

**Can Small Markets Compete?  
Empirics on Competitive Balance in the NHL**

**Craig Wakeford  
3059139  
ECO7997**

**In partial fulfillment of requirements for the Master's Degree  
in Economics at the University of Ottawa**

**Supervisor: Professor Marc Lavoie  
Date: August 22, 2003**

## **Can small-markets compete? Empirics on Competitive Balance in the NHL**

*One of the key ingredients of the demand by fans for team sports is the excitement generated because of the uncertainty of outcome of league games. For every fan who is a purist who simply enjoys watching athletes with outstanding ability perform regardless of outcome, there are many who go to watch their team win, and particularly to watch their team win a close game over a challenging opponent.*

(Quirk & Fort p.243)

Uncertainty is one of the key elements that compel fans and people in general to watch a professional sports game. The competition of the top athletes in a particular sport captures the attention of a lot of people worldwide. They go to a game, or watch it on television because they think they are going to see a close match, one that the athletes fight hard to win until the last second. In the last couple of decades, people have started to complain about a divergence between large-market and small-market franchises. The consensus is that small market teams can not hire enough talent to compete with large-market franchises, mostly from New York and Los Angeles. This is why there has been much concern over the competitive balance of sports leagues.

In a truly competitively balanced league no one team has an advantage over the other. On any given day or night, a team with the lowest revenue, or least talent has as good of a chance of winning as any other team. Perfect competitive balance is only observed if there is complete uncertainty in the outcome of every game. Theoretically this could happen, but in practice, it never will. There is always going to be a few teams every year that are just a little better than the rest. This paper addresses the question of balance in the National Hockey League.

More specifically, the paper addresses the balance of the league over the past 13 years. I have collected data on payroll, revenue, regular season performance and playoff performance from the 1989-1990 season until the 2001-2002 season. With this data I will measure the imbalance in all of the above variables. While strict balance among the listed variables is interesting, what most people want to know is if the imbalance of revenue or payroll directly correlates to imbalance in the performance. Does a rich team have a greater chance of winning than a poor team?

In the sample studied, there has been rising monetary imbalance. Monetary variables are revenue and payroll. Outcome variables are considered to be regular season and playoff performance. The league has not seen the same kind of rise in disparity in the so-called outcome variables. As well, there is a slight relationship between money and regular season performance, while, there is no apparent relationship between money and playoff results.

The paper is organized in the following manner. The following section gives a brief history of the evolution of the National Hockey League (NHL). The second section presents data on some trends that have developed over the course of the sample period. Presented after that, are problems that could be found with the data and the sample size. The fourth section presents the results obtained from the analysis of competitive balance. Balance exhibited by single variables is shown first, followed by correlation between variables and finally the relationships between the monetary variables and the outcome variables. The conclusion offers some areas for future research.

There have been several papers written on this topic. Economists such as Richardson (2000), Zimbalist (2002), Fort (2003), Leifer (1995), and Quirk & Fort (1992), have all examined the issue of competitive balance in the NHL. While these papers did look at balance or imbalance of single variables, very few looked at the correlation between variables and none looked at a regression analysis to determine the relationship between the money variables and outcome variables. This is how my paper extends beyond previous papers.

## **I - A Brief History**

The National Hockey League has been around for over seventy-five years and has been the premier hockey league worldwide since its inception. In its first few years, there were several different teams that competed in the league, but after 1926, the league has been comprised, for the most part, of six mainstay teams. These teams were the Boston Bruins, Chicago Blachhawks, Detroit Red Wings, Montreal Canadiens, New York Rangers, and the Toronto Maple Leafs. In 1967, the league expanded to twice the size.

Since then the league has gone through numerous changes and seen many teams come and go. The league now consists of thirty teams.

The NHL was believed to be at its most competitive when only the original six teams existed. All of these teams had large markets to draw from and each had a long history in the area, with an honest following. There was still, however, an imbalance that existed at that time. This was not due to some teams being able to afford more talent than others, but was instead due, in some part, to the immobility of players.

Nowadays, television contracts contribute large sums to the revenue of each franchise. Television first appeared at mid-century and the networks were not paying the amounts of money they are today for the rights to televise games, thus exerting a much smaller influence on league dynamics. This meant that there was a more level playing field in terms of money. If the teams could generate somewhat the same amount of revenues at the actual game, then they could afford to put the same calibre of team on the ice as their competitors. Other elements such as hard work, team spirit and coaching ability helped determine team performance. Unfortunately these days are long gone.

From the small-market Canadian teams such as Edmonton and Calgary, to the large-market American ones like Los Angeles and New York, there is a vast difference in the amount of revenues generated by each franchise. Teams in large markets may not do as well performance wise, but they can still generate substantial revenues, while teams in smaller markets may fare extremely well while barely grossing enough earnings to stay afloat. Recent examples of this are the Winnipeg Jets and Quebec Nordiques. While they did not win the Stanley Cup – the National Hockey League's ultimate prize – they still performed reasonably well during the years just before they were relocated. Both teams entered the playoffs in the last year they were in their respective original cities. The Nordiques even finished second the year that they moved. At the other end of the economic spectrum, there are teams like the New York Rangers and the Los Angeles Kings. Coming from the two biggest markets in the National Hockey League, or any North American sports league for that matter, both failed to qualify for the playoffs last year. The Rangers have, in fact, not made the playoffs for the past six years. In spite of that, these two teams are among the highest revenue teams in the league.

## **II – Data and Recent Trends**

### *Data*

I have collected data on various aspects of each individual team. The complete data set starts with the 1989-1990 season and continues through till 2001-2002. I also have some data for the 2002-2003 season, but it is an incomplete data set because full information on payroll and revenue are not released until the middle of the following year. The data was collected from various sources. Figures on monetary variables were collected from either Financial World or Forbes. Financial World used to supply the statistic until the mid-1990's when Forbes took over. Regular season statistics were collected from [www.nhl.com](http://www.nhl.com). This website contains all regular season history from the early twentieth century. Data for playoff performance was collected of another website [www.whowins.com/tables/all.htm](http://www.whowins.com/tables/all.htm). This website contains all playoff information for every major North American sports league. It contains data on which teams played in the series, how many games were played in the series and who won the series. It contains data listed back until the 1905 season.

There are some problems that could arise with this data. Payrolls may be misrepresented by calculating team payroll with a different number of players for each team. Because some teams will go through more players than others (due to injuries, etc.), one team's payroll could include thirty players, while another's could include only 24. This could skew the payroll statistics.

However, wherever possible, it was verified that the same number of players were used in the payroll figure of each team (i.e. the top 24 players). Elsewhere, the figure was calculated using the top twenty-four paid players of each team. This makes the relative ranking similar.

Difficulties in measuring revenue can also arise. Unlike the players which have allowed their salaries to be public knowledge, the owners do not legally have to reveal information on revenues and costs. These numbers are estimated by analysts from either Financial World or Forbes.

Nevertheless, the same formula was used in estimating every team's revenue from different sources (i.e. gate revenue, merchandise revenue, etc.), so measurement error should be kept to a minimum in this category.

Problems can also arise in the measurement of winning percentage. Winning percentage is measured by taking the total number of points won by a team and dividing that by the total number of points that could have been obtained. Up until the 1999-2000 season, the NHL used one type of point system. With a win, the team will earn two points, with a tie, one point, and with a loss, zero points. In an eighty-game season, a team has the opportunity of gaining 160 points. However, from the 2000-2001 season a different point system has been used. Two points are still awarded for a win, but, if a team loses in overtime, it is given one point for a regulation tie. This could provide a team with a higher winning percentage than they have actually achieved.

In order to remedy this, I have manipulated the data from the 2000-2001 season onwards. I calculated the total amount of points awarded in the entire league and then divided this number by the total amount of points that would be awarded under the previous point system. Then I took this number and divided every team's points by it. In the end, this manipulation allows me to compare every year equally.

With this data, and more historical data, that I have obtained from other papers, one can discern a trend in disparity arising in the National Hockey League.

### *Payroll*

A large increase in disparity has occurred in the payroll over the course of the sample. During the 1989-1990 season, the average payroll per team was given as \$6,276,190 and the median was given as \$5,800,000. All numbers are given in American Dollars. Although there is a slight skew to the right, it is a very small skew with a sample size like this. The skew ratio (Average divided Median) is 1.08, which means that there are a couple of teams that are pulling the average up, while there remains a majority of teams that are lagging behind, bringing down the median. A perfectly normal distribution – optimal in analyzing economic data – would have a skew ratio of 1.

In this year the maximum payroll belonged to the Los Angeles Kings with a payroll of 9,500,000. This high payroll was due mainly to the fact that Wayne Gretzky, arguably the greatest player of all time, played for the Kings at this point in time. The

minimum payroll belonged to the Vancouver Canucks, which tallied a payroll of \$4,800,000. The difference of \$4,700,000 between these two ends of the pole is a range that would be considered a well-balanced league by today's standards.

<b>Year</b>	<b>Average</b>	<b>Median</b>	<b>Skew Ratio</b>	<b>Maximum</b>	<b>Minimum</b>
1989-1990	\$ 6,276,190.48	\$ 5,800,000.00	1.0821	\$ 9,500,000.00	\$ 4,800,000.00
1990-1991	\$ 8,038,095.24	\$ 8,000,000.00	1.0048	\$11,200,000.00	\$ 4,700,000.00
1991-1992	\$ 10,168,181.82	\$10,050,000.00	1.0118	\$14,500,000.00	\$ 6,200,000.00
1992-1993	\$ 11,804,166.67	\$11,550,000.00	1.0220	\$17,000,000.00	\$ 7,500,000.00
1993-1994	\$ 12,950,000.00	\$13,350,000.00	0.9700	\$20,700,000.00	\$ 7,500,000.00
1994-1995	\$ 15,967,500.00	\$14,733,500.00	1.0838	\$24,320,000.00	\$ 9,365,000.00
1995-1996	\$ 21,638,461.54	\$21,850,000.00	0.9903	\$34,100,000.00	\$11,700,000.00
1996-1997	\$ 22,941,651.00	\$22,429,791.50	1.0228	\$37,908,000.00	\$14,508,100.00
1997-1998	\$ 26,788,269.23	\$25,726,000.00	1.0413	\$44,216,000.00	\$17,547,000.00
1998-1999	\$ 29,797,037.04	\$30,430,000.00	0.9792	\$48,270,000.00	\$13,640,000.00
1999-2000	\$ 31,540,001.89	\$31,307,255.00	1.0074	\$59,419,011.00	\$16,590,000.00
2000-2001	\$ 34,267,433.33	\$32,004,500.00	1.0707	\$57,117,000.00	\$18,400,000.00
2001-2002	\$ 35,077,048.37	\$31,273,750.00	1.1216	\$65,518,750.00	\$18,438,576.00

This is in large contrast to the data that was collected from the later years in the sample. A look at the figures in Table 1 show that the teams in the lower portion of payroll are becoming worse off than they were thirteen years ago. At the beginning of the sample, the minimum payroll was three-quarters the league average, whereas now the minimum is half the league average. Using the same analysis, the team with the highest payroll team has improved their situation. At the beginning of the sample, they were one and one half times the average payroll, whereas they are now over two times the average payroll.

There is also a noticeable amount of inflation during this time period. With the average going from 6.25 million to just over 35 million, the payroll exhibits an inflation rate of 16% per year. This is in contrast to a GDP inflation rate of roughly 3% over this time period.

The skew ratio, on the other hand follows no standard pattern and can be seen as random. As before, a skew ratio above unity signifies that there are a small number of franchises who have very high payrolls, therefore bringing up the average and leaving the median lagging behind. A skew ratio below unity means that there are a few teams that have very low payrolls, thus bringing down the average and leaving the median higher.

A possible reason for the randomness of the skew ratio and its fluctuation around unity is the introduction of new teams, or expansion teams as they are called in the industry. A closer look at the data reveals that the ratio dips below one, either the year that there was an expansion team, or the year directly following it.

It is also interesting to note that the maximum – the team with the highest payroll – climbs faster than the average more often than not. Further examination of the data reveals that the maximum climbs faster than the average 8 out of the 12 years or 66% of the time during the years in the sample. The minimum follows the same pattern although not quite as often. It climbs faster, or just as fast as the average in 7 out of the 12 years. This is most likely due to expansion teams as well. Most of the time, expansion teams will not have a huge payroll because of the fact that they do not immediately possess star players. This also has the implication of pushing up payroll disparity within the league.

Another way that the disparity can be shown is the rise in the range over time. The range is defined as the minimum subtracted from the maximum; a rise in the range will also indicate a rise in disparity. A higher payroll implies that the franchise has, under its payroll, quite a few star players. Two current examples of this are the Detroit Red Wings and the Colorado Avalanche. Both teams possess payroll far over 60 million dollars per annum. This is in excess of more than half of the other franchise's revenues. There is absolutely no way that other teams could possibly hire the star talent that these two teams enjoy. This makes it difficult for teams with small revenues to compete against these teams.

The rapid rise in payroll has mainly been attributed to the player reservation system or the free agent status of a player. The NHL still has fairly strict policies of free agency – the player's allowable mobility after his contract expires – as compared to the rest of the big four North American sports. However, these regulations are loosening, allowing players more mobility after the completion of their contract. This, in turn, means the player can go to where he would be most valuable and therefore earn a higher salary. As compared to 1989-1990 (the start of the sample) where a full team's payroll was at a maximum of 9.5 million, the minimum at the end of the sample, 2001-2002, was \$18.5 million. In fact, there are players now who have signed contracts for \$11 million per season. One person is making more than a full team 12 years ago.

Richardson (2000) noticed the same trend for average players' salaries. His sample period was 1984-1985 to 1993-1994. He found that during his sample period the average salary rose by over 250%. It was also found in his study that it was during the 1989-1990 season that average salaries really started to rise. This is concurrent to the time that my analysis started.

### *Revenues*

There are some similar trends in the revenues obtained by NHL teams as well, as can be seen in Table 2. However, one thing to notice first is the dip in the revenue average and median in the sixth year. This is the 1994-1995 year, or the year that there were only 48 games played due to a lockout or strike. The year was shortened considerably due to disagreement on the collective bargaining agreement (CBA) by the owners and the National Hockey League Player's Association (NHLPA), the governing body (union) which represents the players. This dip though is not apparent in the payroll. This is due to the data on salaries being based on full year salaries. Even though the players were not paid in full for this year, the data was compiled as if they were.

<b>Recent Trends in Revenue</b>					
<b>Year</b>	<b>Average</b>	<b>Median</b>	<b>Skew Ratio</b>	<b>Maximum</b>	<b>Minimum</b>
1989-1990	\$20,947,619.05	\$21,300,000.00	0.9835	\$ 28,800,000.00	\$11,100,000.00
1990-1991	\$20,947,619.05	\$21,300,000.00	0.9835	\$ 28,800,000.00	\$11,100,000.00
1991-1992	\$26,118,181.82	\$23,750,000.00	1.0997	\$ 50,700,000.00	\$17,000,000.00
1992-1993	\$28,908,333.33	\$24,700,000.00	1.1704	\$ 47,900,000.00	\$16,800,000.00
1993-1994	\$31,438,461.54	\$28,900,000.00	1.0878	\$ 55,700,000.00	\$16,800,000.00
1994-1995	\$27,980,769.23	\$27,900,000.00	1.0029	\$ 55,800,000.00	\$12,800,000.00
1995-1996	\$42,284,615.38	\$39,850,000.00	1.0611	\$ 73,100,000.00	\$20,600,000.00
1996-1997	\$51,376,923.08	\$48,350,000.00	1.0626	\$ 80,100,000.00	\$25,100,000.00
1997-1998	\$50,726,923.08	\$46,850,000.00	1.0828	\$ 80,100,000.00	\$25,100,000.00
1998-1999	\$52,722,222.22	\$51,900,000.00	1.0158	\$ 84,200,000.00	\$24,100,000.00
1999-2000	\$60,614,285.71	\$58,700,000.00	1.0326	\$ 97,600,000.00	\$40,400,000.00
2000-2001	\$63,533,333.33	\$61,500,000.00	1.0331	\$103,000,000.00	\$39,000,000.00
2001-2002	\$67,700,000.00	\$61,000,000.00	1.1098	\$114,000,000.00	\$42,000,000.00

Another slight anomaly occurs in the 1997-1998 year. However, this anomaly does not have quite a conclusive reason. There was nothing outside of the league that would make revenues halt, instead of climb. In fact, it was during this time that there was a big boom with the internet and a lot of money being made. This would make one think that the revenues would go up, but according to the average and the median, they went

down at this time. One possible explanation for this would be the retirement of Mario Lemieux. One of the game's very elite players, he retired the season prior, meaning that the game had one less "big-ticket" draw. Conversely, I do not think that his retirement alone could have caused the average revenue of the whole league to drop. One thing that is very interesting to note, is the fact that this anomaly follows the most competitively balanced season over the entire sample, the 1996-1997 season.

The revenue maximum and minimum follow closely the same trend as their payroll counterparts. Nevertheless, there are some things that should be pointed out. The maximum revenue increased faster than the average 9 out of 12 or 75% of the time. At the same time, the minimum climbed faster than the average only 6 out of 12, or only half the time. Again, the disparity in the revenues of the franchises is rising. If the large revenue teams are able to gain more revenue, they are able to sign more and/or better players that will most likely help them in the upcoming seasons. If these teams become even better, it is probable that they will gain more revenues in the future seasons. This could be a cycle with no end, leading a select few teams to be competitive every season and leaving most of the others unviable. This leads directly into one of the main parts of the paper. That is, can the franchises that are not large-market (or rich) be competitive in a league that contains some very powerful teams? This question will be examined later on in the paper.

### **III - Potential Problems**

While problems have already been examined with data, there are some problems that can exist outside of any errors related to data collection. It is worthwhile to examine these discrepancies in order to allow for accommodation.

One problem possibly arising is due to having only 13 years worth of data. Considering that the NHL has existed for more than 75 years, the sample size could be seen as too small, however it is still considerable as compared to other studies on this topic. To use a time series data set that covers only about 20% of the entire time period may lead to the omission of cyclical trends. Take, for example, someone trying to analyze an economy using the Gross Domestic Product (GDP) of a particular country. If

this individual uses only 5 years worth of data it could lead to an incomplete analysis. If only the downward or upward sloping portion of the business cycle is captured, it could lead to the conclusion that the economy has an ever decreasing (or increasing) Gross Domestic Product. This would, in fact, be an error in analysis. A business cycle follows a cyclical trend, which means it has periods of decline followed by periods of incline.

Competitive balance in the National Hockey League could follow a similar trend. There could be heightened times of imbalance followed by times of a very balanced league. Thirteen years, while it may only be 20% of the entire population and I might only be capturing one part of a cycle, is still a significant time period and is considered to be relatively long when examining a time trend.

There is also the problem of unbalanced schedules. In any of the major North American sports, teams do not play each other team the same number of times. This is evident in the NHL, but even more evident in Major League Baseball (MLB) and the National Football League (NFL). In baseball there are two separate leagues: the American and the National. Teams in separate leagues play each other very seldom during the regular season compared to the intra-league teams. In fact, they play around 20 games per year against teams of the other league. These are so-called inter-league games that just started in the past decade. Before this, teams from opposite leagues would never see each other until the World Series (the final round of playoffs in MLB). Even now, with only 20 out of 162 games being played against the other league, there is still a data error that could occur in some measures of competitive balance. If one league was made up completely of large-market teams while the other league was entirely composed of small-market teams, the two leagues themselves may seem very competitive during the regular season. However, during the World Series, there may be no comparison between the representative teams from the two leagues. It could be that the two leagues are totally mismatched with one being ten times better than the other, but this would not be picked up in most of the different metrics used to measure competitive balance.

The NFL has an even more unbalanced schedule. The schedule for the upcoming season is devised after the completion of the previous season. If a team does poorly during the previous season, it will play proportionately more games against other teams

in the same situation. Theoretically this means that the best of the top teams and the best of the poor teams from the previous year make it to the playoffs or have a higher winning percentage in the current year. It also makes some teams look better than they actually are, while making other teams look worse.

The NHL does not have the same amount of imbalance in its schedule; one team will play every other team at least once. Teams will play other teams in their own division most often, then play teams in their conference the second most and teams in the other conference the least. This could mean that there is one extremely strong division, but the weakest team in this so-called strong division may not obtain as many points as the strongest team in a very weak division, even though the former arguably performs much better. This type of scheduling can give the illusion of having a very balanced league, when in fact there is less balance than what is measured.

While there are possibly many more small measurement problems, the last one to be considered here is the introduction of expansion teams into a league. As before, an expansion team is a new team introduced into the league. This team must draft its players either from the NHL entry draft (a draft of amateur players) or from other teams (a franchise can only draft unprotected players from other teams, these players usually not the being of star calibre). Otherwise a franchise must buy its players.

The entry draft usually consists of very young talent who most often, need time to develop at the minor league level. Therefore it is often several years before they will be considered to join the NHL. Drafts from other teams in the league do not often allow an expansion team to acquire an abundance of talent either. Every other team is allowed to protect a certain number of existing players and permit the expansion team to take any player not on their protected list. Typically an existing team will not leave a well loved or high calibre player to go unprotected. This type of system further decreases the amount of success the new team will obtain in the following couple of years. The final way an expansion team can obtain talent is to buy (or trade) a player from another team. Unless the new team has an owner with extremely deep pockets, this is unlikely to happen. New owners are a little more hesitant to agree to lucrative contracts, because they have no realized revenues and are not exactly sure what those realized revenues are going to be.

For the above reasons, expansion teams – at least for the first few years – are not likely to be very successful. They will probably be extremely weak, and therefore make the balance of the league appear to be quite lacking. This could be very untrue because the rest of the league could be very balanced, while one team is not competitive at all. There are many ways in which a competitive balance metric could be skewed. Above are some of the main reasons.

#### **IV - Data on Competitive Balance**

##### ***Single Variable Analysis***

Fort (2003) listed several different types of competitive balance (or imbalance). These are not completely exclusive, but do have separate relevance. The four types are revenue imbalance, payroll imbalance, winning percentage imbalance and championship imbalance. Most likely, there is considerable correlation between payroll and revenue imbalance. As well, there is a good theoretical chance that winning percentage imbalance and championship imbalance are highly correlated as well. To see this, take for example a league where only one team from each division gets into the playoffs or championship round. If there is considerable winning percentage imbalance, especially one with a time trend (to be explained later), then the teams that consistently gain entry to the playoffs will most probably win more championships. While all of the possible correlations will be examined, the main ones of interest are the correlations between revenue or payroll and that between winning percentage and championship performance. If small market teams are winning, then the outcomes of the league are not determined by money. It would be possible to conclude that all teams – large and small market – are balanced when it comes to measuring the uncertainty of winning. This is seen as the ultimate goal in a professional sports league.

To start the analysis, I will examine the exclusive categories to see what, if any, imbalances exist. I will first examine revenue, followed by payroll, winning percentage and the championship. I will then examine the relationships between these variables.

### *Revenue Imbalance*

It is apparent that revenue imbalance currently exists in the National Hockey League. When there are a few teams earning over \$100 million in revenue, while others earn less than \$45 million, it can be fairly stated that there is a considerable amount of imbalance. This type of imbalance is not a recent phenomenon either. Over the course of the study, there has been a considerable increase in the degree of inequity, but disparities were found long ago as well.

Table 3 displays the standard deviations and the ranges (defined as the minimum subtracted from the maximum) of league revenue over the course of the thirteen years under examination. I have used two different measures of imbalance here to show conclusively that revenue imbalance does exist. The standard deviation and the range are divided by the average in the league that year. This is done to show relative and not absolute differences. As an example of why this is done, consider two teams at the beginning of the sample. One has revenue of \$10 million, while the other has revenue of \$5 million. The first team is twice as well off at this time. Now, at the end of the sample, the first team earns \$50 million in revenue while the second team earns \$25 million. In absolute terms the second team is now a lot worse off, because the nominal difference between them and the first team has increased by \$20 million (difference between 25 and 5 million). However, they could be just as well (or poorly) off as they were at the beginning of the sample (half the revenues of the first team). For this reason the standard deviations and ranges were divided by the average revenue of that year to normalize these differences and show the relative and not absolute difference.

From Table 3 it is apparent that the standard deviation of revenues has increased significantly relative to the 1989-1990 season. In the recent years, however, there has been a noteworthy decrease in revenue imbalance. For the six years up to but not including the last year of the study, there was a downward trend. The last year saw another up-turn. Due to the lack of data, we can not see if this is a cyclical trend or just an anomaly. The downturn that has occurred can be attributed to the increase in revenue sharing in the NHL. In an effort to keep the same teams in the league without relocating any of them, the Board of Directors has agreed to have each team share some of their revenues in order to allow teams to be more competitive.

Table 3		
Revenue Imbalance Measure		
Year	Standard Deviation	Range
1989-1990	0.23	0.84
1990-1991	0.23	0.84
1991-1992	0.31	1.29
1992-1993	0.34	1.08
1993-1994	0.35	1.24
1994-1995	0.37	1.54
1995-1996	0.33	1.24
1996-1997	0.28	1.07
1997-1998	0.28	1.08
1998-1999	0.28	1.14
1999-2000	0.27	0.94
2000-2001	0.26	1.01
2001-2002	0.29	1.06
Average	0.29	1.11

As mentioned before, the imbalance is not a recent phenomenon. Jones (1969) found that there was considerable difference in attendance rates even when there were only six teams in the league. At a time when the money brought in at the gate was basically the only money generated for a franchise, there were considerable differences in attendance records. ``[s]ince 1948-49 Montreal has never played to less than 100 per cent of seating capacity . . . since 1946-47 Toronto has never played to less than 100 per cent of seating capacity.`` At this same time, Jones found that Chicago was playing to a seating capacity of sometimes less than 25%. This would generate substantial differences in revenues attained by each team. At the time, there were barely any revenues from broadcasting corporations or merchandise sales. During this period, Toronto and Montreal were consistently at the top of the league, while Chicago was consistently at the bottom. What remains to be seen is if lack of revenue caused a lack of on-ice competitiveness or vice versa. This issue will be considered later on.

#### *Payroll Imbalance*

As will be illustrated later, there is a strong correlation between the payroll and revenue of a team. Simply put, if the revenue of a team goes up, they are able to afford more talented (and more costly) players. These players tend to generate more revenue and therefore the franchise can go out and get even more star players. This whole relationship works in a circle.

From the data in Table 4, payrolls have increased steadily with revenues. In fact, there has been a greater rise in payroll inequity than there was in revenue disparity. For the reasons mentioned in the revenue imbalance section, the standard deviation and range have been divided by the average payroll in that particular year.

Since the 1989-1990 season, there has been a fairly consistent upward rising trend in payroll disparity, however, and it has gone up fairly high. The inequity in payroll has increased by approximately 225%. This is due to a couple of teams generating very high revenues year after year, while there are expansion teams with low revenues and low payrolls. As was stated before, it can be okay to have teams with very different payrolls as long as there is no or very little correlation between payroll and outcomes.

<b>Table 4</b>		
<b>Payroll Imbalance Measure</b>		
<b>Year</b>	<b>Standard Deviation</b>	<b>Range</b>
1989-1990	0.17	0.75
1990-1991	0.24	0.81
1991-1992	0.23	0.82
1992-1993	0.21	0.80
1993-1994	0.28	1.02
1994-1995	0.27	0.94
1995-1996	0.24	1.04
1996-1997	0.27	1.02
1997-1998	0.28	1.00
1998-1999	0.25	1.16
1999-2000	0.32	1.36
2000-2001	0.31	1.13
2001-2002	0.37	1.34
Average	0.27	1.01

The range of payrolls has increased significantly as well. It has in fact increased slightly less than 200%. This also means that there is a greater dispersion now than there was even a decade ago. This problem has manifested itself fairly recently. With the loosening of free agent status or the reserve clause, players earn closer to their marginal revenue product than to their opportunity cost. Broadly speaking, players are taking rents away from the owners.

While there have been no abrupt changes to the rules and regulations governing free agent status, they have been phased in over the past ten to twenty years. A few decades ago, a player would initially sign with a professional team and would stay with that team, unless he was traded or picked up on waivers. Free agency was extremely unusual at this time. With the loosening of regulations and the greater strength of the players union (NHLPA), players were allowed to negotiate better contracts with their owners to allow them more mobility. In other words, it has been increasingly easier to achieve free agency. Obtaining complete free agent status means that a player is free to move wherever he wants to go. Save few exceptions, this means that players go wherever they are offered the most lucrative contract, i.e. where they are valued the most. Where they are valued the most is usually a franchise with very high earnings, because these teams can afford to pay them the most, thus creating an imbalance of talent on each team.

The generalization of free agent status has been proven to increase the salaries of players in the other major North American sports as well.

#### *Regular Play Imbalance: Within a Season*

We are now getting into what most fans of sport deem to be most interesting question. How does all of this affect outcomes in the league? A lot of people only seem interested in what is happening to their team. Is their team competitive with the rest of the league? After all, a fan will not care if his team is the lowest in terms of payroll and revenue unless this statistic carries over to the team's on-ice performance. As mentioned previously, all of the correlations of the variables are going to be analyzed and discussed later. Here I will discuss the disparities in the regular season.

The standard deviation and range of winning percentage have been analyzed in several journal articles. Not only hockey, but all of the major North American sports and the European soccer leagues have been investigated as well for ongoing disparities.

<b>Winning Percentage Imbalance Measure</b>		
<b>Year</b>	<b>Standard Deviation</b>	<b>Range</b>
1989-1990	9.425%	43.750%
1990-1991	10.441%	37.500%
1991-1992	9.594%	41.250%
1992-1993	14.514%	56.548%
1993-1994	10.228%	44.643%
1994-1995	11.118%	48.958%
1995-1996	11.553%	54.878%
1996-1997	7.793%	28.049%
1997-1998	9.618%	39.634%
1998-1999	9.646%	40.854%
1999-2000	9.891%	43.568%
2000-2001	10.260%	38.342%
2001-2002	8.728%	36.033%
Average	10.216%	42.615%

In their 1992 book *Pay Dirt: The Business of Professional Team Sports*, Quirk and Fort following Noll (1998) and Scully (1989) use a more normalized metric. They divide the actual standard deviation by an idealized standard deviation. The idealized standard deviation is one that the authors say would be reached if every team in the league was of equal playing strength. This is equal to  $0.5/\sqrt{N}$ , where N is the number of games played in a season. In this analysis, this will be good because the number of games in a season changes three times. I have computed this calculation for the sample period as well as included the work done by Quirk & Fort, which consists of the average over a decade for the NHL. This data is listed in Table 6.

Over the course of my sample period there has been considerable fluctuation in this statistic. The most competitive the league has been was in the 1996-1997 season, while the least competitive season was 1992-1993. It can not be said however, that competitive balance has increased or decreased over the time period. The same thing can be said for the decade by decade analysis. While it does appear that the NHL has become

more competitive over the past few decades, the league is still less balanced than what it was for the first 40 years of its existence. It is a step in the right direction.

<b>Actual Standard Deviation / Idealized Standard Deviation Ratio</b>			
<b>Year</b>	<b>Ratio</b>	<b>Decade</b>	<b>Ratio</b>
1989-1990**	1.686	1910-1919*	1.71
1990-1991**	1.868	1920-1929*	1.61
1991-1992**	1.716	1930-1039*	1.53
1992-1993**	2.660	1940-1949*	1.69
1993-1994**	1.875	1950-1959*	1.88
1994-1995**	1.541	1960-1969*	1.90
1995-1996**	2.092	1970-1979*	2.54
1996-1997**	1.411	1980-1989*	1.91
1997-1998**	1.742	1990-1999**	1.84
1998-1999**	1.747		
1999-2000**	1.791		
2000-2001**	1.858		
2001-2002**	1.581		
2002-2003**	1.592		

\* denotes analysis from Quirk & Fort  
 \*\* denotes analysis from this paper

From Tables 5 & 6, it is clear to see that there is some disparity in the NHL. However, this disparity has not changed much over the course of the study. In fact, for the period of the full data set, the standard deviation and the range of winning percentages have dropped. This has not been a continuing trend, but one that is worth noting nonetheless. In the first year, the standard deviation and range were approximately 9.5% and 43.5% respectively. These numbers seemed to rise during the early to mid 1990's and fall again in the mid to late 90's and the early part of the new millennium.

The rise in the disparity of outcomes can be partly attributed to the arrival of nine expansion teams in the years studied. As previously mentioned, expansion teams have very limited talent in their first couple of years. As well, they do not have a solid team structure. The ownership of a team usually expects to take a few years to build a team that plays well together. The first couple of years are the so-called building years, where a manager tries to bring together a group of players that compliment each other. An expansion team which starts off with players who have most likely never played together

before, will be hard pressed to do well until they have had time to add to and change their roster so as to have players that work favourably with one another.

The majority of the double-digit standard deviations can be correlated with the arrival of expansion teams in those years. From the 92-93 season to the 93-94 season, four new teams were added to the NHL. This can explain the high amount of inequity in the early to mid 1990's. As well, in the 2000-2001 season, another two teams were added to the roster. Again in this year, we saw a standard deviation of over 10%.

The range is also well correlated with the introduction of expansion teams. The range was consistently high from the 92-93 season to the 95-96 season. This is just before a drop by almost half in the 96-97 season.

The fact that expansion teams can lead to imbalance is demonstrated in Quirk & Fort (1992). They analyzed the winning percentages of expansion teams over their first several years and found that it took, on average, for a team in the NHL 6.3 years to reach the .500 level. They also found that "[t]he extremely weak W/L (won/loss) records of expansion teams are matched by unusually strong W/L records for existing teams during a period of expansion, both of which act to increase the dispersion of W/L percentages for a league." (Quirk & Fort, 1992, p.250)

Although expansion teams are an explanation for the recent fluctuations in regular season outcomes, they are not the only one. There existed significant disparity even when there were the original 6 teams. This was caused by an imbalance of talent. Even before the original expansion in 1967, when the player mobility rules were a lot more strict, and the league tried harder to even out the talent pool, there existed some inequality.

#### *Regular Season Imbalance: Through time.*

Another question to be asked is this; does there exist a time trend in regular season performance? To see this, I used two different methods of measuring the correlation between the current year and previous years. The first method was to use the rank of each team in each year, while the second method used the winning percentage of each year. Every year was then correlated with the previous year and the year before that and so on and so forth. The rank was used because it is how well a team does relative to the rest of the league that really counts and not how many points a team obtains.

Winning percentage was used to see the comparison between this paper and other papers. Teams that have relocated have been considered as the same team in this analysis.

When using a correlation coefficient, one has a number bounded by  $-1$  and  $+1$ . A coefficient of  $+1$  means that there is complete correlation between this year and previous years. Anywhere between  $0$  and  $+1$  means that there is some positive correlation between the two years analyzed, with numbers closer to one signifying stronger correlation. A negative coefficient indicates that how one team placed this year is inversely related to how the team placed in the other analyzed year.

The data, listed in Appendix A, shows that there is significant positive correlation when analyzing a one year difference. This means that how a team placed last year is highly related to how it places this year. The positive correlation remains until one goes five years back. This is when the negative correlation coefficients are first seen. After five-year lags, inverse relations become noticeable. What this means is that if a team did well five years ago or more, it is most likely not to do well this year. Also, if a team did not do well five years ago, it will probably do well this year. This is a good sign of balance in the league. This shows that some teams are not consistently winning while others are losing consistently year after year. A team's fate can change and there is a so-called changing of the guard in the National Hockey League.

Leifer did somewhat the same analysis in his book entitled *Making the Majors*. He used points to demonstrate the time trend however, and took the average lag trend over certain periods: the period before expansion 1917-1969, and the time since expansion 1969-1989. Leifer's analysis was different in some regards. For the first five years his analysis was the same, meaning that there were positive correlations 5 years back. In his study though, the positive correlation continued up until 8 years back where he stopped his analysis. The probable explanation for the difference between my analysis and his analysis is the time period and the dropping of the reservation clause. His analysis terminated in 1989 right around the time when the player reserve clause was weakened. With the subsiding of the reserve clause it is easier for a team to buy star players and increase its performance.

In his book, *Avantage numerique: l'argent et la Ligue nationale de hockey*, Lavoie obtained the same results for a fairly similar time period. He used the same

method as Leifer had over the 1987-1995 period. He found that there was positive correlation up until five years back. The correlations were then negative for the fifth and sixth year, then turned positive for the later years.

Using winning percentage obtains rather different results than using rank correlations. Using this method, positive correlation exists throughout every lagged correlation. From the one year through to the ten-year lag, the average correlation coefficient stays above 0.4. This result portrays two things. The first is that there exists a significant relationship between a team's current standing and their previous outcome. It also exhibits the fact that this sample contains different data from the data used in the Leifer and Lavoie analyses. The sample set used in this paper illustrates that how a team does this year is linked to how it has done in the past.

### *Championship Imbalance*

Another isolated feature of comparing imbalance is to compare the championship imbalance. This measure is used to see if inequality exists amongst the post-regular season. Are there very few teams who win the ultimate prize year after year, or are there a number of teams that have equal chance of winning the Stanley Cup? Again, for the league to create balanced, one would want a large number of teams to win very few times. Ideally, if there are thirty teams in the league (as there are now in the NHL, one would want every team to win at least once every thirty years.

While we know that this will never happen, it would be nice for the fans of every team to know that their team is going to win as often as every other team.

There are at least three possible ways to measure championship imbalance: the concentration of winners (see how many times each team has won the Stanley Cup), the number of years a team goes, on average, between winning, and the number of games played in a post-regular season.

The first two metrics are used to compare just the balance of the championship winners, whereas the third metric is used to compare the balance of all the teams in the post-regular season. For this reason, the first two and the third metric will be analyzed separately.

When comparing the concentration of championships and the average years metrics, I will use my sample, as well as include the history of the entire league, to see

whether or not there is a rise in imbalance. One has to be careful when trying to compare these two metrics. The two measurements are equal when every team has been in the league the same amount of years; however, as mentioned previously, nine teams have been added in the thirteen years of the study. As well, there has been only six teams in the league since its inception. For these reasons, more emphasis will be placed on the metric that uses the average number of years between championships.

When using the number of games played in a playoff series, I will only use the normal sample (my sample). In trying to use the entire league history, there are some problems that could occur, the main one being that there have been different playoff systems in effect since the start of the Stanley Cup. Up to 1967, there were two types of formats used. The first had the top two teams competing for the Cup, and the second saw the top four teams would be involved in a two round elimination with the winner crowned champion. Using the first method the maximum number of games that could be played was seven, whereas in the second format, the minimum number of games that could be played would be 12 while the maximum would be 21. It is impossible to compare the two. As well, between 1967 and the 1989-1990 season, there were variations on the number of teams that would enter the playoff round as well as the number of games played in each round. This again would make a comparison of the entire history impossible.

The data on the number of Stanley Cups that a team has won is presented in Table 7. While analyzing this, if a city had relocated – such as the relocation from Quebec to Colorado – the cups were not amalgamated. However, this made no difference, since no relocated team ever won before their relocation. Also, in the case of Montreal, they have won 22 cups under the name of the Montreal Canadiens and 1 cup under the name of the Maroons. This was in 1935. As well, the Ottawa Senators are listed as winning the Stanley Cup once. This is not the new era Senators, in fact the team folded and then joined the league again in the 1992-1993 season. They actually won the Stanley Cup in the 1926-1927 season. The entire Stanley Cup history is counted from 1926 forward. Even though the Stanley Cup has been around for more than 100 years, it wasn't until this year that the NHL took complete control of the Cup. Before this time, the Stanley Cup was competed for by amateur and professional teams alike.

From the table, the last 13 years have been fairly competitive. Eight teams have won the cup over the course of the study. No team has won it more than 3 times and no team has won it more than twice in a row. This is in stark contrast to the data shown for the league since 1926. There is a concentration of championships with the Original six

<b>Table 7</b>	
<b>Number of Cups per Team</b>	
<b>Sample</b>	<b>1989-1990-present</b>
Colorado Avalanche	2
New Jersey Devils	3
Dallas Stars	1
Detroit Red Wings	3
New York Rangers	1
Montreal Canadiens	1
Pittsburgh Penguins	2
Edmonton Oilers	1
<b>Entire Sample</b>	<b>1926-present</b>
Colorado Avalanche	2
New Jersey Devils	3
Dallas Stars	1
Detroit Red Wings	10
New York Rangers	4
Montreal Canadiens	23
Pittsburgh Penguins	2
Edmonton Oilers	5
Calgary Flames	1
New York Islanders	4
Philadelphia Flyers	2
Boston Bruins	5
Toronto Maple Leafs	11
Chicago Blackhawks	3
Ottawa Senators	1

teams. In total they have won 56 of the entire 75. This is significantly greater than 66%. However, as mentioned before, this is not the best metric to use, because some teams have been in the league for the full 75 years, while others have only been valid for just over 20.

The average number of years between cups would be a more useful measure to see if there is competition. This data is presented in Table 8. In this table, relocation of a franchise is more important. The Colorado Avalanche have technically only been in the league since the 1995-1996 season. However, they were known as the Quebec Nordiques before this and retained most of the players when they moved. It is for this reason that they have been included for all thirteen years of the study. This is the same for the Dallas Stars. They were formerly known as the Minnesota North Stars.

<b>Table 8</b>	
<b>Average # of Years B/W Stanley Cups</b>	
<b>Sample</b>	<b>1989-1990-present</b>
Colorado Avalanche	6.5
New Jersey Devils	4.333333333
Dallas Stars	13
Detroit Red Wings	4.333333333
New York Rangers	13
Montreal Canadiens	13
Pittsburgh Penguins	6.5
Edmonton Oilers	13
<b>Entire Sample</b>	<b>1926-present</b>
Colorado Avalanche	11.5
New Jersey Devils	9.666666667
Dallas Stars	35
Detroit Red Wings	7.5
New York Rangers	18.75
Montreal Canadiens	3.260869565
Pittsburgh Penguins	17.5
Edmonton Oilers	4.6
Calgary Flames	30
New York Islanders	7
Philadelphia Flyers	17.5
Boston Bruins	15
Toronto Maple Leafs	6.818181818
Chicago Blackhawks	25
Ottawa Senators	0

Table 8 also shows that there is a fair amount of imbalance in the league. Only half of the teams in the league have ever won a Stanley Cup. There have been teams in the league for over 35 years who have never won a championship (i.e. St. Louis Blues), whereas the Montreal Canadiens win a Stanley Cup every 3.25 years. Over the course of the study however, the Canadiens have only won the Cup once. On the other hand, there

are the Detroit Red Wings, who, over the course of 75 years have won the cup every 7.5 years, but, over the last 13 have won the Cup 3 times for an average of 1 title every 4 and one-third years.

There is still a lot less balance than what it optimal, that is if we were to have a new champion every year. This level of optimality can not be expected, but it would be better if we saw more teams win, or even be in competition for the championship.

While the above two measures compare the competition for the championship, they say nothing about the competition that lead up to the finals. Richardson (2000) suggests to look at the number of playoff games played. There is a total of 16 teams who

<b>Table 9</b>	
<b>Number of Playoff Games/Season</b>	
<b>Year</b>	<b>Playoff Games</b>
1989-1990	85
1990-1991	92
1991-1992	86
1992-1993	85
1993-1994	90
1994-1995	81
1995-1996	86
1996-1997	95
1997-1998	82
1998-1999	82
1999-2000	86
2000-2001	83
2001-2002	86
2002-2003	90
<b>Average</b>	<b>86.36</b>

make the playoffs who have to endure 4 rounds to win. Each round has a minimum of 4 games and a maximum of 7. The total number of games thus varies from a minimum of 60, where one team wins the first four games of each of round, to a maximum of 105 where all rounds go to 7 games. A small number of playoff games is associated with unequal playing strength and a lack of competitive balance. (Richardson 2000). Table 9 presents the data of playoff games per year.

While there has been a lot of variation in the number of playoff games, it has mainly consisted of up and down fluctuations without a constant increase or decrease.

The lowest amount of games played was 81 whereas the maximum number of games played was 95. One problem with this measure is that it doesn't state who is in the playoffs. It is no wonder that the playoffs exhibit more balance, since the same teams (usually expansion teams) tend to miss the playoffs year after year (Richardson 2000). There could be the same 16 teams that enter the playoffs every year, while the others never make it. This is however untrue. While there may be some teams (i.e. Detroit Red Wings and Colorado Avalanche) that make the playoffs every year, there is a lot of variation in terms of which teams make the playoffs.

One analysis that I didn't include in this section of the paper is the Gini Coefficient. The Gini Coefficient is often used in measuring inequity and disparities namely in welfare economics. It measures on a cumulative basis how much each percent of the population owns. This measure is derived from the Lorenz curve.

Recently, sports economists have used Gini Coefficients to determine the measure of imbalance in a league. However, according to Fort & Utt, this is not a metric for analyzing inequity in sports. In their article *Pitfalls to Measuring Competitive Balance with Gini Coefficients* (2002), they state that there are a number of setbacks to using this type of measurement.

One main setback is the zero sum game of sports. "At most, one team can win all of the games it plays, a fraction of the total number of games played in the league." (Fort & Utt, 2002) This fact requires a reworking of certain data in order to accommodate the problem. Other difficulties in determination, such as "unbalanced schedules (teams do not play the same number of games against all opponents), league expansion, and interleague play, present additional challenges." (Fort & Utt, 2002)

For the above reasons Gini Coefficients will not be used in this paper, due to the computational difficulties that Fort & Utt have mentioned.

### ***Inter-Variable Correlations of Competitive Balance***

Up to now, I have only looked at competitive balance in terms of one variable. I am now going to look at the correlations and dependencies of one variable (i.e. payroll) on another variable (i.e. winning percentage). This is to see if there are any relationships between the two.

This, I believe is the most interesting part, because there could be a substantial amount of imbalance in some variables, although these could be arbitrary. The ideal goal of a league is to have complete balance. Put another way, there should be a high degree of uncertainty in the outcome of a game. There is very little interest shown in teams that are positively going to lose. While this is hard to prove at a season level – no team has gone winless in our sample – it can be portrayed in a different fashion. Take any game where there is a sell-out crowd. This usually means that one team has a huge following or that both teams are evenly matched. If one team is running away with the game in the third period (the last period of regulation time in an NHL game) the attendance dwindles down quite noticeably even 20 minutes before the game is over. This shows a lack of interest from people who have already paid to watch the game. If its outcome was known for certain before the game started a lot of people probably wouldn't even buy tickets to the game. On the contrary, when there is a closely matched game, say a tie, going into the last few minutes, you will hardly notice an empty seat in the arena. This shows that uncertainty increases fan interest.

#### *Winning Percentage and Playoff Games*

The first area of inter-variable interest is the correlation between winning percentage in the regular season and playoff games. This displays a balance in the top part of the league. Playoff games played, as opposed to playoff winning percentage is used because winning percentage can be misleading. For example, take one team that wins their first three rounds in seven games. Each round the team wins four and lose three, giving it twelve wins and nine losses. In the last round, if the team loses all four games straight, it will now have only twelve wins, but thirteen losses. This is a winning percentage that is less than 50%, deemed not to be very good. Moreover, another team could win their first round in four games and lose their second round in seven games. Overall they have a total of seven wins and four losses; a winning percentage of almost 66%. While the latter team has a higher winning percentage, it obviously didn't make it that far. The first team has played more games, had more wins, and has been closer to winning the Stanley Cup, but it will appear to be less successful if the statistics used are winning percentage alone.

Because it is more or less the top 16 teams in the league (according to points) that make it into the post-regular season round, this statistic will only determine if there is balance in the upper portion of the league. It is not exactly the top 16 teams in the league, because the top 8 teams in each conference qualify. There could, theoretically, be some teams in one conference who have a better regular season record than one of the top 8 teams in the other conference, however they do not make the playoffs because they are not one of the top 8 teams in their conference. More or less though, it is the top 16 teams in the league.

I have done this calculation for every year to see if there is a trend that has developed over the past 13 years. As well, I have done an overall calculation to see if there is a strong link between the two variables.

The correlations are shown in Table 10.

<b>Table 10</b>	
<b>Playoff / Winning Percentage Correlation</b>	
<b>Year</b>	<b>Correlation Coefficient</b>
1989-1990	0.32
1990-1991	0.04
1991-1992	0.31
1992-1993	-0.14
1993-1994	0.29
1994-1995	0.24
1995-1996	0.65
1996-1997	0.51
1997-1998	0.28
1998-1999	0.49
1999-2000	0.30
2000-2001	0.47
2001-2002	0.36
<b>Average</b>	<b>0.29</b>

From this table, there is no trend that exists over the complete sample. However, during the first part of the sample (1989-1990 to 1994-1995) there is little or no correlation, while in the later part of the sample, there exists fairly high correlation. The year-by-year statistic shows fluctuations with no more than two years going in the same direction. Twelve out of the 13 years have positive correlation with only the 1992-1993 season having a negative coefficient. As well, the overall coefficient is positive. These two

facts lead to the conclusion that the better that a team performs in the regular season, the better it will do in the post-season. Though there is not a perfect correlation between the two, the data shows there is a positive, relatively strong relationship.

It has already been noted that while there is a significant relationship between payroll and revenue (shown in Table 11), however, it is not a 1 to 1 relationship. Over the last ten years, this correlation has become more significant. In the beginning of the sample there was even one year where the correlation coefficient was negative. The high revenue teams would have lower payrolls than the low revenue teams.

<b>Table 11</b>	
<b>Payroll / Revenue Correlation</b>	
<b>Year</b>	<b>Correlation Coefficient</b>
1989-1990	0.45
1990-1991	-0.01
1991-1992	0.55
1992-1993	0.66
1993-1994	0.31
1994-1995	0.46
1995-1996	0.64
1996-1997	0.55
1997-1998	0.48
1998-1999	0.59
1999-2000	0.74
2000-2001	0.67
2001-2002	0.70
Average	0.52

While this only happened in one year, it is still worth noting that the dependency of one variable on the other is more of a recent phenomenon. For the reason that these two variables are not perfectly correlated I am going to complete the money variable – outcome variable analysis using both revenue and payroll.

The most interesting relationships are the ones between money variables and outcome variables. These will determine if the outcomes of the season are determined before the season even begins. More specifically, can one look at the payrolls or revenue of a team and determine where that team will finish in the standings or in the playoffs? The first of these relationships that I will look at is the relationships between the money variables and the regular season outcome. For this analysis, I am going to show first the straight correlation coefficient and then a regression analysis. Because winning percentage is a continuous random variable, I am able to use a simple OLS analysis.

Using these two different types of statistics will show two interesting details. The correlation coefficient will show how the two variables are directly related. If the coefficient is close to +1 it means that the high payroll or high revenue teams are always in the top of the league. A coefficient of 0 means that there is no direct relationship between the money variable and the outcome variable. A high revenue or low revenue team has equal chances of finishing first or last in the league. A coefficient of negative proportions means that the more revenue or payroll one obtains, the lower in rank a team will finish. The optimal from a league perspective and from that of fans of hockey in general, is a coefficient of zero.

The figures for the relationship between payroll and regular season winning percentage are given in Table 12.

<b>Table 12</b>	
<b>Payroll / Winning Percentage Correlation</b>	
<b>Year</b>	<b>Correlation Coefficient</b>
1989-1990	-0.07
1990-1991	0.31
1991-1992	0.53
1992-1993	0.61
1993-1994	0.55
1994-1995	0.48
1995-1996	0.48
1996-1997	0.49
1997-1998	0.28
1998-1999	0.28
1999-2000	0.51
2000-2001	0.64
2001-2002	0.47
Average	0.43

The table shows that there is significant positive correlation between a team's payroll and its standing at the end of the regular season. The first season, can be seen as an anomaly or an outlier. It has a slight negative correlation, but really close to zero. For the rest of the sample, the coefficient is positive and more than 50% of the time, it is either around or above 0.5. This indicates that the outcome of the season is partially determined by the payroll of a team.

To see how payroll actually helps out a team, it is worthwhile to use regression analysis. I will use a straight univariate linear regression analysis with regular season winning percentage as the dependant variable, while having payroll as the independent or explanatory variable. Therefore the regression equation will look like the following

$$Y_i = x'B + \varepsilon_i$$

or more specifically,

$$\text{Regular Season Outcome} = B_0 + B_1(\text{Payroll})_{j,i} + \text{random error}_i$$

For year  $j=1, \dots, n$  and team  $i=1, \dots, m$

I presume all of the assumptions of the Classical Linear Normal Model (CLNM). This makes the analysis fairly simplified. The output for this analysis is located in Table 13. For this analysis, the payroll has been manipulated by an inflation index. To do this, the last year (2001-2002) was used as a base year. From here, inflation rates were calculated in order to figure how much a payroll in 1989 is valued in 2001-2002 dollars.

This regression analysis explains how much an increase in of one unit of payroll affects the regular season winning percentage of that team. I have divided every team's payroll and revenue by one million. By doing this, one unit of payroll or revenue, is now one million dollars. One thing to note is that the R-squared coefficient is approximately 0.2, which is considered decent in cross-sectional analyses. This is to be expected because there are a lot of other factors that determine outcome. A lower R-squared such as this, means that the explanatory power of this model is not very large. This is exhibited by the high intercept coefficient. An intercept coefficient of 0.33 means that a team with a payroll of zero, wins 33% of the time. This can not be true because a team with no payroll does not exist.

All other things considered, the remainder of the regression evaluation is straightforward. The slope coefficient (the coefficient corresponding to payroll), was calculated to be 0.005. This means that for every extra million that a team spends on its players, their winning percentage will increase by 0.5 percentage points. This can be seen as significant over the course of a season. One half of one percent can be the difference between making the playoffs and not making the playoffs. It can also mean the difference between home ice advantage throughout the playoffs and never having that

advantage. This would not only help with the support of a team, but it will help with revenues for the next season. As another example, a 10-million increase in the payroll of a team could push the performance level of that team from say .500 to .550.

<b>Table 13</b>				
<b>Payroll Regression Analysis</b>				
<b>Year</b>	<b>Estimated Coefficient</b>	<b>T-Ratio</b>	<b>R-Squared</b>	
1989-1990	-0.001	-0.327	0.000	21
1990-1991	0.004	1.434	0.100	21
1991-1992	0.006	2.826	0.258	22
1992-1993	0.010	3.640	0.375	24
1993-1994	0.005	3.226	0.308	26
1994-1995	0.005	2.680	0.230	26
1995-1996	0.007	2.648	0.226	26
1996-1997	0.004	2.779	0.243	26
1997-1998	0.003	1.434	0.079	26
1998-1999	0.003	1.226	0.077	27
1999-2000	0.005	3.047	0.263	28
2000-2001	0.006	4.370	0.406	30
2001-2002	0.003	2.850	0.225	30
<b>Overall</b>	0.005	8.988	0.196	333
<b>Semi-Log Model</b>				
<b>Overall</b>	0.179	9.828	0.226	333

Also included in Table 13 is a year-by-year analysis. This is to show the evolution of the effect that payroll has on winning percentage. As is evident from the table, the payroll coefficient is fairly stable, and it is statistically significant, with a couple of exceptions. In the 1989-1990 season, a higher payroll would actually decrease your winning percentage. However, this could not be tested as statistically significant. In the rest of the sample, the coefficients range from 0.3% to 1%. This denotes that overall, an extra million in payroll can significantly increase a team's regular season performance.

With the strict winning percentage/payroll model having only a decent amount of explanatory power, I have also attempted a semi-log regression. I have manipulated the data by using the natural logarithm of the payroll data. From the regression results, this model has slightly better explanatory power, as can be seen from the higher R-squared coefficient of 0.23. The results of this model are also displayed in Table 13, as the semi-log model. The interpretation of the slope coefficient differs slightly in this analysis as

well. The coefficient of 0.18 means that a 100 percent increase in payroll will increase the regular season winning percentage by 18 percentage points. In other words, whenever payroll doubles, the winning percentage increases by 18 percentage points. For instance, with payroll moving up from 23 to 46 million, around the 2001-2002 payroll average, the winning performance would move from say .410 up to .590.

The analysis between revenue and winning percentage is carried out in the same manner. The results are, as well, much the same as the regression for payroll and winning percentage. The results for revenue and winning percentage are listed in table 14. The R-squared coefficient is again lower than desired. As well, the coefficient on the constant is nearly 36%. The similar number can be attributed to the high correlation of revenue and payroll. The slope coefficient is however, slightly smaller. It has been calculated to be around 0.25%. For every one million dollar increase in revenue, the model predicts that a team's winning percentage will increase by 0.25 percentage points. Again, this is statistically significant, but economically speaking, the coefficient is rather small. If a franchise can increase its revenues it can increase its winning percentage, but not by much. If it increases its winning percentage, it will most likely increase its revenues in the following year. Will that lead to further improvement in winning percentage?

<b>Table 14</b>			
<b>Revenue Regression Analysis</b>			
<b>Model</b>	<b>Estimated Coefficient</b>	<b>T-Ratio</b>	<b>R-Squared</b>
Original Model	0.0025	7.736	0.154
Semi-Log Model	0.135	7.753	0.1537
Lagged Regression Model	0.0008	3.1125	0.0312

To test whether there is a feedback effect between revenue and winning percentage, it is worthwhile to investigate the effect that last year's revenues have had on this year's performance. This can be observed using a lagged regression. The sample size for this regression is slightly less (only 12 years) because now the first year's revenues are not included. This analysis shows that there is little to no correlation between last year's revenue and the current year's winning percentage. A very low R-squared of 0.03 shows that the model has very weak explanatory power. The model also

displays a slope coefficient result of 0.08%; a one million dollar increase in last year's revenue will increase a team's current year winning percentage by less than one-tenth of a percentage point.

Again, trying for a model that demonstrates more explanatory power, I have manipulated the revenue by transforming it using the natural logarithm. The results are very similar to that of the logged payroll regression. The R-squared coefficient has stayed the same at just over 0.15. The slope coefficient is 13%. This says that whenever revenues double, the performance statistic moves up by 13 percentage points. In other words, if team revenues double (from say 45 to 90 millions around the 2001-2002 mean), the performance measure would move up from say .435 to .565.

One likely explanation that a higher payroll will increase a team's performance more than revenue is that payroll is more closely linked to on-ice performance. A team's payroll more or less determines what type of product is being put on the ice day in and day out. A team's revenues, on the other hand, can be totally unrelated to what type of talent the team has. In a city like New York, there are enough businesses and people who have an abundance of money that will go to games no matter how the team is doing. This will contribute to a team's revenues, but could have no impact on its payroll. Therefore, the revenues could be very high, but their payroll could be very low.

*Monetary Variables and Playoff Performance*

The last relationship that needs to be examined is the one between monetary variables and playoff performance. As mentioned before, it is hard to analyze the playoffs, because a team's winning percentage doesn't necessarily show how well the team performed in the post-regular season. This is due to the four, seven-game elimination rounds. For the above reason, winning percentage will not be used in the analysis of playoff performance. Instead, a point system, devised for this paper, will be used. The point system is listed in Table 15.

<b>Table 15</b>	
<b>Playoff Point System</b>	
<b>Round</b>	<b>Points</b>
Eliminated in First Round	1
Eliminated in Second Round	2
Eliminated in Third Round	4
Eliminated in Final Round	8

The point system was designed to be simplistically logical. For every year of the data set there are sixteen teams that make the post season. If a team does not make the playoffs, they receive zero points. If a team loses out in the first round it receive one point. Because only eight teams make it to the second round (half as many as the first round), they receive two points, because it can be seen as twice as hard to make the second round as it is the first round. A team that loses in the third round receives four points because only four teams make the third round, it is twice as hard to make this round than it was to make the second round. Following this logic through the next few steps, the finalist (loser of the final round) would receive eight points, and the champion would receive 16 points.

The first relationship that I will look at between the money variables and the playoff variable, will be the coefficient of correlation. This is done by correlating the playoff point system and the nominal payroll and revenue amounts. I have shown the year to year coefficients for payroll as well as for the overall sample. This was done for revenue as well.

Payroll correlation is listed in Table 16 while revenue correlation is listed in Table 17

<b>Table 16</b>	
<b>Payroll / Playoff Performance Correlation</b>	
<b>Year</b>	<b>Correlation Coefficient</b>
1989-1990	0.250
1990-1991	0.129
1991-1992	0.386
1992-1993	0.487
1993-1994	0.148
1994-1995	0.183
1995-1996	-0.059
1996-1997	0.398
1997-1998	0.008
1998-1999	0.296
1999-2000	0.212
2000-2001	0.373
2001-2002	0.549
Average	0.258

As shown in table 16, the correlation is not as high as one would think. While some years exhibit decent linkage, only one year has a coefficient of more than 0.5. Most of the coefficients can be seen as having some positive relationship, but, there cannot be a fully conclusive statement saying that payroll will determine outcome. This can be seen by two facts. First, in the 1995-1996 season, there was a negative coefficient meaning that the higher the team payroll, the worse the team performed during the playoffs. This is a good thing when looking at competitive balance. As was said before, a league does not want a fan to have the ability to look at team's payrolls and determine how they will place in the playoffs. Secondly, the overall average is not too high, portraying the fact that payroll is not a determining variable for playoff results.

In Table 17, the same analysis is shown between revenue and post-season outcome. The data in this table do not follow a trend either. Over the course of the sample it can not be said that the correlation between these two variables has consistently increased over time. As with payroll, there are some negative coefficients. These happen in the years 1990-1991 and 1995-1996. The coefficients tend to fluctuate quite drastically as well. Contrary to the payroll analysis, there are more years that are over the 0.5 mark. The main reason for this is the fact that the more points you get in the playoffs, the farther you made it. The farther you make it, most likely the more games you have played (not theoretically true, but true practically). The more games that you play the more revenue you are likely to get – you get the gate revenue from home games and some television revenue, if the game is televised. This is an instance of reversed causality. This fact is important for two reasons. One, it makes the negative numbers seem quite surprising. The better you do, the more revenue you should get, not the other way around; and two, it explains why the complete sample coefficient is higher for revenue than it is for payroll.

Although the complete sample relationship is shown as stronger, it still isn't very high, meaning that a team's revenue does not necessarily determine how a team will perform in the playoffs. This is what the league's administrators have as an optimal goal.

Table 17	
Revenue / Playoff Performance Correlation	
Year	Correlation Coefficient
1989-1990	0.222
1990-1991	-0.405
1991-1992	0.168
1992-1993	0.274
1993-1994	0.542
1994-1995	0.295
1995-1996	-0.134
1996-1997	0.629
1997-1998	0.504
1998-1999	0.208
1999-2000	0.264
2000-2001	0.461
2001-2002	0.596
Average	0.279

To determine the effect that revenue and payroll have on the playoff outcome Ordinary Least Squares can not be used. Because of the discrete nature of the dependant variable a probit model must be used. Furthermore, because of the ranked nature of progression through the playoffs, an ordered probit regression will be used.

There are quite a few differences between OLS and ordered probit models, so I see it fit to briefly explain some of these dissimilarities. The probit model is used for discrete dependant variables which uses probabilities that must lie between 0 and 1. The variables are translated into this range by using the cumulative distribution function  $\text{Prob}(Y_t = 1) = F_t = I_t = F(X_t'B)$ . The probit alternative for  $F$  is the distribution function

$$F(X_t'B) = \int_{-\infty}^{I_t} \frac{1}{\sqrt{2\pi}} \exp(-t^2/2) dt$$

This is one of the main differences from the standard OLS model.

The results have to be interpreted differently as well. In the Probit model the coefficients do not explain how one unit increase in the independent variable will increase the dependant variable, instead it reveals the propensity or likelihood of achieving one option over the other in the set. Applied to this paper, the coefficient tells

us how an extra million in payroll or revenue will increase the likelihood of making it to the next round of the playoffs. This is also known as the marginal effect of the independent variable. In this analysis there are five separate distinctions in playoff rounds. The first one is the difference between not making it to the playoffs and making it to the first round. The second difference is being eliminated in the first round and making it to the second round. This continues until the final or championship round. These results are surprising as compared to the rest of the results. From round to round another unit of payroll or revenue (one unit being defined as one million dollars) does not seem to add much extra propensity to making it to the next round.

Looking at revenue to start the analysis, only three out of the five rounds have a positive coefficient. A positive coefficient means that revenue will increase the likelihood that a team will go on to the next series. However, only one of these coefficient tested significant, the second round coefficient. As well, there are two rounds where revenue actually lowered the odds of making it deeper into the playoffs. These were the first and fourth rounds. The coefficients were also very small. An extra million in revenue would range in absolute value from -8% to 7%. Nonetheless, this could translate into a high likelihood of some teams making it to the next round. This is because there are teams with fifty million dollars more revenue than some of their opponents, which make some teams an almost certainty in making it to the next round.

<b>Table 18</b>				
<b>Probit Regression – Revenue / Playoff Performance</b>				
<b>Round</b>	<b>Estimated Coef.</b>	<b>T-Ratio</b>	<b>Elas. at Means</b>	<b>Aggregate Elas.</b>
Round 1	-0.079451	-0.86268	-0.43326	-0.42991
Round 2	0.022015	3.6759	1.6802	1.5728
Round 3	0.58444E-08	0.85813	0.30930	0.31305
Round 4	-0.24647E-08	0.26421	-0.13649	-0.13471
Round 5	0.072655	0.60761	-0.47959	-0.47414

Payroll has a little more influence on the outcome of some rounds of the playoffs. There are four rounds where a higher payroll increases a team's likelihood of progressing it to the next round. The one round where payroll harms a team is the fourth round, but the negative coefficient is not statistically significant. The first, second and fifth round tested very high and significant at the 95% level. The final round has the largest

coefficient. The result states that an extra million dollars in payroll increase a team's odds of winning the post season by 7.5%. This is a substantial increase. In other words, once a team with a big payroll reaches the Stanley Cup finals, it has an excellent chance of winning the Cup.

<b>Table 19</b>				
<b>Probit Regression – Payroll / Playoff Performance</b>				
<b>Round</b>	<b>Estimated Coef.</b>	<b>T-Ratio</b>	<b>Elas. at Means</b>	<b>Aggregate Elas.</b>
Round 1	0.034541	3.5716	1.0058	0.96084
Round 2	0.024032	1.9433	0.95866	0.95024
Round 3	0.77797E-08	0.72800	0.19526	0.19863
Round 4	-0.82349E-08	-0.54062	-0.21365	-0.20497
Round 5	0.074383	2.0193	2.4332	2.0589

Overall, the money variables slightly determine the likelihood of advancement in the post-regular season. This doesn't mean that the playoffs are completely certain, however it does mean that a team could increase its chances of making it farther into the playoffs if they could raise more revenues or have a higher payroll.

## **V - Conclusion**

It is said that a truly uncertain league is one that exhibits complete competitive balance. This is what could be best for joint profit maximization by all teams in the league. As it says in the opening quote, many fans go to a game to see their team play against a challenging opponent. Once there are few challenging opponents there is less interest in the league as a whole. This point is exhibited every game a team is well ahead of the other team. In this situation, a game that could have been sold out to begin with will show a lot of empty seats before the end of the encounter. In order to maintain fan interest, the National Hockey League has to decrease the amount of control that large-market teams benefit from.

Over the past thirteen years the NHL has seen rising inequality off the ice. Payroll and revenue for some teams have skyrocketed over the sample period. On the other hand, some smaller market teams have had their payroll and revenues rise, but have had them rise disproportionately less than the league average. A wide divergence has

begun to occur as was exhibited in the section on payroll and revenue imbalance. This needs to be stopped if the league wants to remain competitive.

The performance variables do not exhibit the same departure from equality that the monetary variables do. During the studied period the actual divided by idealized standard deviation statistic fluctuated up and down finishing a little higher than where it started. The average of this statistic over the nineties decreased when compared to a study done by Quirk & Fort. In addition to this, it was found that whereas payroll and revenue did have a positive effect on regular season performance, the effect is not very substantial. An extra million dollars in revenue or payroll does not increase winning percentage that much; however, when some teams have earnings or payrolls that are twice as big as the other teams, these differences generate substantial performance differentials.

The playoffs exhibit much the same results as the regular season. It can not be said conclusively that there has been a decrease in competitive balance in the post-regular season. The number of playoff games in a season – one measurement of playoff balance – has increased over the sample, portraying that there has been an increase in balance. As well the championship concentration has dispersed relative to earlier times.

Revenue and payroll have a strong positive relationship with playoff outcome. This can be seen in the correlation of the two variables. Conversely, an extra million dollars in revenue or payroll does not increase the likelihood of a team progressing through the playoffs.

While it doesn't seem that increasing money imbalance is causing the same increase in performance imbalance, if the National Hockey League continues in the same direction, it is not hard to imagine a worse-case scenario. Many fans can see the existing revenue and payroll imbalance as a precursor to an inequitable league. There have been some remedies for this used in other professional sports leagues in North America.

The National Football League has long been the most balanced North American professional sports league. One of the main reasons that they have been the most monetarily balanced is the revenue sharing structure that they use. 70% of all the league revenues are shared. This allows for very little inequity between franchises.

The NFL also tries to control performance imbalance. This is done, by using an unbalanced schedule. If a team places poorly in a season, they will face proportionately more lower-placed teams in the following season. While this may portray balance, it can also hide an underlying imbalance.

Leagues are also using a salary cap to control the increasing divergence in monetary variables. A salary cap limits a franchise's payroll to a certain nominal amount. Some leagues are already enforcing this, but its effect depends on the strictness with which they are enforced.

Further investigating these regulations would be a good area for future research. Once the above conventions have been in place for some time it would be worthwhile to investigate their effect on competitive balance. Using a strict theoretical approach, these policies should increase competitive balance, but this remains to be seen.

## Appendix A:

### Rank Correlations in the NHL 1989-2003

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	Average
1 Year Lag	0.75592	0.58407	0.41482	0.49379	0.71160	0.81999	0.58506	0.60798	0.55641	0.54825	0.57042	0.55729	0.48165	59.133%
2 Year Lag	0.34152	0.20747	0.28569	0.58352	0.62804	0.53350	0.43895	0.52402	0.46420	0.34346	0.24402	0.46785		42.185%
3 Year Lag	0.07443	0.15535	0.12915	0.65621	0.42033	0.65000	0.19123	0.33201	0.40815	0.11418	0.36894			31.818%
4 Year Lag	0.05599	0.10176	0.44813	0.34941	0.52777	0.33017	0.11865	0.13159	0.00300	0.28818				23.546%
5 Year Lag	(0.10924)	0.17043	0.30767	0.25767	0.24521	0.27264	(0.15648)	(0.38741)	0.10947					7.888%
6 Year Lag	(0.00765)	0.05432	0.06642	0.06406	0.06252	0.15556	(0.25420)	0.09887						2.999%
7 Year Lag	0.00296	0.04315	(0.13112)	0.00292	(0.09195)	(0.05600)	(0.08550)							-4.508%
8 Year Lag	(0.10924)	(0.07606)	(0.08589)	(0.34068)	(0.23023)	(0.07905)								-15.353%
9 Year Lag	0.09845	(0.15302)	(0.42306)	(0.30632)	(0.10163)									-17.712%
10 Year Lag	(0.29130)	(0.39849)	(0.56279)	(0.22212)										-36.867%
11 Year Lag	(0.52351)	(0.50411)	0.01268											-33.631%
12 Year Lag	(0.51219)	(0.32790)												-42.004%
13 Year Lag	(0.41465)													-41.465%

### Winning Percentage Lag Correlations 1989-2003

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	Average
1 Year Lag	0.96	0.93	0.90	0.76	0.91	0.94	0.87	0.91	0.87	0.89	0.65	0.62	0.50	0.82
2 Year Lag	0.92	0.88	0.74	0.80	0.87	0.86	0.80	0.84	0.84	0.67	0.60	0.51		0.78
3 Year Lag	0.85	0.68	0.72	0.74	0.84	0.87	0.69	0.80	0.65	0.63	0.34			0.71
4 Year Lag	0.68	0.70	0.65	0.63	0.84	0.74	0.71	0.50	0.71	0.34				0.65
5 Year Lag	0.65	0.62	0.58	0.73	0.71	0.77	0.36	0.55	0.36					0.59
6 Year Lag	0.58	0.53	0.71	0.54	0.69	0.50	0.52	0.38						0.56
7 Year Lag	0.52	0.67	0.49	0.53	0.31	0.63	0.25							0.48
8 Year Lag	0.65	0.46	0.53	0.41	0.53	0.25								0.47
9 Year Lag	0.49	0.49	0.39	0.65	0.18									0.44
10 Year Lag	0.47	0.33	0.66	0.19										0.41
11 Year Lag	0.29	0.56	0.10											0.32
12 Year Lag	0.56	0.03												0.29
13 Year Lag	0.01													0.01