CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Do National Hockey League Players Perform Better During Their Contract Years?

Bachelor's Thesis

Author: Jan Liehman

Study program: Economics (B6201)

Supervisor: Matěj Opatrný, MSc.

Year of defense: 2019

Declaration of Authorship

The author hereby declares that he or she compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agrees with the thesis being used for study and scientific purposes.

Prague, July 31, 2019

Jan Liehman

Abstract

This thesis analyses the validity of the contract year phenomenon in the National Hockey League (NHL). This notion, that players increase their performance in the final year of their contract, has been supported by previous findings in baseball and basketball, but ice hockey has been largely overlooked thus far. This thesis further extends the analysis to the second-to-last year on players' contract and distinguishes between unrestricted and restricted free agents. Rigorous analysis of the NHL contract structure is employed to eliminate contract states that would bias the results. By using a within-player fixed effects model in combination with advanced performance metrics, it arrives at an unexpected conclusion that upcoming unrestricted free agents improve their performance in the second-to-last year of their contract, but there is not enough evidence to suggest that they improve in the final one. No performance increase was found for upcoming restricted free agents during the final two years of their contract.

JEL Classification	C23, D01, J30, J41, Z20
Keywords	contract year phenomenon, NHL, incentives, motivation

Title	Do National Hockey League Players Perform Better During Their Contract Years?

Author's e-mail	`jan.liehman@gmail.com`
Supervisor's e-mail	`matej.opatrny@fsv.cuni.cz`

Abstrakt

Tato práce zkoumá přítomnost fenoménu kontraktového roku v National Hockey League (NHL). Tato teorie, která předpovídá zlepšení výkonu hráčů v posledním roce jejich smlouvy, byla podpořena předchozím výzkumem v baseballu a basketbalu, zatímco výzkum tohoto jevu v ledním hokeji byl z velké části opomíjen. Tato práce dále rozšiřuje analýzu o předposlední rok hráčských smluv, a zároveň rozlišuje mezi chráněnými a nechráněnými volnými hráči. Za účelem eliminace smluvních stavů, které by ovlivňovaly výsledky, je použita důkladná analýza kontraktové struktury v NHL. Pomocí modelu s fixními efekty v kombinaci s pokročilými statistikami dochází tato práce k překvapivému závěru, že nadcházející nechránění volní hráči vykazují lepší výkon v předposlením roce smlouvy, zatímco vliv posledního roku smlouvy je nejednoznačný. U nadcházejících chráněných volných hráčů nebyl nalezen žádný efekt ani v jednom z posledních dvou let jejich smlouvy.

Klasifikace JEL	C23, D01, J30, J41, Z20
Klíčová slova	fenomén kontraktového roku, NHL, incentivy, motivace

Název práce	Podávají hráči NHL lepší výkony v posledním roce své smlouvy?

E-mail autora	`jan.liehman@gmail.com`
E-mail vedoucího práce	`matej.opatrny@fsv.cuni.cz`

Acknowledgments

I am grateful to my supervisor Matěj Opatrný, MSc., for his patience, time, and guidance. I am also grateful to my family and friends for their undying support, and many productive discussions. Finally, I would like to acknowledge the hockey analytics community and the team at CapFriendly.com for their work.

Typeset in FSV LaTeX template with great thanks to prof. Zuzana Havránkova and prof. Tomáš Havránek of Institute of Economic Studies, Faculty of Social Sciences, Charles University.

Bibliographic Record

Liehman, Jan: *Do National Hockey League Players Perform Better During Their Contract Years?*. Bachelor's Thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2019, pages 52. Advisor: Matěj Opatrný, MSc.

Contents

List of Tables

Acronyms

AAV Average Annual Value

CBA Collective Bargaining Agreement

ELC Entry-Level Contract

MLB Major League Baseball

NBA National Basketball Association

NFL National Football League

NHL National Hockey League

NHLPA National Hockey League Players' Association

RFA Restricted Free Agent

SPC Standard Player Contract

UFA Unrestricted Free Agent

Bachelor's Thesis Proposal

Author	Jan Liehman
Supervisor	Matěj Opatrný, MSc.
Proposed topic	Do National Hockey League Players Perform Better During Their Contract Years?

Motivation In professional team sports, athletes are given contracts for a definite period of time, typically in years. The last year of their expiring contract is known as the "contract year". In public and media opinion, players are often believed to make extra effort during their contract year to secure a more lucrative future contract. Teams, on the other hand, appear to be willing to pay players the extra money because they tend to overvalue recent performance. This is known as the "contract year phenomenon". In an industry with no indefinite contracts this could be viewed as a form of monetary incentive. In my thesis, I would like to examine this phenomenon by comparing performance indicators (i.e. statistics) of National Hockey League (NHL) players in and out of their contract years. To find out whether the phenomenon is real or not, and if it is real, how big is its impact on player performance.

Contribution The thesis will build on existing literature on the contract year phenomenon. Most of the research on this topic focuses on the other major North American sports leagues – MLB, NFL, or NBA. I would like to extend it in the NHL. I would also like to separate players by position to make statistics more comparable.

Methodology The analysis will use official NHL statistics as performance indicators. They will be combined with data on players' contract status. Although the NHL does not officially publish details of player contracts, the details can be obtained via fan-made websites such as CapFriendly.com. The theoretical part will introduce conditions of labor market in the NHL and look at previous research on this topic. The empirical part will rely on econometric analysis of the created dataset. Conclusions will be based on comparing the results with studies on other sports leagues and other studies related to effect of bonuses on performance.

Outline

1. Introduction
2. Literature overview and theoretical background
3. Description of the data
4. The model
5. Analysis of the results
6. Conclusion

Bibliography

Liu, Ping and Xuan, Yuhai, The Contract Year Phenomenon in the Corner Office: An Analysis of Firm Behavior During CEO Contract Renewals (October 7, 2016).

Park, S., and Sturman, M. C. The relative effects of merit pay, bonuses, and long-term incentives on future job performance, 2009.

Perry, Dayn. "Do Players Perform Better in a Contract Year?" Baseball between the Numbers: Why Everything You Know about the Game Is Wrong. Ed. Jonah Keri. New York: Basic, 2006.

Ryan, Julian. 2015. Show Me the Money: Examining the Validity of the Contract Year Phenomenon in the NBA. Bachelor's thesis, Harvard College.

Vollman, Rob. Stat Shot : The Ultimate Guide to Hockey Analytics. Toronto: ECW Press, 2016.

Wooldridge, Jeffrey M. Introductory Econometrics : A Modern Approach. Mason, OH: South-Western Cengage Learning, 2013.

Author

Supervisor

Chapter 1

Introduction

Professional sports provide an attractive setting for statistical analysis with their wealth of information. One of the intriguing topics that the setting provides is the one of the so-called contract year phenomenon. Unlike in a typical office setting, athletes in professional team sports cannot receive contracts for an indefinite period, meaning that all contracts between players and teams are fixed-term contracts. Moreover, players' careers are limited in time and are under the ever-present threat of an injury. This motivates them to maximize their career earnings while they can. The contract year phenomenon stems from the notion that players on expiring contracts have an increased incentive to perform to the absolute best of their abilities in order to improve their leverage in negotiations to help them secure as lucrative a contract as possible. Previous research has found compelling evidence for the existence of this phenomenon in the Major League Baseball (MLB) and the National Basketball Association (NBA) (Perry 2006; O'Neill 2014; Ryan 2015).

However, the research has largely overlooked the study of the contract year phenomenon in the National Hockey League (NHL), even though its labor market conditions possess the traits that make the presence of this phenomenon likely. Just like in the MLB and the NBA, all NHL player contracts involve guaranteed money. Not only that, but players also cannot be fired in the traditional sense of the word, barring a severe material breach of contract on the part of the player. Teams have the option to trade players to a rival team, in which case their current contract is picked up by the new team. Teams may also choose to buy out severely underperforming players, although that is a costly option available only under certain circumstances. It is therefore likely that NHL players will see any contract through and can subsequently reap the rewards of any contract they previously earned.

The compelling evidence from the other sports, combined with the NHL labor market conditions, and the lack of studies on this phenomenon in the NHL, along with recent emergence of new ways of measuring player performance in the NHL, all stand behind the motivation for this thesis. It hopes to extend the previous research from other sports to the NHL, where it wants to present the first comprehensive study on the contract year phenomenon. This is to be done by analyzing the NHL player contract structure in order to understand what motivations might stem from different contract characteristics. This allows for proper determining of the contract year and non-contract year states. It also leads to the inclusion of second-to-last years of players' contract into the analysis. This information, along with the new measurements of player performance, is then to be used in a within-player fixed effects estimation to answer the questions whether players perform better in their contract years, and as mentioned, if that is the case in the second-to-last years of their contract.

The rest of the text is organized as follows: Chapter 2 provides an overview of literature on the topics of the contract year phenomenon, NHL contract structure, and player evaluation. Chapter 3 describes the dataset consisting of performance data as well as contract information on NHL players from 2013 to 2019. Chapter 4 delves into the methodology and suggested structure of the model. The results of the used econometric models are presented and discussed in Chapter 5. Chapter 6 concludes.

Chapter 2

Literature Review

This section is divided into three subsections. Section 2.1 provides an overview of the previous research on the topic of the contract year phenomenon. Section 2.2 delves into the workings of the NHL labor market and discusses incentives arising from contract structure that are crucial for this analysis. Finally, Section 2.3 takes a look at measurements of individual performance in the NHL and ways of adjusting for factors outside of player's control.

2.1 Contract Year Phenomenon

Previous research on the contract year phenomenon has, for the most part, focused on its study in the MLB, followed by the NBA. In our effort to extend the research to the NHL, we mainly draw inspiration from two previous studies.

Primarily, it is O'Neill (2014), who focuses on MLB hitters with at least six years of experience and uses a within-player fixed effects estimation on a six-year sample with 1016 player-year across observations 256 players. The author finds a 6.7 percent increase in performance for MLB hitters in their contract years. Noteworthily, the author uses an estimate of the probability of retirement as an explanatory variable, which is discussed further in Section 2.2.

The second one is the work of Ryan (2015), who provides perhaps the most comprehensive study on the contract year phenomenon in the NBA. The author rigorously discusses the contract states in the NBA in order to eliminate potential bias and uses advanced baseball metrics to measure individual performance, something this thesis hopes to replicate in the NHL. The author arrives at a conclusion that the contract year effect translates into a 3-5 percentile boost for median player.

Curry & Drummond (2014) provide one of the lone attempts at measuring the contract year phenomenon in the NHL. They use a simple analysis by taking every NHL forward who signed a new contract over a seven-year period. They look at their production, measured by points per 60 minutes during 5-on-5 play, over a period of four seasons: two seasons prior to the contract year, during their contract year, and one season after the contract year. First, they simply take means of production of all forwards for each of the four aforementioned seasons. They later extend this by splitting the dataset into quartiles based on production, and then take means for each quartile in each season separately. They find that the bottom half of the dataset consists mainly of declining players, with drops in production with every additional season. The second best quartile exhibited stable production over the first three years, including no increase in production during contract year, with a slight drop during the first year of the new contract. The top 25% is the most interesting group, with stable production during the two years prior to their contract year, followed by a sharp increase during the contract year. This increase in production is then carried over to the first year of the new contract.

The analysis of Curry & Drummond (2014) is problematic as it takes player production at face value, and fails to control for factors such as age, quality of teammates, player usage, and so forth, all of which might have severe impact on production. Moreover, many potential other issues of this analysis are unfortunately left unexplained by the authors. These issues include whether players are required to be on a at least three-year contract prior to signing of the new one, or whether only players who were able to secure the next contract are included. While the analysis of Curry & Drummond (2014) severely lacks in rigor, there are still takeaways from their work. Primarily it suggests that the contract year phenomenon concerns mainly better players, who have the ability to stay in the league despite perhaps not giving their best effort every single year.

The contract year phenomenon is not limited to professional sports. Liu & Xuan (2016) examine the behavior of 159 CEOs of S&P 500 companies on fixed-term contracts during the years from 2001 to 2010. The authors find compelling evidence for opportunistic behavior during the final year of CEOs' contract. They also suggest that "the upcoming contract renewal can also have disciplinary effects on potential value destroying behaviors of CEOs". More importantly for us, Liu & Xuan (2016) conclude that this behavior also translates into better employment terms on new contracts.

This is supported by Hochberg (2011), who reaches a similar conclusion in the MLB. The author finds that teams overvalue players' performance from the most recent season disproportionately to its role as a predictor of future performance. These findings indicate that the presence of this phenomenon in the NHL is not unlikely, which brings us to the next section.

2.2 Contract Structure in the NHL

This section provides an introduction into the workings of NHL labor market. More importantly, it discusses the structure of player contracts that is relevant to the investigation of the contract year phenomenon. Understanding the nuances of player contracts and the incentives arising from them is crucial for correct estimation the relationship between contract years and performance. Failing to do so would result in including player-seasons in which player's performance might be incentivized by other factors, biasing the results.

The terms and conditions of employment of all professional hockey players playing in the NHL are set out in the Collective Bargaining Agreement (CBA). This contract is signed between the NHL and the National Hockey League Players' Association (NHLPA), which serves as the collective bargaining representative for all current NHL players. The CBA also determines the respective rights of the NHL Clubs, the NHL and the NHLPA.

Team spending on player salaries is limited by a salary cap that is set for each season. The NHL salary cap system was introduced by the 2005 CBA in an attempt to improve competitive balance in the League and prevent player salaries from escalating out of control. In effect, the salary cap limits the amount of dollars available for player salaries in each season, and players are in a sense competing against each other for a higher share. Each player's cap hit, the amount he takes away from his team's salary cap in each season, is based on the Average Annual Value (AAV) of his contract, which limits teams' ability to front-load or back-load contracts in order to circumvent the cap system. The current CBA was ratified in January 2013, concluding a four-month lock-out that resulted in a a shortened 2012-2013 regular season from the typical 82 games to merely 48. The 2013 CBA is set to remain effective until September 15, 2022, with both the NHL and the NHLPA having the right to terminate the CBA after the 2019-2020 season.

Players enter the NHL labor market via the annual NHL Entry Draft, unless they are no longer draft-eligible, in which case they enter via unrestricted free agency (CBA, Article 10.1(d)). The NHL Entry Draft is typically held at the end of June after the conclusion of the previous season. During the process, teams select rights to draft-eligible players in a predetermined order. In its current format, the draft has seven rounds with each team receiving one selection per round, barring any compensatory draft selections. Draft order is determined based on team standings after the conclusion of regular season, the results of the Stanley Cup Playoffs, and draft lottery. This system favors the weakest teams of the previous season and rewards them with the highest selections in order to maintain the League competitive. All players become draft-eligible in the year in which they reach their eighteenth birthday between January 1 preceding the draft and September 15 following the draft (CBA, Article 8.4(a)). Players under the age of 25 signing their first NHL contract must sign a so-called Entry-Level Contract (ELC). The CBA severely limits maximum base salary on ELCs and determines their length depending on the signing age of the player. For players aged between 18 and 21 years the ELC length is set to three-years, players who are 22 or 23 years old must sign two-year ELC, and 24-year old players must sign a one-year ELC (CBA, Article 9.1(b)).

Prior to the 2013 CBA, all other NHL player contracts had no formal limit on maximum length imposed on them. The 2013 CBA changed that, as it introduced a limit on maximum contract term of seven years, with teams re-signing their own players receiving the luxury of being able to offer an additional eighth year (CBA, Article 50.8(b)). This restriction was part of the League's reaction to several controversial contracts which tried to circumvent the salary cap system under the 2005 CBA. This typically involved adding extra years with minumum salary to the end of a player's contract in order to bring the cap hit down, while it seemed fairly reasonable to assume that the player would retire before the expiration of his contract and would not therefore play in those final years of the contract. The most prominent case of this behavior was a contract signed by Ilya Kovalchuk and the New Jersey Devils in the summer of 2010, which was actually blocked be the League for cap circumvention.[1]

As a result, new measures were added to the 2005 CBA to control contracts with term of five years or more (Bernstein 2011). The 2013 CBA additionally capped salary variance at 35 percent from year to year, with an extra provision

[1] Kovalchuk would have been 44 at the end of a 17-year contract, and while his cap hit was $6 million, he was slated to receive $6 million in the first two seasons, $11.5 million for the following five seasons, $10.5 million in the 2017-18 season, $8.5 million for the 2018-19 season, $6.5 million in 2019-20, $3.5 million in 2020-21, $750,000 the following season and only $550,000 for the final five years (ESPN 2010).

limiting the variance between last year and highest-salaried year to no more than 50 percent (CBA, Article 50.7).

These changes might have several implications. Firstly, players are now likely to sign more contracts over their career than in the past, making the analysis of the contract year phenomenon more meaningful. On the other hand, players are now less likely to sign a big contract that would take them to the end of their career early on in their career, which might increase the pressure to perform in non-contract years in order to secure the future contract.

Once a player's contract expires, he becomes a free agent. Teams have the option to protect their young players by tendering them a "qualifying offer" before their contract expires. Any such player then becomes a Restricted Free Agent (RFA). A qualifying offer is a one-year Standard Player Contract (SPC) and its NHL salary must be at least between 100 and 110 percent of the player's NHL salary in the prior season, depending on the value of that prior salary (CBA, Article 10.2(a)). A player, who receives a qualifying offer, can choose to accept it or not, but qualifying offers are mostly means of retaining player's rights. This stems from the fact that the team gains right of first refusal or is eligible to receive compensation in the form of draft picks should the player sign with another team (CBA, Article 10.2(a)). If the player does not sign the qualifying offer, he remains an RFA. This protection gives teams an advantage in negotiations with RFAs.

An RFA can either negotiate a new SPC with the team that holds his rights, or he can sign an "offer sheet" with any other NHL team. In that case, his prior team is notified and has seven days to decide whether to match the offer sheet or to receive draft choice compensation based on the AAV of up to first five years of the contract (CBA, Articles 10.3(a) & 10.4). The prior team can no longer negotiate a contract with the RFA or trade his rights.

However, offer sheets are not a very effective way for players to secure a lucrative contract. For several reasons, structuring an offer sheet that works for the new team is quite difficult, making offer sheets quite rare. In fact, there were no offer sheets signed in any off-season from 2013 to 2018. It should be noted that the potential offer sheet becomes an offer sheet only after the player signs it, meaning that RFAs can receive offers that the public does not know about. Moreover, the mere presence of the offer sheet threat might put general managers under pressure to sign the player before his contract expires.

There is also a more effective CBA resolution mechanism available to RFAs, and that is player-elected salary arbitration. If an RFA and his employer are unable to negotiate a new contract, the player who is eligible can register for a salary arbitration, in which a neutral third-party arbitrator sets a fair amount of money and term for both sides. The team can then decide if they want to sign the contract awarded by the independent arbitrator or not. In the latter case, the player becomes an Unrestricted Free Agent (UFA).

The rules for eligibility depend on the age of the player when we signed his first SPC, and number of years of professional experience (CBA, Article 12.1). Based on the rules, player-elected salary arbitration is not available to all players coming out of their ELC. Most arbitration cases actually get settled before the arbitrator reaches a verdict. A settlement allows both the player and the team to retain control over the contract term and salary, while they have no direct control over the verdict. Arbitration hearing can also severely damage the relationship between the player and the management. Put simply, players and their agents try to convince the arbitrator how great the player is, while management groups try to convince the arbitrator how terrible the player is. Nevertheless, once the player elects for arbitration, the case gets resolved over the summer. Therefore, RFAs with arbitration rights are absolutely incentivized to put themselves in the best possible position for the negotiations.

RFAs must sign a contract before December 1 in order to be eligible to play in the remainder of the season (CBA, Article 11.4). This means that players without arbitration rights, who want to only sign with their current team and have the willingness to wait and miss the start of the season, can put general managers under pressure. This tactic was most notoriously employed by William Nylander during the 2018 off-season, when his negotiations with the Toronto Maple Leafs went right to the December 1 deadline (Shilton 2018).

Players, who are no longer eligible to be RFAs after their contract expires, become UFAs. In the most typical way, players become UFAs through the rules of the so-called Group 3 unrestricted free agency. Under those rules, they become eligible after seven years of NHL experience, defined as being on the active roster for a minimum of 40 games in a season, or when they become 27 years of age prior to June 30 at the end of the season (CBA, Article 10.1(a)). This is sometimes referred to as the ‘27-or-7’ rule. UFAs are allowed to freely negotiate with all NHL teams as the prior team loses rights for that player as soon as his contract expires (CBA, Article 10.1(a)). This freedom to negotiate with any NHL team incentivizes UFAs to perform to the best of their abilities in order to effectively create a bidding war for their services.

There are two other groups of UFAs that need to be considered further.

Restricted free agents, who do not receive a qualifying offer from their prior team, become unrestricted free agents at the end of their contract. For the purposes of our dataset, however, they are treated as RFAs as they are entering the final year of their contract with the knowledge that the team has the option to protect their rights via a qualifying offer. So-called Group 6 UFAs are players over 25 years of age, with three or more professional seasons and less than 80 career NHL games (CBA, Article 10.1(c)). Based on those rules, our analysis treats the Group 6 UFAs same as UFAs.

While the incentives stemming from the final year of a contract might not be as strong for RFAs as in the case of UFAs, it is still in their interest to put themselves in the best possible negotiating position. For these reasons, we distinguish between RFA and UFA expiry status in our analysis but include them both.

With this in mind, it is also crucial to look at the rules concerning contract extensions. Players on multi-year contracts can sign a contract extension with their current team with less than one year remaining on their contract. This is not a contract renegotiation, meaning that players must play out the remaining portion on their contract, and the extension kicks in only after that contract expires. The date from which they can sign an extension is set to July 1, while the NHL season typically commences its regular season in early October. Players on one-year deals can re-sign from January 1 onwards. These rules allow for a situation where a player is entering the final year of his contract with an extension already signed, effectively turning his contract year into a non-contract year. It is therefore necessary to control for these situations, the process of which is described in detail later in Chapter 3.

The rules for contract extensions present another challenge. The second-to-last year of any multi-year contract is a quasi-contract year (term coined by Ryan (2015)), as players have the opportunity to impress their current team and sign a potentially lucrative contract extension a year prior to the expiration of their current contract. While in this case players might only impress their current employer, and their leverage is limited compared to free agents, they are still incentivized to perform on a higher level. Risk-averse players, who do not want to experience the uncertainty of an expiring contract, might especially be motivated by this scenario, particularly if they are happy in their current team. And while Ryan (2015) eventually treats quasi-contract years as non-contract years for the rarity of extensions in the NBA, this thesis treats them as an incentive state separate from both the contract year and the non-contract year.

At first glance, it does not matter whether the player is set to become a restricted or unrestricted free agent in his quasi-contract year. Formally, the negotiating position after season's conclusion is the same in both cases. There is, however, an important distinction to make. While teams have the option to protect their RFAs via a qualifying offer in case they do not sign them to an extension, UFAs can simply walk off with no compensation for the team. This increased sense of urgency suggests an increased willingness of teams to extend players set for unrestricted free agency before they enter their contract year. Alternatively, teams may seek to trade the player and obtain some form of compensation for him rather than lose him to free agency. The probability of this happening increases when the team is out of the playoff race by the annual NHL trade deadline, which occurs in late February or in early March, when each team has roughly around 20 games left to play in the regular season. Players on expiring contracts traded from teams out of playoff contention to a team aspiring for the Stanley Cup are often referred to as 'rental players'. So it is also in the player's interest to sign an extension with the team early if he wants to stay put and avoid being traded. For these reasons, our analysis also further separates quasi-contract years into two categories based on the type of free agency the player is facing at the end of his contract, for a total of four contract year states..

Another extremely important distinction to make is the one between players playing on one-way and two-way contracts. Players on one-way contracts receive the same salary regardless of whether they play in the NHL or whether they are loaned to a minor league team, while players on a two-way contract have a second salary assigned to them for when they play in a minor league. Essentially, a one-way contract is a vote of confidence from the team to the player. It is a sign that the team expects the player to perform well enough to stay in the NHL. The minor league salary is determined as a percentage of NHL salary and is stipulated in player's contract. The percentage varies from player to player but it seldom exceeds 25% in our dataset and is often significantly lower. This means that players on two-way contracts have a very strong incentive to perform well enough to secure and maintain an NHL roster spot, which rewards them with a much higher salary. For this reason, player-seasons played on two-way contracts are not included in our analysis. This includes all ELCs as they are essentially two-way contracts by default. Moreover, ELCs can include significant performance bonuses.

The CBA allows performance bonuses to be present on an SPC in three cases. Firstly, performance bonuses are allowed for players on ELCs. Second, for players who are 35 years or older as of June 30 prior to the season start and who signed a one-year contract for that season. And finally, for players who signed a one-year contract after returning from a long-term injury, and who have played 400 or more games in the NHL, and spent 100 or more days on the Injured Reserve in the last year of their most recent contract (CBA, Article 50.2(b)). In the second and third case, both types of players can sign contracts longer than one-year, but performance bonuses can only be added to their salary if they are on one-year contracts.

When it comes to the 35-plus contracts, there is actually another more important reason why not to include them. Players over the age of 35 are much more likely to consider retirement, which translates into a lack of motivation for securing a lucrative future contract. O'Neill (2014) uses an estimate of the probability of retirement based on experience, injuries, and performance in the main model on the effects of the contract year on MLB hitters. O'Neill (2014) then concludes that "retiring players show a decline in their contract year performance and any models which ignore retirement will be misspecified". Our admittedly simplistic solution is to exclude not only player-seasons played on 35-plus contracts, but also all player seasons involving players over the age of 35 by June 30 before the season start in order to avoid this potential bias.

2.3 Measuring Individual Performance

While the hockey analytics movement is certainly not as venerable as the more prominent analytics movement in baseball, it has garnered more and more attention and recognition over the past couple of years. Analytics movements aim to extend the analysis of their respective sport by coming up with new ways of quantifying it, and thus helping to eliminate subjective biases. The most prominent of these movemements is that in baseball, where it is often referred to as sabermetrics, a term coined by Bill James, who is often recognized as its most prominent figure (Vollman 2016). The influence of the sabermetrics eventually spread to other sports, hockey not excluding. The hockey analytics community came into the spotlight during the 2014 'summer of analytics', after several of its correct predictions led to an increased interest from NHL teams and mainstream media (Vollman 2016). Number of analysts from related websites and magazines were hired by NHL teams, and the NHL itself started a partnership with SAP to include some of the statistics developed by outside analysts on its website (Vollman 2016). This thesis aims to take advantage of this evolving movement by involving its creations, sometimes lightheartedly referred to as 'fancy stats', into the analysis of player performance.

Most of the traditional metrics for measuring player performance in ice hockey are related to offensive production. Defensive contribution is a lot more difficult to capture. And while defensive metrics such as hits or blocked shots are viewed positively by traditional hockey minds, the analytics community exposed their inherent weakness. When a player hits an opponent or blocks a shot, it automatically implies that his team is not in the possession of the puck. For obvious reasons, lack of puck possession is associated with higher probability of giving up the next goal and lower probability of scoring it. So while a player with high numbers in these categories might be praised for his 'grit', it might simply mean that his lack of skill forces him into defensive situations in the first place. And as forwards have greater responsibility over offensive production, which is typically their primary assignment, the dataset includes forwards only in order to get an 'apples to apples' comparison.

The primary purpose of this section is to define a dependent variable that can be used to measure player performance. This requires boiling down player performance into a single statistic, a very difficult task in a game as complex as ice hockey. One of the key issues is the assignment of credit for on-ice actions to individual players. Scoring a goal has obvious merit, but how much credit should the passer get, or the player who was screening the goaltender? Such task is extremely difficult, much more difficult than in basebasll (Vollman 2016).

Nevertheless, there have been several attempts by the hockey analytics community to create an all-in-one statistic. One such attempt that suits the needs of this thesis is Game Score, developed by Luszczyszyn (2016). It is based on the original baseball version of Game Score developed by Bill James, and its later basketball equivalent created by John Hollinger (Luszczyszyn 2016). Its aim is to "give a rough measure of a player's productivity for a single game" with a similar scaling to points scored to increase familiarity. Its calculation is based on several relatively common and simple statistics, with each statistic having a weight assigned to it. The statistics included are goals (G), primary assists (A1), secondary assists (A2), shots on goal (SOG), blocked shots (BLK), penalties drawn and taken (PD & PT), faceoffs won and lost (FOW & FOL), 5-on-5 corsi for and against (CF & CA), and 5-on-5 goals for and against (GF

& GA). Most of the statistics are results of individual actions by the player. However, goal differential and corsi differential are results of what happens while the player is on the ice. The complete calculation can be seen in the following equation (Luszczyszyn 2016):

$$\begin{aligned} GS = (0.75 * G) + (0.7 * A1) + (0.55 * A2) + (0.075 * SOG) + (0.05 * BLK) \\ +(0.15 * PD) - (0.15 * PT) + (0.01 * FOW) - (0.01 * FOL) \\ +(0.05 * CF) - (0.05 * CA) + (0.15 * GF) - (0.15 * GA) \end{aligned} \tag{2.1}$$

Although the original intention of Game Score is for it to be used as a single game statistic, there is nothing that prevents it from being used on a season-level. In its basic form, Game Score is a cumulative statistic and in order to make it more indicative of a player's performance at any given moment, we use Game Score per 60 minutes of ice time (GS/60) as the measure of individual performance. The main advantage of Game Score is that it is based on common box-score statistics, which one might expect to be used as arguments during contract negotiations, as well as the relative readability that comes with it. Another crucial advantage is that it is repeatable from year to year, with goal differential being the only statictic that might suffer from elevated amount of randomness (Luszczyszyn 2016). The only statistic that cannot be found in a typical box score is corsi.

Named after Jim Corsi, it is perhaps the most well-known advanced statistic (Vollman 2016). In a way, corsi is similar to goal differential, also referred to as $+/-$, but instead of goals it uses shot attempts. Corsi for is the sum of team's shot attempts while the player is on the ice, while corsi against is the same for opposition's shot attempts. These sums can be used to create corsi for percentage, which is a de facto measurement of puck possession. The problem of goals is their relative rarity, shot attempts are much more frequent and for this reason they are considered to be a better predictor of future success (Vollman 2016).

The drawback of Game Score, as its author admits, is that it is completely devoid of context. Luckily, we can use several advanced statistics as explanatory variables to contextualize it. Quality of teammates, zone starts, and quality of teammates are considered the three big factors affecting player's raw shot-based metrics (Vollman 2016). Vollman (2016) also notes that "possession is the biggest repeatable factor in the success of players and teams in the NHL today". For these reasons, all three statistics are used as explanatory variables.

Controlling for quality of teammates is very important as ice hockey is a team sport, so if we want to evaluate individual performance, we should try to control for the effect that teammates have on the player. Ceteris paribus, players playing with better teammates are at an advantage as they benefit from the talent of their teammates. Controlling for quality of teammates has an advantage over simply taking into account team quality, as players on good teams who do not play with good teammates are not penalized. Once we control for quality of teammates, controlling for quality of competition becomes of lesser importance (Vollman 2016). Measurements of quality of teammates and competition are typically done through corsi-based metrics or metrics based on allocation of ice time. In our analysis, we choose the one based on corsi.

Zone starts allow us to control for player utilization. In other words, whether he is primarily used in an offensive or defensive role. Any time the game is blown dead by the referee, it starts again with a faceoff in one of the three zones of the ice – offensive, neutral, or defensive. Any time a player is on the ice for a faceoff in the offensive zone, it is recorded as an offensive zone faceoff. If the player's shift starts with that faceoff, it is also recorded as an offensive zone start (OZS). Analogously, we have defensive and neutral zone faceoffs and starts. Offensive zone start percentage (OZS%) is calculated as a share of OZS on all non-neutral zone starts (OZS + DZS). High OZS% indicates that the player is primarily utilized for offence, which puts him in favorable positions with higher likelihood of incresing his GS/60. Contrarily, if a player is utilized primarily for defensive purposes, it puts him in a disadvantage.

Finally, we need to control for changes in league-wide scoring throughout the years. One of the reasons why some of Wayne Gretzky's records seem absolutely insurmountable is because today's game is quite different from the high-scoring 1980's. Hockey is an evolving game, and the difficulty of scoring changes from year to year thanks to changes to the rulebook or other factors such as advancements in player training & equipment or development of team strategies. In recent years, the League introduced several waves of restrictions on goalie equipment in hopes to increase scoring. The 2010's also saw a gradual decrease of physicality in the game, with increased awareness on the issue of concussions and partly because the salary cap system and analytics forced players who relied solely on their physicality out of the game. All these factors have an influence on how difficult or easy it is to score in any given season.

For this purpose, we can use league-wide goals against average (GAA) for each season as a proxy for the difficulty of scoring in any given season, which allows us to control for time heterogeneity.

Chapter 3

Data Description

Unfortunately, a dataset fitting all of the needs of this research was not readily available, and its creation required combining datasets from several sources together, which required overcoming the inconvenience of different naming systems across the sources.

The first part of the dataset is a combination of three datasets from *naturalstattrick.com* for NHL forwards from six consecutive NHL regular seasons, from the 2013-2014 to the 2018-2019 season, compiled from official game reports. One on individual statistics in all situations, second on on-ice statistics during 5v5 play only, and the last one on player bios. Any NHL forward who has appeared in at least one NHL game during any of the six regular seasons is included in the dataset. The individual statistics in each season include basic information such as player name, position, games played (GP), and time on ice as well as performance data on number of goals, primary assists, secondary assists, shots on goal, blocked shots, penalties drawn, penalties taken, faceoffs won, and faceoffs lost. The 5v5 on-ice dataset was included mainly for its inclusion of corsi for, corsi against, goals for, goals against, and offensive zone starts percentage (OZSp). This merge allowed us to compute Game Score, and Game Score per 60 minutes of ice time (GS60) developed by Luszczyszyn (2016) as discussed in Section 2.3.

Player bios were included mainly for players' date of birth, which allowed us to compute their age as of June 30 before the start of each season. A dummy for centers (Center) is added to differentiate them from wingers, as their responsibilities are more complex and they take the vast majority of faceoffs. Information on players' position differs across sources and can include more positions over the course of any season. For this reason, the dummy Center is equal to one when a player is listed solely as a center for a given season either in the dataset from *naturalstattrick.com* or *corsicahockey.com*.

While the NHL itself does not officially publish data on player contracts, thanks to the excellent work of the team at *capfriendly.com*, which assembles NHL contract details from several sources, they are publicly available. Its database contains contract information on each player for the contract on which the player has played that season on. This includes, but is not limited to, contract type and whether it is a one-way or a two-way contract, contract length, signing date, expiry status, and expiry year of the contract. Unfortunately, contracts that were bought-out or terminated are not included in the contract database and had to be input manually from individual player pages on *capfriendly.com*. In case of a few players, this information was not available from *capfriendly.com* and it was instead taken from *spotrac.com* and media reports on each signing. Contract information on sixteen players is missing, but as these players only appear in the 2013-2014 season, they never make it to datasets that are used in the models. Collecting this data for seasons 2013-2014 through to 2019-2020 allowed us to include relevant contract information to each player-season in our original dataset. The additional season is included to cover for contract extensions signed in the summer of 2018, which kick in for the 2019-2020 season.

This information was later used to calculate two auxiliary variables. Years remaining (YrsLeft) indicates the number of years remaining on a player's current contract. This is then extended in what we call 'years left real' (YLR), which is equal to YrsLeft for every season of a player's contract except for the last one. If a player on a multi-year contract signs an extension with his current team before October 1 preceding the start of the final season on his contract, then the length of the extension is added to the remaining year.

The foregoing cut-off date was set to October 1st for several reasons, but admittedly, there seems to be no perfect solution to this conundrum as players might sign an extension at any point beyond July 1. With October 1, there is a risk that a player might sign just before the season starts or shortly after that. But as training camps start in September, teams perform a physical evaluation of their players. So, for negotiations that drag into October, teams have an opportunity to evaluate if a player shirked his off-season training and potentially back out of the extension and can instead wait and see how the player performs during the final season. On the other hand, contract negotiations are not done overnight, so a contract extension signed in late September indicates that the player likely had an inkling that he will sign an extension during his off-season preparation. And as players prepare individually during the off-season, they are not under team supervision. Moreover, the grueling 82-game regular season schedule combined with a lot of travel between games limit the opportunities for improvement during the season. Therefore, it seems reasonable to assume, that off-season diet and training play a role as a source of potential disparity between performance during contract and non-contract years. For these reasons, October 1 seems like a good compromise.

Having all this information allows us to create our dummy variables for contract year statuses. Namely, contract-year with UFA expiry status (**UCY**) and contract-year with RFA expiry status (**RCY**), for both of which YLR is equal to one. Similarly, we define quasi-contract year with UFA expiry status (**UQCY**) and quasi contract-year with RFA expiry status (**RQCY**), for both of which YLR is equal to two. Finally, non-contract year (**NCY**) is the performance baseline where YLR is equal to three or more, and is independent of expiry status. As discussed in Chapter 2, all such player-seasons are required to be played on one-way contracts by players under the age of 35.

The next step is to add variables controlling for quality of competition and quality of teammates. These were taken from *corsicahockey.com*, and are included as CFpQoT along with CFpQoC, and represent the weighted average corsi for percentage of a player's teammates and opponents faced.

We also include two additional advanced catch-all statistics that can be used as dependent variables for robustness checks. Firstly, Wins Above Replacement (WAR) developed by Younggren & Younggren (2019a;c;b) and obtained from *evolving-hockey.com* in the rate form per 60 minutes, WAR60. And Point Shares, an estimate of number of team points contributed by a player, developed by Justin Kubatko (Paine 2011). Point Shares were obtained from *hockey-reference.com* and then converted to per 60 rate statistic, PS60. Additionally, we convert total points into a rate statistic (P60) as a contrasting traditional metric.

Finally, league-wide regular season average of GAA was added for each of the six seasons to each player-season in our dataset. This information was also taken from *hockey-reference.com*.

The period from the 2013-2014 season to 2018-2019 season was determined by several factors. Firstly, there is a benefit to a dataset that is covered by a single CBA with the same set of incentives stemming from it. Next, the lockout-shortened season of 2012-2013 with its late start increases the risk of players performance fluctuating from their standard. Moreover, contract information for all players becomes increasingly difficult to obtain when going back in time beyond the 2013-2014 season. And finally, excluding the 2012-2013 season means that the dataset starts after the first round of compliance buyouts that were given to NHL teams to help them comply with the rules of the 2013 CBA. Each team was allowed to carry out two buyouts that would not count against the salary cap, which might have reduced the feeling of contract security for certain players. Although this change dates further back, it is also worth mentioning that the age requirement on eligibility for unrestricted free agency was gradually lowered by the 2005 CBA, and the current '27-or-7' rule came into effect only in the 2008 off-season (2005 CBA, Article 10.1(a)).

The full dataset is a unbalanced panel consisting of 3504 player-seasons from 1059 players, where different players might have different number of observations. As the statistics used are rate statistics, we need to eliminate player-seasons with limited amount of ice time as they are likely to be unsustainable had the player played a full season. By law of large numbers, the higher the number, the closer the statistics get to their true value. But it is important to consider that there are two likely reasons for a low number of minutes played in any player-season − injuries and poor performance. While injuries are somewhat exogenously determined, poor performance falls onto the players themselves. In an effort to find a suitable compromise, more limits were tested before ultimately deciding on a 200-minute time on ice minimum. This reduces the sample to 2676 player-seasons from 799 players. Summary statistics for this sample can be seen in Table A.1. This sample still includes player-seasons played on ELCs and two-way contracts, as well as by players over the age of 35.

Chapter 4

Methodology

Based on the theory outlined in Section 2.3, the suggested regression model for GS60 for player i in season t is:

$$GS60_{it} = \beta_1 CY_{it} + \beta_2 QCY_{it} + \beta_3 Age_{it} + \beta_4 Age^2_{it} + \beta_5 Center_{it} + \beta_6 GP_{it} + \beta_7 CFQoT_{it} + \beta_8 OZSp_{it} + \beta_9 CFQoC_{it} + \beta_{10} GAA_t + \alpha_i + \epsilon_{it} \quad (4.1)$$

where α_i represents individual player intercept, CY represents either UCY or RCY, and QCY either UQCY or RQCY, depending on the subsample.

The primary interest for us lies in β_1 and β_2, where the notion that players respond to incentives suggests $\beta_1 > 0$, as well as $\beta_2 > 0$, although the effect of the quasi-contract year is expected to be smaller or even negligible. Similarly, the effects are expected to be stronger for UFAs. We also anticipate $\beta_3 > 0$ and $\beta_4 < 0$, as players are expected to improve as they gain experience before slowing down towards the end of their careers (Younggren & Younggren 2017). $\beta_5 > 0$ is expected as centers take on a slightly more prominent role than wingers. Following O'Neill (2014), games played are included as a control variable to mitigate potential bias. Playing more games helps players gain confidence, which might lead to higher GS60. And players with higher GS60 are likely play in more games. This association indicates $\beta_6 > 0$. Better teammates should lead to better performance, suggesting $\beta_7 > 0$. Conversely, $\beta_9 < 0$ is expected as tougher competition is likely to have a negative effect on performance. Higher usage in the offensive zone is expected to positively influence GS60, i.e. $\beta_8 > 0$. Finally, higher league-wide scoring is expected to have a positive effect on GS60 given its prevalence of offensive statistics, indicating $\beta_{10} > 0$.

As discussed in Chapter 3, the dataset is an unbalanced panel. It seems reasonable to assume presence of individual unobserved effects. Not only that, it also seems likely that the unobserved player characteristics in the error term α_i are correlated with some of the independent variables. As an example, player's underlying ability, captured in α_i, is expected to be positively correlated with the quality of teammates he plays with, CFQoT, as better players are more likely to play with better teammates. Similarly, we can expect positive correlation of the player's unobserved ability with his percentage of offensive zone starts, OZSp, as players with higher skill level are more likely to be used for faceoffs in the offensive zone.

Under these assumptions, OLS estimation suffers from omitted variable bias (Greene 2012). The expectation that the unobserved player traits are correlated with the independent variables also suggests preference of fixed effects model over random effects model, which would be suitable if they were uncorrelated (Greene 2012). This expectation is also supported by the findings of O'Neill & Hummel (2011) who test both, random and fixed effects models, in their estimation of the contract year phenomenon in the MLB and subsequently use Hausman's specification test to compare between them, concluding that the fixed effects model is the appropriate choice.

Moreover, we are primarily interested in how is each player's behavior and performance affected by entering a state with increased incentives − contract or quasi-contract year. And as O'Neill (2014) suggests, changes in his performance should, therefore, be measured by his metrics, not those of other players. Crucially, better players are more likely to receive longer contracts, thus appearing in contract years less often. For these reasons, we are using a within-player fixed effects estimator.

As the fixed effects within-player estimation demeans the data, the constant unobserved traits in α_i do not affect estimation results O'Neill (2014). On the other hand, it prevents us from using observable time-invariant variables (Greene 2012). Another drawback is that within-player variation in variables is likely to be much smaller than variation across the whole sample (Mummolo & Peterson 2018). This, combined with the reduced degrees of freedom for the individual intercepts, diminishes the probability of finding statistical significance for the estimated coefficients (O'Neill 2014).

Although the theory suggests that fixed effects are the appropriate technique given our dataset, we still intend to re-estimate the models using pooled OLS and random effects. This then allows us to test for the presence of fixed effects using an F-test, and use Hausman's specification test to verify that fixed effects are indeed the correct choice over random effects.

Chapter 5

Estimation Results

5.1 Unrestricted Free Agency

We start by examining the effects of contract year and quasi-contract year associated with unrestricted free agency, where the effect is expected to be stronger. Table 5.2 presents results from two player fixed effects models, described in Equation 4.1, on two different subsamples.

The sample in model (1) consists of player-seasons that are classified as UCY, UQCY, or NCY, with each player required to have at least two observations in order to be able to affect the coefficients in the fixed effects estimation. This specification includes players on one-year contracts, and is resemblant to the one used in O'Neill (2014). Sample in model (1) might be a truer representation of the NHL player population, however, the notion of the potential increased performance in contract years implies that at some point the player finds himself in a state in which he is not pressured to perform. At the same time, we want to focus on players who have appeared in multiple contract states. Therefore, players in the sample of model (2) are required to have at least one NCY, and at least one of either UCY or UQCY over the six-year period in order to be included. This is a somewhat similar to reduced sample from Ryan (2015), who requires one contract year and one non-contract year from each player.

Table 5.1 shows how the structure of both samples differs based on performance quartiles of the original full sample presented in Table A.1. The shift from the full sample to the sample from (1) comes with a noticeable increase in quality. This is to be expected, as the full sample includes player-seasons played on two-way contracts, ELCs, and player-seasons from players over the age of 35. However, the shift from (1) to (2) does not appear to have significant effect on the composition. Summary statistics of both samples divided by contract state can be seen in Table A.2 & Table A.3.

Table 5.1: Structure of UFA Samples

Performance Quartile	Number of Player-Seasons				
	Full Sample	Model (1)		Model (2)	
First Quartile	669	280	(20.9%)	169	(19.1%)
Second Quartile	669	307	(22.9%)	212	(24.0%)
Third Quartile	669	324	(24.2%)	235	(26.6%)
Fourth Quartile	669	428	(32.0%)	269	(30.4%)
Total	2676	1339		885	

Both models indicate a surprising result. While the coefficient on the UCY dummy is positive, it is not statistically significant at the 10% level. It is also less than half of the coefficient on UQCY, which is significant at the 1% level in both models. Its signs and magnitudes indicate that for an upcoming UFA in his quasi-contract year, GS60 is expected to be 0.096 to 0.100 points higher compared to his non-contract year, holding all else constant. This corresponds to a percentile boost of 3-5 for median player, or an about 5% increase from the mean value of GS60.[1]

In both models, all explanatory variables have the expected signs, as discussed in Chapter 4, and, with the exception of dummy variable Center (and the aforementioned UCY), are significant at least at the 5% level. Both models also show similar results.

Based on these results, the next logical step is to look at when players actually sign the following contract. Out of the total number of 885 player-seasons from model (2), 208 of them are classified as quasi-contract years. For the purpose of this analysis, 32 UQCY player-seasons that occurred during the 2018-2019 season are excluded as these contracts expire on June 30, 2020 and the players' future is yet to be determined at the time of this writing.[2] Out of the remaining 176 player-seasons, in 7 instances players drop out of the dataset after their UQCY season as they turn 35 before the start of their UCY season.

Let us examine when the 169 players signed their following contract. Twenty-

[1]Note that GS60 is slightly negative for some players.
[2]Contract information is as of July 15, 2019.

Table 5.2: UFA Models: GS60

	Dependent variable:	
	GS60	
	(1)	(2)
UCY1	0.040	0.043
	(0.038)	(0.040)
UQCY1	0.096***	0.100***
	(0.036)	(0.037)
Age	0.368***	0.411***
	(0.106)	(0.127)
Age2	−0.008***	−0.009***
	(0.002)	(0.002)
Center1	0.152*	0.177*
	(0.086)	(0.097)
GP	0.005***	0.005***
	(0.001)	(0.001)
CFpQoT	0.050***	0.048***
	(0.007)	(0.008)
OZSp	0.010***	0.010***
	(0.002)	(0.002)
CFpQoC	−0.188***	−0.175**
	(0.057)	(0.068)
GAA	0.695***	0.555**
	(0.198)	(0.228)
Observations	1,339	885
No. of Players	354	201
R^2	0.256	0.280
Adjusted R^2	−0.021	0.056
F Statistic	33.580*** (df = 10; 975)	26.200*** (df = 10; 674)
Note:		*p<0.1; **p<0.05; ***p<0.01

five players signed a contract extension between July 1 and September 30 following the conclusion of their UQCY season. This means that their following season was not assumed to be an UCY, but it was classified as an NCY instead. The rest of the players entered the final year of their contract with it classified as an UCY. Nine players signed an extension between October 1 and December 31, with further thirteen players signing between January 1 and the NHL trade deadline. Another fifteen players signed an extension before reaching the free agent market. Most of these signings occurred after season's conclusion, shortly before the opening of the free agent market, meaning that they likely had no effect on player's performance in that season. Seventy players tested the waters of free agency and signed a contract on or after July 1. This does not necessarily mean that they signed with a different team, although that is typically the case. Thirty-seven players were left without a subsequent contract. A total of six players, (not counting the players over 35) for which the 2018-2019 season was the last year of their contract, remain unsigned and can still potentially find a team. However, this number is in line with the number of players left unsigned from previous seasons.

For players that sign an extension before or during the final season of their contract, the quasi-contract year represents the last full season that they can use during contract negotiations. This might spur them to perform to the best of their abilities. On the other hand, most players still sign their future contract after the conclusion of the final season on their contract. The in-season signings might partially help explain the unconvincing effects of the UCY as players might exhibit an increased propensity to shirking following the signing, while their season is still assumed to be an UCY. However, this theory needs to be further tested before jumping to a conclusion.

Another potential explanation of the results is that the quasi-contract represents a year when players start to get into contract uncertainty. A solid UQCY season might provide them with a safety net in case they do not perform as well as they would have hoped during the final year of their contract. Conversely, a good UQCY might lead to self-satisfaction resulting in a decreased motivation to perform well during the UCY.

Nonetheless, the results are still surprising. Perhaps not the increased performance during the quasi-contract year, more so the lack of significance of the coefficient on the following contract year. One thing to consider is that the NHL uses a 'hard cap' system, unlike MLB and NBA which use a 'soft cap'. This puts NHL general managers under increased pressure not to overpay players (both in terms of money and length), especially when combined with the crack-down on contracts that circumvent the cap system. Under this system, general managers cannot simply evaluate players based on their performance; they are forced to evaluate players based on their performance relative to their cap hit. And every extra dollar given to a player represents not only an expense, but also an opportunity cost as that dollar takes away cap space that could be used on other players, even if the team has the wherewithal to sign that player. The hard cap system also limits the amount of potential destinations for any UFA, as teams simply might not have the cap space to sign them. All this means that the free agent market in the NHL is likely not as lucrative for the players when compared to other sports, or for that matter an office environment with seemingly unlimited possible destinations.

In order to check if the results change when a different dependent variable is used, three additional models with sample from (2) are presented in Table A.6. The dependent variables are Point Shares per 60 minutes (PS60), Wins Above Replacement per 60 minutes (WAR60), and simply total points per 60 minutes (P60). In all three instances, the model loses some of its explanatory power. In case of models (1) & (2), it is likely due to better inherent contextualization of both dependent variables. In case of model (3), it might be down to the one-dimensional nature of P60, which might also be a reason for the lack of statistical significance for both contract dummies. Interestingly, the coefficients on UCY are much closer to those of UQCY in all three estimations. Moreover, UCY is significant at the 10% in model (1), and even at the 5% level in model (2), suggesting that there might be an effect after all. However, tt is still not enough to prove the presence of the contract year phenomenon given the results from Table 5.2, and so the effect of the contract year remains ambiguous. In models (1) and (2), the significance of UQCY slightly drops but remains under the 5% level, and the coefficients remain positive. Their magnitude translates into 3-4 percentile boost for median player in the case of PS60, and 5-7 in the case of WAR60.

We also separate GS60 into Offensive Game Score per 60 minutes[3] (OGS60) and Defensive Game Score per 60 minutes[4] (DGS60) in an effort to find out whether the boost comes primarily from offensive or defensive contributions. Results can be seen in Table A.7. The coefficient on UQCY is more signif-

[3]Based on goals, assists, shots on goal, penalties drawn, corsi for, and goals for with the original weights.

[4]Based on blocked shots, penalties taken, corsi against, and goals against with the original weights.

icant and corresponds to a higher percentile boost for median player in the case of DGS60 (4-6 for DGS60 vs. 2-3 for OGS60), suggesting that defensive contributions play a larger role. This might partially help explain the lack of significance of UQCY in model (3) from Table A.6, as points are a purely offensive statistic.

5.2 Restricted Free Agency

The next task is to examine the effects of contract year and quasi-contract year associated with restricted free agency. Results of models analogous to those in Table 5.2 are presented in Table 5.4. Shifts in sample composition based on performance quartiles are presented in Table 5.3. The change in sample size from (1) to (2) highlights the main issue of these models − the lack of observations of RCY & RQCY player-seasons from players who also have an NCY season. Summary statistics of both samples separated by contract state can be seen in Table A.4 & Table A.5. The caveat to the sample (1) can be seen by looking at the age and performance disparity between contract and non-contract years in Table A.4. NCY seasons are seemingly played by older, more skilled players. These limitations negatively affect our ability to correctly estimate the relationship and draw meaningful conclusions.

Table 5.3: Structure of RFA Samples

Performance Quartile	Number of Player-Seasons				
	Full Sample	Model (1)		Model (2)	
First Quartile	669	138	(14.1%)	23	(7.8%)
Second Quartile	669	215	(22.0%)	72	(24.3%)
Third Quartile	669	269	(27.5%)	102	(34.5%)
Fourth Quartile	669	355	(36.3%)	99	(33.4%)
Total	2676	977		296	

The coefficient on RCY in models (1) & (2) is actually negative, albeit very far from statistical significance. The same applies to the coefficient on RQCY in model (1). Although it is positive in model (2), it is close to zero and very far from statistical significance. With the exception of the three dummies, all variables in model (1) are significant at the 1% level and their coefficients have the expected signs. However, several variables in model (2) with the reduced

Table 5.4: RFA Models: GS60

	Dependent variable:	
	GS60	
	(1)	(2)
RCY1	−0.050	−0.049
	(0.061)	(0.071)
RQCY1	−0.039	0.016
	(0.072)	(0.091)
Age	0.518^{***}	1.096^{***}
	(0.125)	(0.378)
Age^2	-0.011^{***}	-0.022^{***}
	(0.002)	(0.007)
Center1	0.019	−0.124
	(0.087)	(0.121)
GP	0.004^{***}	0.003
	(0.001)	(0.002)
CFpQoT	0.052^{***}	0.098^{***}
	(0.009)	(0.016)
OZSp	0.009^{***}	0.005
	(0.002)	(0.004)
CFpQoC	-0.233^{***}	−0.069
	(0.075)	(0.138)
GAA	0.896^{***}	0.742^{*}
	(0.243)	(0.407)
Observations	977	296
Observations	300	80
R^2	0.226	0.266
Adjusted R^2	−0.132	−0.051
F Statistic	19.530^{***} (df = 10; 667)	7.456^{***} (df = 10; 206)
Note:		*p<0.1; **p<0.05; ***p<0.01

sample are not significant even at the 10% level, which can likely be attributed to the smaller sample size.

The only way to increase sample size with the available data is to include player-seasons played on ELCs, which leads us to another potential issue with this specification. Elite rookies are more likely to sign long-term contracts straight out of their ELC that take them directly to unrestricted free agency. Coincidentally, these players are the ones for whom it can be reasonably assumed that they are most likely to improve during their contract year (Curry & Drummond 2014). However, as previously discussed in Section 2.2, this is problematic as ELCs are two-way contracts, can include significant performance bonuses, and at this age players undergo an accelerated development compared to the rest of their career that is not necessarily homogeneous across all players. Differentiating the contract year effects from all of these factors is difficult and beyond the scope of this text. The small number of observations on these states prevents us from making any conclusions, but based on the results it is not unlikely that contract years associated with restricted free agency do not provide strong enough incentives for the players to perform above their standard. Research from Peck (2012) also indicates that there is a difference in salaries between RFAs and UFAs to the benefit of the latter.

5.3 Comparison of Methods

Finally, all four model specifications from Table 5.2 & Table 5.4 were re-estimated using using pooled OLS and random effects. This subsequently allowed us to perform an F-test for the significance of individual effects and a Hausman's specification test respectively. In all eight cases, the null hypothesis was rejected ($p < 0.0001$). These results indicate presence of the individual effects that are correlated with the independent variables, supporting the selection of the fixed effect methodology.

Results from the models are not reported, but it is worth mentioning that the POLS estimation of models from Table 5.2 indicated that UCY has negative effect on performance that is significant at the 10% level for model (2) and at 1% for model (1), not unlikely due to the increased presence of worse players in contract years.

Chapter 6

Conclusion

This thesis set out to analyze the effects of the contract year phenomenon in the NHL, based on the compelling evidence from the MLB and the NBA, as well as the lack of previous research on this phenomenon in the NHL. Previous literature also primarily focused on the last year of players' contract. However, the theoretical discussion on the workings of the NHL labor market in Section 2.2 led us to include the second-to-last year of players' contract, referred to as quasi-contract year, into the analysis as well, based on the potential incentives associated with it. It also led us to further distinguish between restricted and unrestricted free agents, for a total of four different contract states with potentially increased incentives. Their effects on player performance were then estimated using the within-player fixed effects estimator as discussed in Chapter 4, on a dataset consisting of NHL forwards over six seasons from 2013-2014 to 2018-2019. The dataset takes advantage of advanced statistics measuring and contextualizing player performance from the emerging hockey analytics movement, which is discussed in detail in Section 2.3.

The estimation of the effects of contract and quasi-contract years associated with restricted free agency was hindered by the lack of observations on those states. Nevertheless, the results do not indicate that players boost their performance during the last two years of their contract when they are set to become restricted free agents. Similarly, based on our analysis there is not enough evidence to conclude that players perform significantly better in their contract years when they are set to become unrestricted free agents, despite some minor indications to the contrary. However, upcoming UFAs were found to improve their performance during their quasi-contract years. The effect was found to translate into a 3-5 percentile boost in performance for median player from the sample, which involved only players on one-way contracts, based on conclusions made in Section 2.2. So in a sense we can say that the effect is a 3-5 percentile boost for a median player on a one-way contract. Potential reasons that might help explain the results are discussed in Section 5.1. Based on our results, this increase is more driven by defensive, rather than offensive, contributions.

The biggest contribution of this thesis lies in extending the research on the contract year phenomenon to the NHL, which was largely overlooked in the past. In that sense, this thesis could serve as a groundwork for future research on this phenomenon in the NHL. The established framework can be adjusted to adopt new ways of measuring player performance, which might further improve the estimates. Just as importantly, this thesis also contributes by discovering the 'quasi-contract year phenomenon'. Based on the findings from previous research it does not seem unlikely that the boost in performance in quasi-contract years is specific to the NHL. Nevertheless, the results presented in this thesis might inspire future research on quasi-contract years in other professional sports leagues as well.

Potential extensions might encompass quantile regression with fixed effects, which might be enticing given the results presented by Curry & Drummond (2014). Another potential extension could lie in estimating the probability of retirement similarly to O'Neill (2014) over excluding players over certain age from the dataset as carried out here. Further improvements might include comparison of player's projected performance to his actual performance or differentiating between performance in meaningful situations when the score is close and in situations where the game is already 'decided' ahead of time.

Bibliography

Bernstein, S. (2011): "Salary caps in professional sports: Closing the kovalchuk loophole in national hockey league player contracts." *Cardozo Arts & Ent. LJ*, **29**: p. 375.

CapFriendly (2019): "Database." *capfriendly.com.*

CorsicaHockey (2019): "Database." *corsicahockey.com.*

Curry, P. & M. Drummond (2014): "Top nhl players really do improve in contract year, statistics show." *depthockeyanalytics.com.*

ESPN (2010): "Kovalchuk gets 17-year deal from devils." *ESPN.com.*

Evolving-Hockey (2019): "Database." *evolving-hockey.com.*

Greene, W. H. (2012): "Econometric analysis, 71e." *Stern School of Business, New York University.*

Hochberg, D. (2011): *The Effect of Contract Year Performance on Free Agent Salary in Major League Baseball.* Ph.D. thesis, Haverford College.

HockeyReference (2019): "Database." *hockey-reference.com.*

Liu, P. & Y. Xuan (2016): "The contract year phenomenon in the corner office: An analysis of firm behavior during ceo contract renewals." *Available at SSRN 2435292.*

Luszczyszyn, D. (2016): "Measuring single game productivity: An introduction to game score." *Hockey-Graphs.com.*

Mummolo, J. & E. Peterson (2018): "Improving the interpretation of fixed effects regression results." *Political Science Research and Methods*, **6(4)**: pp. 829–835.

NaturalStatTrick (2019): "Database." *naturalstattrick.com.*

NHL (2005): "Collective bargaining agreement."

NHL (2013): "Collective bargaining agreement."

O'Neill, H. M. (2014): "Do hitters boost their performance during their contract years?" *Business and Economics Faculty Publications*, **7**.

O'Neill, H. M. & M. J. Hummel (2011): "Do major league baseball hitters come up big in their contract year?" *Business and Economics Faculty Publications*, **8**.

Paine, N. (2011): "Point shares." *Hockey-Reference.com.*

Park, S. & M. C. Sturman (2009): "The relative effects of merit pay, bonuses, and long-term incentives on future job performance (cri 2009-009)." *Compensation Research Initiative.*

Peck, K. (2012): "Salary determination in the national hockey league: Restricted, unrestricted, forwards, and defensemen."

Perry, D. (2006): "Do players perform better in contract years." *Baseball between the numbers. New York: Basic Books.*

Ryan, J. (2015): *Show Me the Money: Examining the Validity of the Contract Year Phenomenon in the NBA*. Bachelor's thesis, Harvard College.

Shilton, K. (2018): "Nylander signs in last second negotiation with leafs." *TSN.ca.*

Spotrac (2019): "Database." *spotrac.com.*

THN (2010): "Nhl rejects kovalchuk's 17-year deal with devils because of salary cap violations." *TheHockeyNews.com.*

Vollman, R. (2016): *Hockey Abstract Presents... Stat Shot: The Ultimate Guide to Hockey Analytics*. ECW Press.

Younggren, J. & L. Younggren (2017): "A new look at aging curves for nhl skaters (part 1)." *Hockey-Graphs.com.*

Younggren, J. & L. Younggren (2019a): "Wins above replacement: History, philosophy, and objectives (part 1)." *Hockey-Graphs.com.*

Younggren, J. & L. Younggren (2019b): "Wins above replacement: Replacement level, decisions, results, and final remarks (part 3)." *Hockey-Graphs.com*.

Younggren, J. & L. Younggren (2019c): "Wins above replacement: The process (part 2)." *Hockey-Graphs.com*.

Appendix A

Appendix

Table A.1: Full Sample Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Age	2,676	26.633	4.458	17.828	23.240	26.026	29.667	45.375
GP	2,676	64.411	18.421	13	54.8	71	80	84
GS60	2,676	1.756	0.894	−1.309	1.131	1.765	2.375	4.738
PS60	2,676	0.159	0.123	−0.170	0.071	0.155	0.243	0.553
WAR60	2,676	0.033	0.065	−0.244	−0.006	0.035	0.076	0.273
P60	2,676	1.707	0.706	0.000	1.177	1.667	2.166	4.691
OGS60	2,676	4.352	0.766	1.723	3.835	4.358	4.870	7.338
DGS60	2,676	−2.595	0.353	−4.492	−2.794	−2.564	−2.360	−1.594
CFpQoT	2,676	49.440	2.667	36.750	47.797	49.520	51.222	56.870
OZSp	2,676	54.097	10.602	5.760	48.310	54.740	60.402	97.120
CFpQoC	2,676	49.839	0.355	47.680	49.640	49.860	50.070	51.320

Note: Minimum of 200 minutes played in each player-season required.

Table A.2: UFA (1) Summary Statistics

UCY

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Age	342	29.765	2.478	25.189	27.745	29.701	31.552	34.944
GP	342	65.997	16.483	15	58	71	79	84
GS60	342	1.484	0.882	−1.309	0.849	1.500	2.135	3.689
PS60	342	0.129	0.118	−0.170	0.043	0.132	0.208	0.430
WAR60	342	0.021	0.066	−0.244	−0.016	0.020	0.062	0.270
P60	342	1.526	0.673	0.139	0.986	1.517	1.996	3.339

UQCY

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Age	288	28.913	2.631	23.397	26.872	28.606	30.756	34.608
GP	288	68.868	15.073	15	63	75	80	82
GS60	288	1.682	0.829	−0.555	1.093	1.721	2.271	3.535
PS60	288	0.150	0.112	−0.105	0.064	0.144	0.233	0.442
WAR60	288	0.027	0.057	−0.150	−0.004	0.034	0.067	0.184
P60	288	1.627	0.653	0.263	1.112	1.638	2.085	3.417

NCY

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Age	709	27.952	3.111	20.828	25.608	27.919	30.275	34.994
GP	709	71.674	12.845	13	66	77	81	82
GS60	709	2.167	0.871	−0.816	1.590	2.216	2.776	4.344
PS60	709	0.216	0.124	−0.117	0.128	0.216	0.300	0.553
WAR60	709	0.047	0.060	−0.164	0.015	0.051	0.086	0.196
P60	709	2.064	0.726	0.235	1.536	2.073	2.545	4.691

Table A.3: UFA (2) Summary Statistics

UCY

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Age	195	29.700	2.395	25.397	27.856	29.550	31.374	34.667
GP	195	68.364	15.668	15	64.5	74	80	84
GS60	195	1.690	0.893	−1.147	1.122	1.729	2.377	3.689
PS60	195	0.158	0.119	−0.170	0.073	0.160	0.250	0.430
WAR60	195	0.031	0.065	−0.244	−0.006	0.028	0.074	0.201
P60	195	1.705	0.681	0.158	1.229	1.701	2.159	3.339

UQCY

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Age	208	28.842	2.540	23.397	26.872	28.565	30.541	34.608
GP	208	68.899	15.388	15	63.8	75	80.2	82
GS60	208	1.796	0.814	−0.446	1.191	1.836	2.365	3.511
PS60	208	0.165	0.111	−0.098	0.076	0.165	0.256	0.442
WAR60	208	0.032	0.057	−0.150	0.000	0.041	0.073	0.156
P60	208	1.728	0.658	0.417	1.166	1.737	2.179	3.417

NCY

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Age	482	28.363	2.782	21.122	26.274	28.392	30.358	34.994
GP	482	71.459	13.006	16	66	76	81	82
GS60	482	1.997	0.848	−0.816	1.402	2.030	2.565	4.243
PS60	482	0.195	0.122	−0.117	0.107	0.191	0.282	0.553
WAR60	482	0.039	0.060	−0.164	0.008	0.040	0.081	0.196
P60	482	1.929	0.701	0.235	1.358	1.889	2.427	3.949

Table A.4: RFA (1) Summary Statistics

RCY

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Age	186	24.368	0.987	21.122	23.765	24.447	25.183	25.994
GP	186	66.032	16.780	19	57.2	72.5	80	82
GS60	186	1.672	0.714	−0.936	1.139	1.741	2.158	3.313
PS60	186	0.146	0.108	−0.156	0.079	0.151	0.221	0.461
WAR60	186	0.036	0.058	−0.189	0.007	0.040	0.070	0.185
P60	186	1.592	0.588	0.000	1.204	1.573	1.955	3.104

RQCY

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Age	119	23.494	0.845	21.028	22.974	23.461	24.165	24.969
GP	119	70.513	12.341	33	61	76	81	82
GS60	119	1.719	0.767	−0.663	1.245	1.708	2.262	3.914
PS60	119	0.156	0.113	−0.093	0.086	0.137	0.230	0.481
WAR60	119	0.037	0.058	−0.210	−0.005	0.035	0.068	0.214
P60	119	1.656	0.601	0.244	1.263	1.606	2.027	3.784

NCY

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Age	672	27.737	3.121	20.828	25.330	27.650	30.106	34.761
GP	672	71.811	12.796	13	67	77	81	82
GS60	672	2.224	0.847	−0.816	1.669	2.296	2.812	4.344
PS60	672	0.223	0.122	−0.117	0.140	0.223	0.306	0.553
WAR60	672	0.049	0.059	−0.164	0.018	0.055	0.088	0.188
P60	672	2.107	0.715	0.235	1.593	2.116	2.576	4.691

Table A.5: RFA (2) Summary Statistics

UCY

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Age	80	24.440	0.897	22.294	23.881	24.424	25.132	25.994
GP	80	71.987	12.494	33	67.8	77.5	81	82
GS60	80	1.997	0.628	0.191	1.596	2.070	2.443	3.313
P60	80	1.878	0.508	0.433	1.551	1.868	2.174	3.102
PS60	80	0.203	0.085	−0.076	0.154	0.204	0.263	0.359
WAR60	80	0.061	0.047	−0.107	0.034	0.060	0.087	0.165

UQCY

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Age	59	23.412	0.874	21.383	22.889	23.344	24.133	24.950
GP	59	74.780	9.617	42	71	79	82	82
GS60	59	2.075	0.701	0.077	1.626	2.026	2.421	3.914
P60	59	1.934	0.576	0.465	1.544	1.928	2.204	3.784
PS60	59	0.215	0.103	−0.093	0.137	0.221	0.272	0.481
WAR60	59	0.060	0.061	−0.210	0.031	0.053	0.096	0.214

NCY

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Age	157	25.988	1.695	21.811	25.014	25.917	26.919	30.194
GP	157	72.439	12.858	20	70	77	81	82
GS60	157	2.102	0.742	−0.059	1.666	2.200	2.546	4.301
P60	157	1.967	0.651	0.648	1.481	2.006	2.426	4.691
PS60	157	0.200	0.105	−0.058	0.123	0.205	0.265	0.535
WAR60	157	0.048	0.058	−0.104	0.019	0.053	0.087	0.167

Table A.6: UFA Models: PS60, WAR60, P60

	Dependent variable:		
	PS60	WAR60	P60
	(1)	(2)	(3)
UCY1	0.012*	0.009**	0.032
	(0.007)	(0.004)	(0.036)
UQCY1	0.014**	0.008**	0.045
	(0.006)	(0.004)	(0.033)
Age	0.057***	0.025*	0.374***
	(0.022)	(0.014)	(0.115)
Age^2	−0.001***	−0.001**	−0.008***
	(0.0004)	(0.0002)	(0.002)
Center1	0.020	0.013	0.050
	(0.017)	(0.011)	(0.088)
GP	0.001***	0.001***	0.003***
	(0.0002)	(0.0001)	(0.001)
CFpQoT	−0.001	−0.002*	−0.007
	(0.001)	(0.001)	(0.007)
OZSp	0.001***	0.0004	0.006***
	(0.0004)	(0.0003)	(0.002)
CFpQoC	−0.005	−0.023***	−0.040
	(0.012)	(0.007)	(0.062)
GAA	0.011	0.007	0.741***
	(0.039)	(0.025)	(0.207)
Observations	885	885	885
No. of Players	201	201	201
R^2	0.182	0.144	0.116
Adjusted R^2	−0.073	−0.123	−0.159
F Statistic (df = 10; 674)	14.966***	11.343***	8.846***
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table A.7: UFA Models: OGS60, DGS60

	Dependent variable:	
	OGS60	DGS60
	(1)	(2)
UCY1	0.030	0.013
	(0.037)	(0.018)
UQCY1	0.064^{*}	0.036^{**}
	(0.034)	(0.017)
Age	0.232^{**}	0.166^{***}
	(0.118)	(0.057)
Age^2	-0.006^{***}	-0.003^{***}
	(0.002)	(0.001)
Center1	0.107	0.068
	(0.090)	(0.044)
GP	0.003^{**}	0.002^{***}
	(0.001)	(0.001)
CFpQoT	0.027^{***}	0.021^{***}
	(0.007)	(0.004)
OZSp	0.011^{***}	−0.001
	(0.002)	(0.001)
CFpQoC	−0.083	-0.086^{***}
	(0.063)	(0.031)
GAA	1.091^{***}	-0.534^{***}
	(0.211)	(0.103)
Observations	885	885
No. of Players	201	201
R^2	0.174	0.264
Adjusted R^2	−0.083	0.035
F Statistic (df = 10; 674)	14.243^{***}	24.192^{***}
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$	