication and for all such observations the Amemiya estimator was used. This type of regression prevented the use of a Huber-White

Table 4: Censored Normal Estimator Regression Results

	(1)	(2)	(3)
Metro Population of home team	-555.6***	-577.6***	-802.17***
	(41.38)	(41.34)	(42.98)
Home win percentage prior to game	2,916***	3,025***	3,694.3***
	(112.1)	(111.3)	(106.35)
Away win percentage prior to game	170.6	118.8	418.18***
	(110.9)	(110.2)	(104.77)
Home team defending champions	1,571***	1,609***	1.841.61***
	(88.31)	(88.43)	(89.23)
Away team defending champions	474.8***	446.0***	555.96***
	(84.81)	(84.80)	(84.93)
Rivalry game	738.1***	731.4***	745.5***
	(45.63)	(45.62)	(45.66)
Weekend game	1,207***	1,205***	1,206.59***
	(30.48)	(30.46)	(30.46)
All-star votes received by home team	1.935***		
	(0.0698)		
All-star votes received by away team	0.794***		
	(0.0666)		
All-star players on home team		621.4***	
		(23.44)	
All-star players on away team		311.0***	
		(22.76)	
Gretzky playing for home team			3,664.8***
			(181.7)
Gretzky playing for away team			1,546.4***
			(138.7)
Lemieux playing for home team			1,460.4***
			(214.1)
Lemieux playing for away team			327.2
			(222.7)
Crosby playing for home team			295.9
			(185.2)
Crosby playing for away team			459.2***
			(197)
Ovechkin playing for home team			2173.3***
			(183.9)
Ovechkin playing for away team			844.2***
			(168)
Jagr playing for home team			-66.3
			(125.9)
Jagr playing for away team			207.3
			(129.2)
Sakic playing for home team			2546.6***
			(166.4)
Sakic playing for away team			162.6
			(161.6)
Hull playing for home team			2182.1***
			(168)
Hull playing for away team			479.8***
			(139)

Messier playing for home team			-103.2
			(158.8)
Messier playing for away team			305*
			(161.6)
Multiple superstars present			1642.2***
			(123.5)
Observations	38,149	38,149	38,149
Number of Home Teams	37	37	37
rho	0	0	0
sigma_u	5.17e-08	0	0
sigma_e	2655	2655	2659

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

correction and a standard cluster correction, instead an observed information matrix estimator is used.

The regression results show that all controls are statistically significant to the 1% level for the third model examining individual superstar effects. Interestingly, population appears to have a negative effect on attendance such that attendance falls by 802 fans for every million residents in the metropolitan area. This is contrary to the expected result where larger market teams will have a larger audience to attract fans from. An important point is that these results may be skewed by the results of extremely large markets such as New York City, Los Angeles, and New Jersey whose immense population encompasses every year in the sample and includes multiple teams, potentially creating bias in the results. The effects of win percentages, for both home and away teams, are positive as expected reflecting the attractive nature of a better performing team.

The variables encompassing stardom are most significant to this analysis and are reflected in the remaining variables. All-star votes and all-star award winners are separated in the results as all-star votes should reflect the overall quality of a team while all-star ward winners describe the presence of high caliber players. For each, the effect on both home and away games is statistically significant and positive. The effects appear to have a more significant effect on home games with all-star votes and all-star award winners having nearly twice the effect on home attendance as away attendance. These results may be interpreted in two ways. First, the results of all-star votes imply that player quality is indeed a determinant of demand in professional hockey for both home and away teams. Second, the all-star award results imply that, aside from the overall quality of a team, individual high-caliber players are an additional determinant of demand. This implies that there is a larger component to team performance when discussing demand for professional hockey beyond a team's capacity to produce wins.

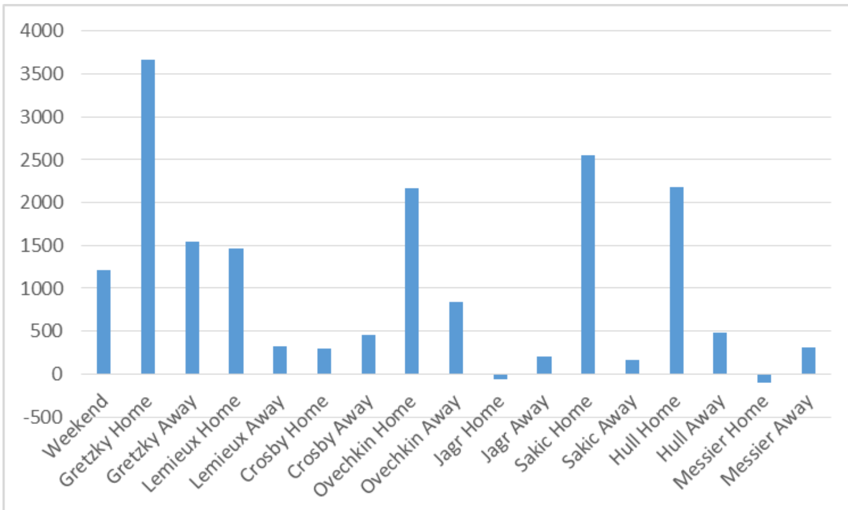
Using a similar method to Berri and Schmidt (2006), the marginal effect on attendance for a single win on attendance is the coefficient for win percentage divided by the number of games in a season, 3,694.3/82. This shows that for every additional win on average, a team will see an increase in attendance of 45. Comparing this to Wayne Gretzky's average point shares, the additional wins contributed by his play caused an average increase of 546.8 fans per game. This shows that the presence of a star player has a larger effect on game attendance than their capacity to produce wins, opposite the findings for the NBA. This does not necessarily show that a *team's* capacity to win is not the most significant determinant of demand.

The results for each individual player are listed in figure 2. Wayne Gretzky, as expected, has the largest effect for both home and away attendance. There are a number of instances where players generate substantial increases in attendance for home games but comparatively small effects on away attendance. This result is reminiscent of Brandes, Franck, and Nüesch (2008) regarding local heroes and superstars. These types of results are visible for Mario Lemieux, Joe Sakic, and Brett Hull. Sidney Crosby, who is commonly considered the premier talent in the NHL for several years, has a smaller effect than rival competitor Alexander Ovechkin. It is interesting that he does not exhibit a superstar effect given this perception. While Crosby has more point shares and a higher point per game average, Ovechkin is considered the premier goal scorer in the league. These results may imply that fan preference is given to the ability to create goals rather than the overall performance of a player. Another result of interest is the absence of observable superstar effects for Jaromir Jagr and Mark Messier despite consistent discussion of their “generational” talents.

Johnson and Humphreys (2017) use the effect of weekend games as a comparison for individual superstar effects. They determine that a player whose observed superstar effect is less than the effect of having a game moved from a weekday to a weekend cannot be considered a superstar. Using this criteria, only Wayne Gretzky exhibits a true superstar effect while several others have a home superstar effect. This criteria will not be used for the same reasons that different characteristics are used in an estimation for professional hockey; the games are fundamentally different.

The effects of Jaromir Jagr and Mario Lemieux are particularly interesting. Mario Lemieux is often considered to be one of the top five greatest players of all time in company with Wayne Gretzky, yet there is no significant effect on away games across his career. Lemieux is well known for taking the once miserable Pittsburgh Penguins to Stanley Cup contention which may explain his significant

Figure 2: Individual Superstar Effects



effect on home attendance. This may be evidenced by the Penguins' previous stadium, which was colloquially referred to as "The House Lemieux Built." It is perhaps for this reason that Lemieux is considered a superstar player; however, his effect on away attendance indicates that he cannot be considered a true superstar despite his elite performance. Jaromir Jagr exhibits some of the lowest effects of any player for both home and away games. While Jagr won multiple MVP trophies and multiple Stanley Cups, it is possible that the later part of his career which saw him play for eight teams across eleven seasons contributes to this. A similar problem arise for Mark Messier. Messier is commonly recognized for his exceptional leadership as a captain. His six Stanley cup victories are often cited as a testament to this. It is perhaps this reason why he is commonly mentioned among conversations regarding the best players in hockey; however, it would appear that exceptional leadership is not a significant determinant of demand and thus does not generate a superstar externality. Given these results, statistically significant effects are observed for Wayne Gretzky, Brett Hull, and Alexander Ovechkin indicating a superstar effect. However, Brett Hull's away results are not large enough to warrant consideration as a true superstar; rather he appears to be a local star.

Adler vs. Rosen Model

As discussed, a player's superstar effect will be compared to their performance across their career. As Mario Lemieux, Jaromir Jagr, Joe Sakic, Sidney Crosby, and Mark Messier failed to exhibit statistically significant superstar effects, they will not be examined. Brett Hull exhibited a quantifiable effect for both home and away games but his away effect was too low to consider a true superstar. Wayne Gretzky and Alexander Ovechkin exhibit strong career superstar effects and as such will be examined for an explanation of the superstar effect's dominating theory.

Adjusted point shares will serve as the performance metric that a player's superstar effect will be compared to. Point shares are calculated using a player's overall performance in a season using values such as goals and assists. This is compared to the league average to determine the player's performance relative to the rest of the league. The difference between the player's performance and the average is then compared to the performance of the team which is represented by the aforementioned standing points where a win is counted as two point and an overtime loss, shootout loss, or tie is counted as one point such that their individual contribution to team wins is measured.[4]

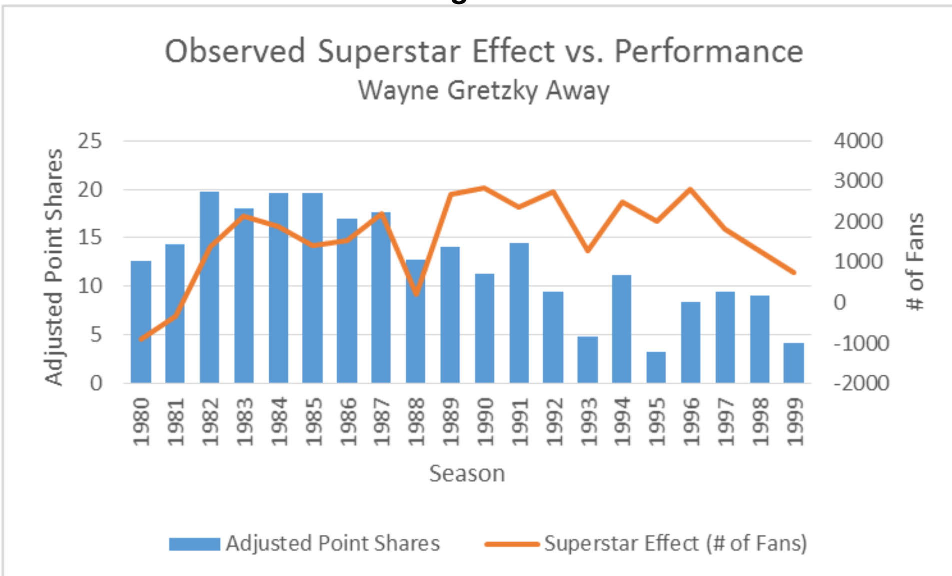
To compare a player's superstar effect to their season-by-season performance, their season-by-season superstar effect must be used. This is done by simply running a regression that uses an interaction term of the player's presence, represented by the STAR variable, and a dummy variable for each season. This provides individual coefficients for each season.

Figure 3 below shows the relationship between Wayne Gretzky's on-ice performance, measured by adjusted point shares, compared to his observed superstar effect. The labels for season in all graphs refers to the year in which the season concluded such that 1980 represents the 1979-1980 season. These data points are used only for away games as the movement of Gretzky to different teams throughout his career may affect the perceived effect whereas away games should have relatively consistent conditions. The same method is conducted for Ovechkin.

Despite being considered an outstanding player prior to entering the NHL, Gretzky did not exhibit a quantifiable superstar effect until his third year, the 1982 season, after winning consecutive MVP awards. The negative value exhibited in Gretzky's first two seasons may be the result of a confounding variable like Edmonton's status as an expansion team following the NHL-WHA merger prior to the 1979-1980 season. The significant drop in value during the 1987-1988

[4] Adjusted point shares account for missed games by calculating the number of point shares a player would have earned if they had played an entire season at the same level of performance. This is done so that a players overall performance may be standardized the in the same way their superstar effect is.

Figure 3

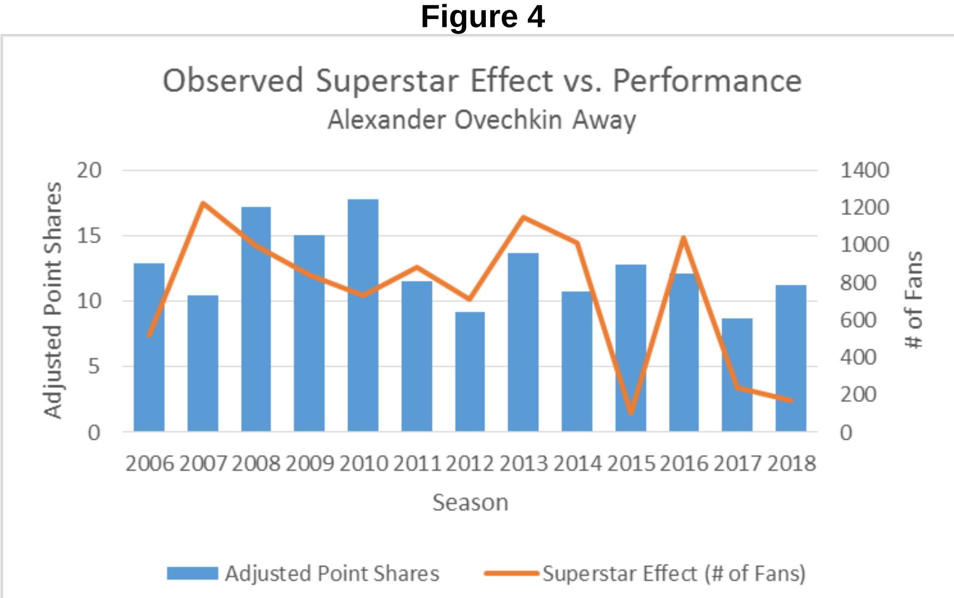


season could be the result of several factors such as the economic conditions following the Black Monday stock market crash in 1987 at the beginning of the season or the rise of fellow superstar Mario Lemieux who would end Gretzky's 8-year MVP streak that season.

Gretzky's effect remained relatively stable aside from a drop in his final season in Edmonton in the 1987-1988 season which culminated in a Stanley Cup victory. Following his trade to Los Angeles prior to the 1988-1989 season, Gretzky began to exhibit a larger superstar effect than ever before. While there does appear to be a small downward trend with relative performance in Gretzky's later years, his observed effect on attendance remained at a magnitude much higher than his performance. The trend following performance should not be considered insignificant. It is perhaps the case that superstars, especially in sports, are driven by their talent which evolves into popularity. As observed with Gretzky, the overall effect increases over time but is still subject to his own performance. Given this, it would appear that Wayne Gretzky may be considered an Adler superstar who derives their superstar status from popularity but is effected by the constraints of Rosen's model. This seems to be a reflection of MacDonald's model which discusses the benefits older superstars receive via information effects, a cause of popularity, but superstar status being uniquely derived from talent.

Figure 4 illustrates the superstar effect of Alexander Ovechkin across his career. While Ovechkin has not yet completed his career, there is an observable diminishing effect on performance over time consistent with the Brander, Egan, and Yeung's (2014) age curve model. Unlike Gretzky, we see an immediate effect on attendance in Ovechkin's rookie year for which he was awarded the Calder trophy given to the rookie of the year. This may be an indication of the improvements to communication and information sharing between Gretzky's era and the modern era. In the 2008, 2009, and 2013 season, Ovechkin was awarded the Hart Memorial trophy, the NHL's MVP award. Ovechkin's observed effect appears to regress relatively consistently with his on-ice performance and

Figure 4



does not appear consistent or upward-trending across is career unlike Gretzky indicating that Ovechkin is likely a Rosen superstar whose impact on attendance is driven more by production on the ice than it is by popularity that is disconnected from on-ice production.

Conclusion

The superstar effect, as first proposed by Rosen (1981), describes the phenomenon whereby the demand for the services of certain individuals in a given market exceeds that of their competitors by a magnitude greater than can be explained by their differences in ability or talent. The element of popularity and information effects was added by Adler (1985) as another cause of this phenomena. Using similar methods to previous studies, this paper provides evidence that there is a quantifiable superstar effect in the National Hockey League. As expected, the effect is smaller in magnitude than what is observed in professional basketball. Additionally, Wayne Gretzky is clearly the greatest superstar in league history based on his observed effect on attendance.

The National Hockey League has never been analyzed for the presence of a superstar effect prior to this paper and it is the first of its kind to use a dataset encompassing such a large period of time in any sport. Using this data , the entire careers of superstar players may be analyzed to determine their effect on team attendance. This effect has important implications for teams that may wish to maximize revenue when a superstar player is either drafted or acquire in trade or free agency . An important observation is that no player will likely ever have the same effect as Wayne Gretzky.

Using a new method to analyze the two theories of superstar status, it is possible that there are different determinants in different eras. In the information age, it seems clear that a player such as Alexander Ovechkin has a superstar effect largely driven by talent and performance whereas Gretzky, at least later in his career, appeared to derive most of his superstar effect from residual popularity. There appears to be a presence of both effects in each superstar indicating that the true basis of the superstar effect lies in both theories and that they are not mutually exclusive, at least in professional sports. This is contrary to the findings of Johnson and Humphreys (2017) in a similar study conducted for the NBA. This could imply that the effect varies in nature across types of leagues as well. Despite hockey’s inherent unpredictability, there still emerge players that exhibit superstar effects. Using these results, it may be the case that superstar effects are more prevalent than previously thought and as such may help explain growing issues of income inequality not only in sports but across markets as Rosen and Adler imply.

References

Adler, M. (1985). Stardom and talent. The American Economic Review, 75(1):208–212.

Berri, D. J., Schmidt, M. B., & Brook, S. L. (2004). Stars at the gate: The impact of star power on NBA gate revenues. Journal of Sports Economics, 5(1), 33-50.

Berri, D. J., & Schmidt, M. B. (2006). On the road with the National Basketball Association's superstar externality. Journal of Sports Economics, 7(4), 347-358.

Borland, J., & MacDonald, R. (2003). Demand for sport. Oxford review of economic policy, 19(4), 478-502

Brander, J. A., Egan, E. J., & Yeung, L. (2014). Estimating the effects of age on NHL player performance. Journal of Quantitative Analysis in Sports, 10(2), 241-259.

Brandes, L., Franck, E., & Nüesch, S. (2008). Local heroes and superstars: An empirical analysis of star attraction in German soccer. Journal of Sports Economics, 9(3), 266-286.

Fort, R. (2000). Market power in pro sports. The Economics of Sports, 1-54.

Franck, E., & Nüesch, S. (2012). Talent and/or popularity: what does it take to be a superstar?. Economic Inquiry, 50(1), 202-216.

Hansen, H., & Gauthier, R. (1989). Factors affecting attendance at professional sport events. Journal of sport management, 3(1), 15-32.

Hausman, J. A., & Leonard, G. K. (1997). Superstars in the National Basketball Association: Economic value and policy. Journal of Labor Economics, 15(4), 586-624.

Idson, T. L., & Kahane, L. H. (2000). Team effects on compensation: an application to salary determination in the National Hockey League. Economic Inquiry, 38(2), 345-357.

Jane, W. J. (2016). The effect of star quality on attendance demand: The case of the National Basketball Association. Journal of Sports Economics, 17(4), 396-417.

Johnson, C., & Humphreys, B. R. (2017). The Effect of Superstar Players on Game Attendance: Evidence from the NBA. SSRN Electronic Journal.

Jones, J. C. H., Nadeau, S., & Walsh, W. D. (1999). Ethnicity, productivity and salary: player compensation and discrimination in the National Hockey League. Applied Economics, 31(5), 593-608.

MacDonald, G. M. (1988). The economics of rising stars. The American Economic Review, 155-166.

Paul, R. J. (2003). Variations in NHL attendance: the impact of violence, scoring, and regional rivalries. American Journal of Economics and Sociology, 62(2), 345-364.

Rosen, S. (1981). The economics of superstars. The American economic review, 71(5), 845-858.

Vollman, R., Awad, T., & Fyffe, I. (2016). Hockey Abstract presents stat shot: The ultimate guide to hockey analytics. Toronto, ON: ECW Press.