

Compensation, Free Agency, and Future Performance in the NFL: A Market Analysis

Wagish Bhartiya
Duke University
Durham, NC
Spring 2004

This paper empirically investigates the driving factors of hiring and compensation in the NFL quarterback market in an attempt to study three fundamentally different objectives. The first is to improve upon prior empirical results through methodological enhancements to the actual compensation model. The second aim is to answer a series of market specific questions with regard to the effect of free agency on NFL salaries. The third goal is to generalize the behavioral implications of this investigation to other organizations through ex-ante/ex-post analysis. The creation of a viable compensation model suggests that compensation in the NFL QB market is a function of both performance-based and non-performance-related variables and that the value of these variables decays over time. Moreover, the inclusion of free agency variables yields results contrary to common wisdom and suggests that the true benefits of free agency are indirect. Finally, a behavioral assessment of the predictability of performance suggests that NFL franchises, like many other organizations, tend to overestimate their predictive abilities given that past performance is often not a reliable indicator for future performance.

*Honors thesis submitted in partial fulfillment of the requirements for
Graduation with Distinction in Economics in Trinity College of Duke University.*

Acknowledgements

Many, many thanks to Professor Cade Massey for his insights, guidance, and intellectual prodding throughout the course of this study.

Introduction

In any labor market, the ability for an employer to accurately value a potential hire is critical.¹ Employers in most industries typically conduct resume reviews and behavioral interviews to better understand the qualifications of an applicant. By reviewing the past experiences and qualifications of each applicant, the firm is better equipped to make a decision as to future performance of an applicant. Moreover, such a review can enable a firm to better determine the salary parameters of a potential hire by accurately valuing the applicant's contributions to the company in the future. The market price of an applicant is estimated by combining the traits of an applicant with the value the firm has for each trait. In its simplest form, a firm should decide to hire applicants if their value to the firm is greater than the wage being demanded.

The National Football League (NFL) represents a labor market in which such hiring practices can be adequately evaluated. A study of the employment methods within the NFL, particularly those involving quarterbacks (QBs), is interesting for three reasons. First, free agency² and the seemingly exorbitant salaries that players receive have come under increased scrutiny by the media in recent years. A model that can shed light on such hiring practices can potentially highlight the carriers of value and derive structure from a seemingly chaotic market. Moreover, a comparison between ex-ante and ex-post compensation models will better illustrate the mindset with which NFL franchises act. Such a task can be achieved by comparing the determinants of salary prior to the hiring decision with the productive value received from the hire in the future. One would expect that teams typically pay a quarterback today for the associated value of his performance tomorrow. That is, the hiring and wage determination process in the off-season should reflect the predictive capability teams have over pricing a QB's performance during the upcoming season.

A second reason for undertaking such an investigation centers on the fact that the structure of the NFL offers the unique opportunity for the econometrician to investigate a largely performance-based market. Unlike most other labor markets, NFL teams are able to see the past performance of a player with nearly perfect accuracy. As a result, one would expect teams to generally place a high value on past performance. Before offering a player a new contract, his statistics from prior seasons can be evaluated and priced.

A third motivating factor for this paper is the existence of inherent value to both teams and agents in developing a model that explains which quarterbacks are signed, and for how much. Instead of relying upon a priori information and hearsay, the value of all QBs can be estimated more systematically by using econometric techniques. By using such a paradigm, a dollar value can be attached to each attribute or trait a player exhibits, and the degree to which it is exhibited. The development of such a model will allow for greater

¹ Value can be loosely defined as helping a firm achieve its goals; be it revenues, victories, etc.

² Free agency is the opportunity for a player to receive salary offers from other franchises once his contract has expired.

inquiry into the nature of the hiring process in the NFL and with regard to the degree of rationality with which franchises act.

This paper will lay the foundation for investigating the dynamics of the NFL quarterback market by developing a comprehensive compensation model for veteran quarterbacks³ from 1996-2002. As a result, the overarching aim of this paper is two-fold. One, it is meant to provide a working blueprint for future analysis of the NFL and markets similar to it. Understanding compensation in an arena where past performance is perfectly observable and quantifiable may provide insight into other related or similarly structured labor areas as well. Two, this investigation will utilize the developed models to answer a set of wide-ranging inquiries into the nature of the NFL QB market. These inquiries can be grouped as follows:

- How do NFL franchises truly value QB talent? Do they strictly use a statistical performance-based model or do they employ an augmented model with other non performance-based considerations?
- How does free agency affect salaries in the NFL? Is it a direct effect – i.e., a player must participate in free agency to benefit – or an indirect effect?
- How well does past performance predict future performance? Do NFL franchises realize the difference between past and future performance? Are highly paid QBs really worth it based upon their future performance?

At the highest level, this study has three fundamentally different objectives. The first is to improve upon prior empirical results through methodological enhancements to the actual compensation model. The second aim is to answer a series of market specific questions with regard to the affect of free agency on NFL salaries. The third goal of this study is to generalize the behavioral implications of this investigation to other organizations through ex-ante/ex-post analysis. In general, ex-ante compensation models can be thought to explain what organizations are paying for in their hires. Conversely, ex-post compensation models illustrate what organizations get in return for their money.

For the purposes of clarity, this paper will be divided into five sections. The first segment briefly overviews the reasoning for investigating the NFL QB market, reviews the scope of prior sports-related labor market literature, and highlights the improvements of the approach used in this analysis. The second piece presents the data and gives the reader a better understanding of the underlying performance fundamentals within the NFL. The third section systematically develops the compensation model by improving upon past methodology. It is at this point that the framework for answering the primary inquiries of this study is constructed. The fourth portion represents a discussion of the results. The fifth and final part provides concluding remarks and suggestions for further research.

³ Veteran quarterback is being defined in this paper as being in the league for more than five seasons.

1.1 - Why Study NFL Quarterbacks?⁴

This investigation into the nature of the NFL quarterback labor market is meant to serve two purposes. The first aim of such an endeavor is to improve the learning capacity of the chief actors within the NFL by bringing systematic analysis to a domain often ruled by anecdotal evidence and cliché.⁵ A model that describes even a few of the determining components of wages is valuable given the nature of the position's compensation. Moreover, while football is indeed a team sport, the quarterback is the most visible player of a franchise. It is not uncommon to have the starting quarterback among the highest paid players of a team and serve as a team captain. Thus, a data-driven paradigm that allows general managers (GMs) to systematically price QBs is of considerable importance. A benchmark model could be used to highlight outliers or those players who were significantly under or overpaid. Such observations would serve to better calibrate the compensation expectations of both players and franchises.

The second reason for studying NFL QBs is more exploratory. From the point of view of the researcher, understanding such exorbitant compensation figures in an arena where past performance and skill sets are perfectly visible begs investigation. The availability of player attributes and past performance characteristics lends itself to an assessment of the quality of decision making by NFL executives. By building labor models that rely upon empirical evidence, not only can one determine whether or not a player is being "overpaid" but also whether he "pans out" over time.

1.2 – A Review of Related Literature

While the specific techniques utilized in this paper may be unique, the basis for examining the NFL labor market is based upon prior research. This section highlights the key additions that several economists have made to the study of sports compensation models. Moreover, by juxtaposition, a distinction between the methodologies and ambitions of this paper and past work will be apparent and summarized in the ensuing section.

In their study of the National Hockey League (NHL), J.C.H Jones and William Walsh (1988) attempt to use econometric models to determine the impact of player skills, franchise characteristics, and ethnic background on compensation. Using salary data from 1977-1978, Jones and Walsh hope to determine whether it is the player (monopoly hypothesis) or the

⁴ The models developed in this analysis are by no means meant to explain all the observed variation within the NFL compensation structure. They serve only as guides for future research and provide insights into the minds and motivations of players and executives alike.

⁵ There are numerous factors for why the NFL labor market acts differently than most other industries. First, the set of employers within the NFL is quite small. There are currently only 32 NFL teams. Moreover, each team has a maximum number of players available for the final roster. Such an arrangement lends bargaining power to franchises when negotiating with most players. Second, the NFL, like many sports leagues, oversees the hiring practices of all franchises. This is primarily evident through the adaptation of a salary cap in 1993. Third, the NFL and the NFL Player's Association (NFLPA) have agreed to a dynamic minimum salary requirement for all players. Fourth, the NFL employs complex rules with regard to free agency, signing bonuses, performance bonuses, and re-structuring contracts. Thus, while the NFL may appear to act like many other markets at its core, it is much more complex.

franchise (monopsony hypothesis) that yields greater power during contract negotiation. Running regressions by position, the authors find that neither the monopoly nor the monopsony hypothesis is overly evident for most positions. However, they find that hockey forwards, the typical goal scoring and glamour-receiving players on a team, have some monopoly power. In addition, Jones and Walsh corroborate prior sports research and conclude that skills appear to be the prime determinant of player salaries.

Jones and Walsh motivate their investigation by looking for discrimination in the hiring process. While this paper has no such aim, the authors' findings with regard to the impact of skills and bargaining power on wage have larger implications for a study of NFL quarterback salaries. The models developed in this paper will go beyond the usage of skills and utilize actual past performance statistics as the basis for compensation. Ultimately, employers are interested in the application of skill to produce output. In addition, the value and negotiating power of high-quality quarterbacks is augmented with Jones and Walsh's findings and should be evident in the models constructed henceforth.

Building upon earlier models similar to those employed by Jones and Walsh, Lawrence Kahn (1992) uses data from the 1989 NFL season to examine the issue of racial discrimination. Given the absence of free agency at the time, or the ability for players to move from one franchise to another freely, the author hints at the increased likelihood of the monopsony hypothesis. With regard to assessing the pervasive nature of racial discrimination, the author prefers the sports industry, because unlike most other labor markets, individual performance is controlled for. Employing an econometric model, Kahn uses performance-related variables like games started, experience, and days injured to value NFL players. He validates past findings and determines that there is a slight but insignificant difference in compensation between races. More importantly, Kahn's assessment of demand factors suggests that the racial composition of the local market plays a role in determining the wage of NFL players. That is, white players earn more in metropolitan areas with larger than average percentages of whites than in other areas, and vice versa with blacks.

Kahn's investigation took place prior to free agency and consequently, his results will vary from those of this study. The bargaining landscape and compensation levels in the NFL may have shifted once free agency and the ability for talent to move freely became commonplace. Indeed, the National Football League Players Association 2003 Salary Trends publication states that the average salary in the NFL has been rising steadily since the advent of free agency. As a result of such trends, little emphasis is placed on the demand characteristics of the franchise in this investigation beyond the team's performance the previous season. Furthermore, the models developed in this paper will go beyond the performance-related variables Kahn used, and directly address past performance.

Closely related to Kahn's article in methodology, Dennis Ahlburg and James Dworkin (1991) attempt to answer whether it is player performance or experience, position, and draft that truly determines compensation. Using data from 1982, the authors develop a model based upon both performance and non-performance variables in which seniority plays a key role in salary. They argue, however, that draft should be insignificant if it is only an indicator for professional performance and should decline as a player ages. In the end,

Ahlburg and Dworkin find that both player performance and non-performance-related characteristics impact wage determination. Their models indicate that salary is significantly related to seniority and draft round, and that career performance explains two to four times as much of the variance as the previous season's data.

Like Kahn, Ahlburg and Dworkin examine the NFL prior to free agency and like most studies in the field, they look at compensation across all positions. However, Ahlburg and Dworkin's methodology is robust and useful for the goals of this investigation. The econometric distinction between career performance and previous season's performance is a technique that will be used in models developed in this paper. Likewise, the usage of both performance characteristics and non-performance characteristics will also be employed.

Eric Martin (1999) attempts to untangle the NFL market for quarterbacks using a bilateral labor model. Like many other experts, Martin argues that a quarterback's salary is a function of both on-field performance and team-related variables. Martin stresses that although the QB is the leader of the offense and receives more of the blame and credit than he deserves, football is still a team game and that the talent around a player matters. Using the salaries of 32 players from 1987-1993, the author develops an econometric model for starting QBs consisting of three unique estimations implementing both performance and team variables. Underlying much of the analysis is the assumption that players play to win and that the amount of revenue a franchise generates is a function of winning. In the end, Martin concludes that both past season and career performance are significant and that teams highly value quarterbacks with a track record of winning.

Martin's investigation is one of the few that revolves solely about quarterbacks. And while the author utilizes both performance characteristics and team-related variables, his analysis is limited to only 32 starting quarterbacks. A robust model would include data on all quarterbacks during the time period studied. Indeed, the models developed in this paper hope to go a step further and illustrate the differences in quality and compensation among different tiers of quarterback. A full understanding of the market for quarterbacks must encompass the varying degree of value a player can bring to a team depending upon his designated role.

Wallace Hendricks, Lawrence DeBrock, and Roger Koenker (2003) also analyze the hiring process in the NFL. However, the authors take a novel approach and focus more on the role of uncertainty on statistical discrimination. Seeking to better explain the linkage between ex-ante hiring patterns and ex-post productivity, the researchers look to the NFL draft. Built upon option value mathematical models, Hendricks et al. validate past research that suggests "groups may be at a disadvantage when the reliability of the test instrument used to predict their performance is less than the reliability of this instrument when it is used to predict a competing group's performance" (2003, p. 883). The end result is fewer individuals from the disadvantaged group are selected. However, research indicates that employers should take on the increased risk of the disadvantaged group if they can eliminate the poorer performers prior to selection. In the NFL draft, the disadvantaged group consists of players from lesser known colleges and athletic programs. The data show that players taken earlier in the draft typically hail from better known football teams despite the fact that conditional upon

selection in one of the early rounds, “disadvantaged” players have better careers. The opposite is visible in the later rounds of the draft.

The notions of uncertainty and ex-ante hiring versus ex-post productivity encompass one of the objectives of this paper. In determining what attributes and performance characteristics are truly valued by NFL franchises, an ex-ante model is developed. In assessing the value of a player to his team in any given season, an ex-post model is constructed. One of the central goals of this study is to better understand and explain how NFL management deals with the uncertainty of signing QBs. Establishing whether GMs are over-confident or under-confident in their predictive ability is a topic of great speculation. More importantly, such results can be generalized from the sports arena to many other types of decision-making.

1.3 – Improvements/Modifications in Methodology

Building upon past empirical frameworks, this investigation will implement three novel principles into the development of its compensation models. The first modification is establishing a distinction between performance in the previous season and performance over the length of a player’s career. By distinguishing between performance one season ago and the aggregate performance prior to last season, the model establishes a “decay” factor into the analysis. The uncoupling of past performance will enable further investigation into the nature of “memory” with regard to NFL executives. Indeed, such a model will highlight whether player compensation is driven by more recent statistics or by career performance.

The second adjustment entails correcting for the selection problems almost all prior work ignores. Unlike many other labor markets, the NFL QB market does not allow for all talent to be signed to a contract. That is, despite the fact that many players may be willing to sign with a team for a minimal amount, the nature of the NFL bargaining agreement and team structure prohibits such a move. Such an arrangement leads to selection problems when modeling for salary because the data lack completeness. For example, suppose that quarterback X wishes to play in the league but is only valued at Y. Then, if the minimum prevailing salary exceeds Y, quarterback X will not be signed. Consequently, the dependent variable is censored.

Such an arrangement is only exacerbated by that fact that a football team is comprised of only 53 players. Any one team is not likely to have more than three quarterbacks on their roster at any one time. Thus, despite the fact that hundreds of potential hires exist, only a fixed number will get signed. To correct for this selection problem, this investigation will utilize Tobit regressions in addition to Ordinary Least Squares (OLS). However, a second related but different selection problem occurs when quarterbacks leave the league yet are still valued by franchises. To correct for this secondary selection problem, players who leave the league are left in the consideration set.

The third enhancement involves the development of two unique but connected models – ex-ante and ex-post. The ex-ante model determines the value of the characteristics that go into the hiring decision and encompasses both last season’s performance and career

performance. The ex-post model ascertains the value, or lack thereof, a player brings to a franchise in a given season. While the ex-ante model aims to establish the methodology by which franchises sign QBs, the ex-post model examines whether the same methodology bears fruit post signing. In this manner, we shall be able to determine whether or not past performance is a suitable predictor for future performance.

2.1 – A Basic Description of the Data

This section reviews the performance statistics, performance-related attributes, and team-related items utilized in the models developed in this study at a high level. The next section will provide a more detailed look into the importance of each of the attributes and their predicted influence on compensation.

The bulk of the data were obtained from two resources. The statistical data were acquired from www.stats.com. These data include all performance statistics for all quarterbacks from 1976-2002. The salary data for QBs were collected from the web version of the *USA Today* for 1996-2002. The baseline criterion for quarterbacks in this dataset is that all were on a team roster sometime after 1992. In this manner, QBs are valued in the labor market even after having left the league.⁶ For the purposes of our analysis, however, all QBs in consideration have played more than five seasons in the NFL.⁷ Veteran QBs were chosen to avoid the influence of the draft.⁸

Table 1 shows the summary statistics for each season's performance. Given the structure of the dataset, each observation is a player year. Table 2 shows the same summary statistics as Table 1 except aggregated over the length of a player's career. This aggregation occurs every season such that for example, a seven-year veteran's career statistics in his second season in the NFL is the combination of his first two seasons' performance. Table 3 summarizes the performance-related characteristics and player attributes in the data. Information on whether a player was a free agent in the prior off-season and his height and weight are captured here. Some of these variables are listed on a yearly basis and others are also aggregated over a player's career. Table 4 is a summary of the adjusted salary data from 1996 to 2002.⁹

In addition, data are available or were constructed for the following: whether or not a player was on a roster the previous season, the franchise for which a player played with in a given year, the year he was drafted or first observed in the league, the percent of the total offense a player represented, the length of time a player has been with his current team, whether or not a player is playing for a new team, and a team's winning percentage in a

⁶ A number of quarterbacks return to the league after missing a year or two.

⁷ It should be noted that for the purposes of the ensuing OLS regressions, we restrict our consideration set to QBs with salary. This restriction is lifted for Tobit regressions.

⁸ Empirical evidence suggests a strong positive relationship between the earlier a player is draft and his first contract.

⁹ All salary figures have been converted to 2002 dollars.

given year. Many of these variables will be utilized in the construction of the compensation model.

Table 1. Season Performance Description¹⁰

The data summarized below pertains to the set of quarterbacks who were on a roster after 1992, have salary from 1996-2002, and have more than 5 years of experience. Each observation represents a single player-season, or the season of a single player.

Variable	N	Mean	sd	Min	p25	p50	p75	Max
s_gs	555	6.23	6.15	0.00	0.00	4.00	13.00	16.00
s_gp	640	8.67	5.87	0.00	3.00	9.00	15.00	16.00
s_p_atts	640	212.12	192.11	0.00	20.00	162.50	389.00	655.00
s_p_comp	640	124.18	114.68	0.00	10.50	92.50	227.00	418.00
s_p_yds	640	1481.43	1385.32	0.00	110.50	1118.50	2778.50	4690.00
s_ints	640	6.73	6.25	0.00	0.00	5.00	11.00	25.00
s_r_atts	640	17.53	18.42	0.00	2.00	12.00	27.00	118.00
s_r_yds	640	61.85	97.89	-19.00	0.00	24.00	82.50	942.00
s_r_tds	640	0.58	1.13	0.00	0.00	0.00	1.00	7.00
s_fum	640	1.45	1.89	0.00	0.00	1.00	2.00	9.00
s_bowl	640	0.12	0.32	0.00	0.00	0.00	0.00	1.00
s_per_com p	538	0.58	0.10	0.00	0.55	0.58	0.61	1.00
s_per_yds	538	6.81	2.14	0.00	6.11	6.85	7.44	36.00
s_per_tds	538	0.04	0.02	0.00	0.03	0.04	0.05	0.22
s_per_ints	538	0.04	0.05	0.00	0.02	0.03	0.04	1.00
s_rating	538	77.01	18.99	0.00	70.14	78.83	87.23	158.33

¹⁰ Note the smaller number of observations for variables based upon pass attempts. This is because several QBs never attempt a pass during the season and these variables are consequently incalculable.

NOTE: The variables are defined as follows: *s_gs* = the number of games started that season; *s_gp* = the number of games played in that season; *s_p_atts* = the number of pass attempts that season; *s_p_comp* = the number of completed passes that season; *s_p_yds* = the number of passing yards that season; *s_p_tds* = the number of touchdowns passed for that season; *s_ints* = the number of interceptions thrown that season; *s_r_atts* = the number of rushing attempts that season; *s_r_yds* = the number of rushing yards gained that season; *s_r_tds* = the number of rushing touchdowns scored that season; *s_fum* = the number of fumbles lost that season; *s_bowl* = a dummy variable coded as 1 if player is a Pro Bowler; *s_per_comp* = the percentage of completed passes that season; *s_per_yds* = the yards gained per attempted pass that season; *s_per_tds* = the number of touchdowns scored per attempted pass that season; *s_per_ints* = the number of interceptions thrown per attempted pass that season; *s_rating* = the quarterback rating that season.

There are three key observations from the season performance summary table. One, given the distribution of games started, full-time starting quarterbacks in the NFL are somewhat rare. Only a small fraction of players start all 16 games in a season adding value to those players who do. Two, based upon the distribution of rushing yards, many QBs are rather immobile and do not gain many yards running the ball. It is quite possible that most teams dismiss the running capabilities of a quarterback as inconsequential to his overall ability. Three, Pro Bowl seasons are infrequent occurrences in the data potentially adding value to those players named all-stars.

Table 2. Career Performance Description

The data summarized below pertains to the set of quarterbacks who were on a roster after 1992, have salary from 1996-2002, and have more than 5 years of experience. Each observation represents the career-aggregated value of that variable at a given point in time for a single player.

Variable	N	Mean	sd	Min	p25	p50	p75	Max
<i>c_gs</i>	640	33.15	34.59	0.00	3.00	22.00	53.00	173.00
<i>c_gp</i>	640	84.05	50.35	0.00	47.00	79.00	115.00	242.00
<i>c_p_atts</i>	640	2047.77	1596.47	0.00	711.00	1789.00	2893.00	8358.00
<i>c_p_comp</i>	640	1170.51	942.54	0.00	398.00	993.00	1699.50	4967.00
<i>c_p_yds</i>	640	14328.46	11621.42	0.00	4881.50	12072.00	20651.00	61361.00
<i>c_p_tds</i>	640	85.64	76.01	0.00	26.00	66.00	120.50	420.00
<i>c_ints</i>	640	70.52	53.82	0.00	27.00	63.00	98.00	252.00
<i>c_r_atts</i>	640	184.06	154.04	0.00	65.00	153.00	248.00	775.00
<i>c_r_yds</i>	640	723.75	852.00	-18.00	128.00	424.00	1047.00	4928.00
<i>c_r_tds</i>	640	6.99	7.58	0.00	1.00	5.00	10.00	43.00
<i>c_fum</i>	640	8.79	9.06	0.00	1.00	6.00	14.00	46.00
<i>c_bowl</i>	640	1.19	2.03	0.00	0.00	0.00	1.00	9.00
<i>c_per_bowl</i>	640	0.11	0.18	0.00	0.00	0.00	0.14	0.83

c_per_com								
p	638	0.56	0.05	0.22	0.54	0.56	0.58	1.00
c_per_yds	638	6.83	0.77	1.67	6.48	6.91	7.16	15.00
c_per_tds	638	0.04	0.01	0.00	0.03	0.04	0.04	0.20
c_per_ints	638	0.04	0.01	0.00	0.03	0.03	0.04	0.22
c_rating	638	75.28	11.13	0.00	69.55	75.61	81.03	158.33

NOTE: The variables are defined exactly as above except over a player's career to that point in time. Thus, the further along in a player's career that these variables are looked at, the larger these values become. In addition, c_bowl is the total number of Pro Bowls a player has been to while c_per_bowl is that total divided by the number of seasons that player has been in the league.

There are three notable observations from the career performance summary table. First, the significantly different distributions of games started and games played indicate that many more quarterbacks come off the bench and play than start any given game. Second, the nature of the career Pro Bowl variables suggests that players named to the Pro Bowl in any given season are not guaranteed a return trip later in their career. Third, the high density of career passer rating observations about the mean may suggest that QBs who are able to achieve ratings greater than 85-90 can distinguish themselves from their more mundane peers.

Table 3. Player Attributes Description

The data summarized below pertains to the set of quarterbacks who were on a roster after 1992, have salary from 1996-2002, and have more than 5 years of experience. Each observation represents the value of that variable at a given point in time for a single player. All data relating to draft are static.

Variable	N	Mean	sd	Min	p25	p50	p75	Max
draft_round	514	4.10	3.03	1	1	3	6	12
draft_pick	514	11.95	9.11	1	3	12	20	28
draft_over~l	514	99.08	86.00	1	26	82	146	319
exp	640	9.56	3.11	6	7	9	11	19
freeagent_~ e	640	0.09	0.28	0	0	0	0	1
move	639	0.36	0.48	0	0	0	1	1
tenure	640	3.48	3.09	1	1	2	5	17

ht	640	74.95	1.47	70	74	75	76	80
wt	640	216.17	11.43	180	210	215	224	248

NOTE: The variables are defined as follows: draft_round = the round in which a player was drafted; draft_pick = the pick within the round a player was drafted; draft_overall = the overall draft pick of player; exp = the number of seasons a player has been in the NFL up to that year; freeagent = a dummy variable equaling 1 if the player signed a free agent contract in the off-season; move = a dummy variable equaling 1 if the player moves to a new franchise; tenure = the length of time a player has spent with the current team; ht = a player's height in inches; wt = a player's weight in pounds.

Three items are particularly striking in the summary table of player attributes. One, quarterbacks are chosen with high frequency in all rounds of the draft despite their relatively greater importance to the team. Two, most QBs move from team to team over the length of their career with regularity. It may be determined that those players who remain with a given franchise for a longer period of time are somehow valued higher than those who do not. We shall return to this issue below. Three, the distribution of height indicates that most quarterbacks fall into a pre-conceived range of 6' 2" to 6' 5". It could be that there is a pre-screening process that takes place in college or earlier that limits shorter players from pursuing the quarterback position.

Table 4. Adjusted Salary Description (in thousands of dollars)

The salary data summarized below pertains to the set of quarterbacks who were on a roster after 1992. All quarterbacks, not only those with greater than 5 years of experience are included. This table conditions upon salary being greater than zero.

Year	N	Mean	sd	Min	p25	p50	p75	Max
1996	95	2,036	2,533	237	373	724	3,242	9,726
1997	98	1,956	2,513	248	429	751	2,134	11,100
1998	103	2,012	2,531	263	459	740	2,606	13,200
1999	108	2,004	2,472	278	531	895	2,680	10,600
2000	98	1,969	2,405	263	525	915	2,483	11,600
2001	96	2,022	2,839	24	396	709	2,404	15,400
2002	99	2,072	2,919	35	378	673	2,348	14,700

As expected, the distributions of adjusted salary from the summary table have essentially the same means. However, the overall distribution of compensation seems to be changing slightly over time. The increasing standard deviation from 1996 to 2002 suggests greater variation in the contracts players are signing. The increased diffusion in compensation may be occurring because of two reasons. First, teams may have devised a method to better calibrate player value in more recent years and distinguish between marginally different prospects. Second, as free agency became a greater and more important part of the hiring process, the dichotomy between those players who reach the free agent market and those who do not has widened.

2.2 – Performance Assessment and Predicted Influence

This section will break down the compensation driving factors into three categories – direct performance statistics, performance-related variables, and other attributes. Tables 1 and 2 from the previous section represent the set of direct performance statistics. Table 3 captures performance-related variables and other important attributes. The primary function of this section is to acclimate those unfamiliar with the NFL with the basics of quarterback performance. However, an understanding of the predicted impact these variables have on compensation will underlie the consequent models. These predictions are provided within the parentheses following each variable.

Direct Performance Metrics:

The impact of these performance variables is directly related to value. It can be argued that past performance is a strong indicator of future performance and prior research suggests that teams are likely to use past performance measures when determining a player's value and compensation level. The magnitude of these traits is perhaps the most significant aspect of salary determination.

- Passing yards (positive relationship) – The accumulation of yardage is perhaps the second most important aspect to offense in the NFL. The ability for a QB to move his team up and down the field in order to get into scoring position is highly valued.
- TDs passes (positive relationship), TDs scored (positive relationship) –The ability for a QB to not only move his team offensively but also score is important. Touchdowns-scored is a more encompassing measure that takes into account touchdowns thrown and ran for.¹¹
- Interceptions thrown (negative relationship), fumbles (negative relationship) – One of the ways a QB can stifle his offense is by turning the ball over to the other team. Interceptions are generally the most common and tangible way this can happen. Fumbles are a secondary way for a quarterback to lose possession. The market will devalue a quarterback with a higher propensity to throw interceptions or fumble.
- Completions (positive relationship), completion percentage (positive relationship) – Completions and completion percentage are characteristics indicative of the efficiency with which a player produces. As such, teams seeking systematic production desire such traits.
- QB rating (positive relationship)- The QB rating is a highly visible composite of several factors including those listed above. Used widely as the first line of comparison between quarterbacks, the rating serves as an indicator of aggregate QB performance.

Performance-related Indicators:

¹¹ The vast majority of touchdowns scored by quarterbacks are passing touchdowns.

While not identical to performance traits, these related variables measure the success a QB has had during his career either directly or indirectly. While these traits are seemingly not as important as performance, one would expect that a player's longevity, winning percentage, and national recognition to positively affect his salary.

- Victories (positive relationship), winning percentage (positive relationship) – Given that winning is the primary aim of all sports franchises, the ability to hire a proven winner is tantamount. A QB's winning percentage may also serve as a proxy for leadership ability, an attribute requisite for the leader of an offense.
- Number of starts (positive relationship), number of games played in (positive relationship) – Unless signed only to serve as a backup or mentor, most QBs are paid to produce on the field. The number of starts and games played in are good indicators of the talent level a player brings to a team.
- Experience (changes from positive to negative over time)¹² – As with most labor markets, those players who have spent several years in the league can rely upon their past experiences to assist them going forward. Experience may also serve as a proxy for intangible assets like leadership, maturity, and dedication.
- Pro Bowl selections (positive relationship) – Each year six quarterbacks are selected to the Pro Bowl. Fans, coaches, and GMs conduct voting for the Pro Bowl and those players selected are typically thought to be the best at their position. Consequently, a high value is placed upon past Pro Bowl selections.

Other Attributes:

Beyond performance statistics and related variables, there may be other signals in the market that affect compensation. Justifiable or not, these variables tend to center on the “labels” a player receives throughout the length of his career.

- Position taken in the draft (no discernible prediction) – While past literature suggests that draft is merely an indicator for performance in the NFL, there is some evidence that the influence of the draft persists throughout a player's career. It is possible that the draft may pick up on some characteristics that performance measures do not.
- Becoming a free agent (negative relationship), moving to new team (negative relationship) – A central argument of this study is that players that are valued highly by their current teams will never reach free agency. Their contracts will be re-negotiated before other teams can make offers to them. Thus, the act of not resigning a player may serve as a signal to other potential hirers that this player is currently overpaid and potentially deficient in some area.

¹² Suggests the usage of both an experience and experience-squared term in the model. Such a notion is supported by prior research.

- Tenure with current team (positive relationship) – Franchises may pay some premium for “firm specific knowledge.” Moreover, given the uncertainty of hiring that pervades much of the NFL, teams that retain talent may be willing to pay more for decreasing the associated hiring risk.¹³

3.1 – The Baseline Model

In order to simultaneously maximize both the reader’s understanding and the overall viability of the compensation model for NFL QBs, we shall present its construction through a series of additions or stages of development. This section shall formulate the baseline model from which all deviations will follow. In line with the compensation models developed in prior research, this model shall encompass mostly past performance statistics as well as examine the influence of experience and draft on player wage. However, to be consistent with the ex-ante/ex-post dichotomy uniquely developed in this investigation, the baseline model developed will be based upon lagged statistics. That is, when examining the salary in season t , we shall only consider performance and other pertinent variables from season $t-1$ and earlier.

Regression 1 is an initial attempt at developing a baseline model through the inclusion of all potentially significant performance statistics. All the variables in the regression behave as predicted except for the experience and experience-squared terms, which were predicted to have the opposite sign.¹⁴ The lagged number of games started over the length of a QBs career is positively related to wage. Lagged completion percentage, yards per pass attempt, and touchdowns per pass attempt are all also positive in compensation. Lagged interceptions per pass attempt is negative in compensation as are lagged fumbles. The influence of the draft on compensation is negligible. The key facet of this regression is that the factors of QB rating are decomposed into independent regressors.¹⁵ However, given the lack of significance on most of the decomposed variables, a more suitable model may use the actual rating instead of the individual statistics as is done in Regression 2.

Not surprisingly, the results of Regression 2 are superior to those of Regression 1. The aggregation of statistics into lagged QB rating turns out to be significant at the 1% significance level. Moreover, neither the economic interpretation of the other regressors and nor their significance changes when QB rating is used. Thus, given the better overall results of Regression 2, we shall proceed to the next section with it as our baseline model.

Table 5. Baseline Model OLS Regressions

The dependent variable is the log of the adjusted salary. Each observation represents the career-aggregated value of that variable at a given point in time for a single player. Observations are included only if a player was on a roster after 1992, has salary for the years

¹³ The role of free agency and the related market signals is an unexplored area in the much of the literature. Its usage in this paper is based upon the research goals of this investigation.

¹⁴ We shall overlook this unanticipated result until the complete model is developed.

¹⁵ The quarterback rating is calculated using completion percentage, average yards gained per pass attempt, the percentage of touchdown passes, and the percentage of interceptions.

regressed (1996-2002), and has more than 5 years of experience. * indicates t-stat significant at the $p < .05$, ** indicates significant at $p < .01$.

Variable	(1)	(2)
exp_lag	-0.144 (-1.49)	-0.153 (-1.62)
exp_sq_lag	0.004 (0.81)	0.004 (0.88)
C_gs_lag	0.017 (4.68)**	0.016 (4.40)**
C_rating_lag		0.022 (5.21)**
C_per_comp_lag	1.258 (0.66)	
C_per_yds_lag	0.004 (0.03)	
C_per_tds_lag	6.464 (1.06)	
C_per_ints_lag	-8.196 (-1.90)	
C_fum_lag	-0.037 (-3.14)**	-0.034 (-3.00)**
C_per_bowl_lag	1.979 (5.45)**	1.866 (5.33)**
draft_inv	-0.028 (-0.14)	0.036 (0.18)
Constant	14.133 (19.56)**	13.266 (27.80)**
Observations	251	251
R-squared	0.46	0.48

NOTE: games played was insignificant when included in this regression and thus dropped to avoid collinearity; 36 observations are lost when using draft_inv as a regressor given the existence of players in the league who were not drafted.

3.2 – Improving the Methodology of the Model

There are two major modifications to the baseline model that enhance its overall viability – the inclusion of performance statistic decay and the utilization of a Tobit model instead of traditional OLS.¹⁶ A third, yet relatively minor addition to the baseline model is the insertion of winning percentage. Such modifications are in line with the enhancements sought for fulfilling the first major objective of this study. As mentioned, the second and third major objectives of this investigation are to inquire into the impact of free agency on compensation and to determine the predictive capability of hiring practices through the construction and evaluation of ex-ante and ex-post models. These two major methodological improvements of previous research should allow for more accurate inquiry into the nature of free agency and facilitate the building of a better end model.

The first modification will call for the specification of a model with both the previous season's statistics and career statistics lagged two seasons. By uncoupling performance from the previous season (stats from t-1) from performance earlier in a player's career (aggregate stats up to t-2), it may be possible to better understand the decay structure embedded in the ex-ante model. The magnitude and significance of these variables may suggest how far back in time a GM is willing to venture in order to make the best hiring decision possible. It should be noted that the relative importance of performance from the previous season and career performance has been confounded in previous studies.

Utilizing the decay assumption in the model significantly improves Regression 3. When compared to the baseline model, Regression 3 can better explain the variance in the data (R-squared of 0.48 versus 0.62). Again, with the exception of the experience and experience-squared terms, all the variables - career lagged two seasons and the past season - behave as predicted. Compensation is increasing in the number of career games started lagged two seasons, career passer rating lagged two seasons, and the frequency of Pro Bowls lagged two seasons. Conversely, compensation is decreasing in the number of career fumbles lagged two seasons. As with lagged career statistics, compensation is positive in the number of games started last season, last year's passer rating, and making the Pro Bowl and negative in fumbles. The impact of the draft is again negligible.

In comparing the significance of lagged career performance with past season's performance, determining which set of statistics is more salient in compensation is not clear. For both the number of games started and passer rating, last season's statistics are more significant. However, for fumbles and Pro Bowl, career statistics prevail. Such results are consistent with prior findings and suggest that both past season's performance and career performance matter in the hiring decision. However, such results may lend credence to the "recency effect" for actual statistics and the "label effect" for career stats. Since QB rating and games started may serve to better predict direct performance next year, their more recent values are more important. Conversely, since the Pro Bowl and fumbling are more categorical by nature, their effects, for better or for worse, persist over time.

The inclusion of career and season winning percentage has a strong impact on the model. Based upon Regression 4, both variables are positively related to compensation with

¹⁶ For a more encompassing overview of Tobit regressions, the reader is strongly urged to look in a basic econometrics textbook like Gujarati's [Basic Econometrics](#).

career winning percentage significant at the 5 percent level. That said, the direction of causality here is somewhat ambiguous. It is possible that winning quarterbacks are paid more relative to losing ones because of the unstated characteristics they bring to the table. But it is also plausible that winning franchises pay more for QBs given the importance of the position to the on-field performance of the team. Further investigation into this question is needed.

Regression 5 is essentially Regression 4 as a Tobit model. It is important to note that since many of the quarterbacks now in the selection set did not play the previous season, the past season's QB rating and team winning percentage are removed from the model. However, the addition of quarterbacks not previously on a team's roster begs for the addition of a did-not-play variable. Thus, Regression 6 includes a did-not-play variable among the regressors, capturing the devaluation that occurs when a player is out of the league and compensating for the loss of rating and winning percentage.

The usage of a Tobit regression instead of OLS in Regression 6 yields some fascinating results. First, the economic interpretation of the experience and experience-squared terms flips and agrees with their previously predicted directions. Such a result can be explained by examining the type of players excluded from the OLS regression. The OLS model, which was conditional upon players receiving salary, did not capture players who were no longer in the league because of their declining value in age. Thus, since the OLS framework only included players who were on rosters, the marginal influence of experience was positive.

A second observation from the Tobit model is that many of the career performance metrics lose their significance. Only winning percentage over a player's career remains important. A closely related third observation is that the significance of games started last season and passer rating last season increases even while fumbles and Pro Bowl remain insignificant. The driving force behind such a result is the increased role of the did-not-play variable discussed above. It is important to note the large magnitude and high significance of the did-not-play lagged variable in the Tobit model. With the inclusion of all quarterbacks and not only those who were on a roster the prior season, greater weight is given to recent performance. Because a significant number of quarterbacks that did not play the season before are now valued under Tobit (note the increase in observations between Regression 4 and 6), those QBs that did play become more valuable. Conversely, those quarterbacks that did not make a roster last year are seen as less valuable.

The growing statistical dichotomy between quarterbacks who played last season and those who did not places increasing weight on recent performance. Therefore, one may conclude that the shifting importance of the performance variables when using a Tobit regression signifies that GMs have a shorter memory than expected and are more interested in performance last season than performance two or more seasons ago.

With the aim of parsimony, Regression 7 is a pared-down version of Regression 6. First, all of the career performance variables – games started, quarterback rating, fumbles, and Pro Bowl - will be dropped from the regression given their declining importance. Career winning percentage will be retained given its significance and economic meaning. Second,

the number of fumbles from last season, whether or not a player made the Pro Bowl, and last season's winning percentage will also be removed given their relative insignificance. A last consideration (not shown here) is the usage of games played instead of games started to better capture performance. The end result is a model with all of its variables significant at the 5 percent level except for experience lagged, experience-squared lagged, and inverse draft.

Table 6. Augmented OLS versus Augmented Tobit

The dependent variable is the log of the adjusted salary. Each observation represents either the player-year or career-aggregated value of that variable. Observations for the both regressions include only players who were on a roster after 1992 and have more than 5 years of experience. * indicates t-stat significant at the $p < .05$, ** indicates significant at $p < .01$.

Variable	(3)	(4)	(5)	(6)	(7)
exp_lag	-0.054 (-0.63)	-0.062 (-0.71)	-0.075 (-0.74)	0.061 (0.59)	0.048 (0.48)
exp_sq_lag	0.004 (0.09)	0.004 (0.11)	-0.004 (-0.94)	-0.008 (-1.56)	-0.007 (-1.43)
C_gs_lag2	0.008 (2.37)*	0.008 (2.30) *	0.016 (2.99)**	0.006 (1.34)	0.004 (1.71)
C_rating_lag2	0.008 (1.70)	0.006 (1.14)	0.009 (2.27)*	0.003 (0.67)	
C_fum_lag2	-0.038 (-3.41)**	-0.034 (-2.99) **	-0.010 (-0.58)	-0.014 (-0.85)	
C_per_bowl_lag2	1.066 (3.56)**	0.856 (2.78) **	-0.545 (-1.02)	0.457 (0.97)	
S_gs_lag	0.097 (8.82)**	0.092 (8.33) **	0.200 (10.39)**	0.141 (8.26)**	0.138 (11.73)**
S_rating_lag	0.005 (1.84)	0.004 (1.40)			
S_fum_lag	-0.053 (-1.84)	-0.044 (-1.53)	0.02 (0.01)	-0.051 (-1.09)	
S_bowl_lag	0.175 (1.23)	0.094 (0.64)	0.17 (0.63)	0.267 (-1.15)	
draft_inv	0.075 (0.43)	0.173 (0.99)	-0.255 (-0.84)	-0.029 (-0.11)	0.097 (0.37)
C_team_winper_lag2		1.154 (2.29) *	1.891 (2.79)**	1.88 (2.96)**	2.282 (3.78)**
S_team_winper_lag		0.389 (1.60)			
dnp_lag				-2.098 (-11.32)**	-2.098 (-11.31)**
Constant	13.291 (25.69)**	12.857 (23.53)**	11.206 (18.56)**	11.935 (20.46)**	11.911 (21.31)**
Observations	219	216	593	593	594

R-squared	0.62	0.62	0.31	0.42	0.42
-----------	------	------	------	------	------

3.3 – The Impact of Free Agency

Conventional wisdom suggests that the advent of free agency in the NFL in 1993 was quite beneficial for player compensation.¹⁷ Many believe that the ability to move from one team to another increases the negotiating power a player has, and thus leads to increased salary. Such a result is supported by the theory of labor mobility. Accordingly, one would expect the unconditional effect of free agency to be positive in compensation. The purpose of this section is to determine whether or not such conventional thought holds empirically through the development of compensation models that include free agent variables.

This analysis of free agency and its impact on wage will systematically build from the very basic to a model encompassing the significant performance variables and non-performance traits used previously in Regression 7. In this manner, the general impact of free agency can be determined without conditioning for performance. For example, Regression 8 is a simple model that has only whether or not a player was on a roster the previous season and whether or not a player is in the first year of a free agent contract. In this model, the did-not-play variable is vital as it conditions for those players who are joining new teams because they were not in the league the previous year. Regression 8 suggests that being in the first year of a free agent contract negatively impacts compensation. While it is premature to speculate, such a result may hold because only marginal players reach the free agency market. A player that a team truly values typically has his contract extended or re-negotiated before he reaches the open market.

Regression 9 bolsters the model and the developing hypothesis by including a variable representing whether or not a player moved to a new team in a given year. The negative coefficient suggests that players signing on with new teams receive less compensation. The unwillingness of a team to re-sign a quarterback may serve as a market signal for other teams to devalue the free agent. Regression 10 adds tenure, or the length of time a player has been with his current team to the model. The strong positive relationship tenure has with salary implies the importance of firm specific knowledge in the hiring process. The franchise benefits from retaining a known commodity and avoiding the associated uncertainty from hiring an outsider. In addition, the initial costs of training a player in order to create familiarity and establish a comfort level for interacting with a given coaching staff and front office is avoided.

Regression 11 augments the model by including an interaction term between move and tenure. Building upon the rationale for why tenure is so important, the movement of a player who has spent several seasons with a respective franchise is probably an unfavorable signal in the open market. A player that is not re-signed by a team for which he has played for several seasons consecutively may not be worth signing in the free agent arena. However, given the relative insignificance of the coefficient, no such claim can be validated statistically.

¹⁷ The 2003 NFLPA omnibus on player salary trends highlights this notion.

Finally, Regression 12 is the full compensation model with free agency variables included. The weakening impact of these additional variables when regressed alongside performance and performance-related statistics lends support to the hypothesis established previously. If only those players who are not highly valued by their current teams reach the free agent market, it is safe to assume that these free agents will have poorer performance statistics relative to those players who never reach the open market. The decreasing significance of the free agent variables in the face of past performance statistics validates this argument. Players who reach the open market are not paid less because they are free agents, but because their performance statistics are weaker than their non-free agent counterparts. Free agency is simply an indicator of this poorer performance.

Table 7. Augmented Tobit with Free Agent Regressors

The dependent variable is the log of the adjusted salary. Each observation represents either the player-year or career-aggregated value of that variable at a given point in time. Observations include only players who were on a roster after 1992 and have more than 5 years of experience. * indicates significant at the $p < .05$, ** indicates significant at $p < .01$.

Variable	(8)	(9)	(10)	(11)	(12)
freeagent_code	-0.384 (-1.91)	-0.218 (1.12)	-0.127 (-0.68)	-0.127 (-0.68)	0.064 -0.65
move		-0.78 (-5.46)**	-0.318 (-1.95)	-0.335 (-1.73)	-0.01 (2.09)*
tenure			0.137 (5.19)**	0.137 (5.19)**	0.071 -0.42
tenxmove				0.007 (0.16)	-0.17 -0.99
exp_lag					0.022 -0.8
exp_sq_lag					-0.052 -1.44
c_gp_lag2					0.006 (2.96)**
s_gp_lag					0.132 (10.22)**
dnp_lag	-3.999 (-18.34)**	-4.084 (-19.92)**	-4.054 (-20.71)**	-4.054 (-20.71)**	-1.886 (9.24)**
draft_inv					0.14 -0.54
c_team_winper_lag2					2.45 (4.07)**
Constant	13.861 (161.33)**	14.19 (142.19)**	13.584 (88.92)**	13.584 (88.94)**	11.544 (20.89)**
Observations	750	749	749	749	594
Pseudo R-squared	.29	.31	.32	.32	.43

NOTE: The variables are defined as follows: freeagent_code = 1 if the player is the first year of a free agent contract; move = 1 if a player is playing for a new team this season; tenure = the length of time a player has been with his current team; tenxmove = the interaction term equal to move * tenure.

3.4 – Predicting Future Performance

Much of the behavioral literature of the psychology of prediction focuses on the notion that decision makers do not incorporate chance or statistical theory into their predictions. Daniel Kahneman and Amos Tversky, two experts in the field of behavioral decision making, argue that many decision makers rely upon heuristics and other a priori information to make predictions about future events. Kahneman and Tversky (1973) find that the utilization of such techniques often leads to severe and systematic errors in prediction. Thus, the purpose of this section is to analytically assess the predictive accuracy of NFL franchises' hiring practices. In order to accomplish this, we shall compare the results of an ex-ante and an ex-post compensation model. In simplest terms, an ex-ante model reflects those variables or attributes being paid for while an ex-post model illustrates the variables or attributes received in the future.

At the heart of the matter is the question of whether or not past performance is a good indicator for future performance. Since it has been shown that NFL teams pay players for past performance, it can be ascertained that they feel confident in the predictive capabilities of performance. This hypothesis can be tested empirically by modeling compensation as a function of strictly on-field performance variables.

Regression 13 is the ex-ante model of compensation as a function of only performance variables conditioning for experience. The model includes the four major performance statistics of interest – games played, winning percentage, QB rating, and Pro Bowl – from both last season and from a player's career two seasons prior. The direction of the coefficients on all of the variables is in line with the predictions made earlier. Teams traditionally pay more for quarterbacks who have been on the field and have produced in the past. In addition, the “recency effect” and “label effect” hypotheses established above seem to prevail in the ex-ante model as well. The number of games played and a player's QB rating from last season have more significance than those career variables lagged two seasons satisfying the recency claim for direct performance. Conversely, both a quarterback's career winning percentage and his Pro Bowl frequency lagged two seasons are more significant than those variables last season are in the model. Given the categorical nature of both these statistics, the label effect seems to hold.

Regression 14 is the ex-post model of compensation as a function of performance that season conditioning for experience. To maintain symmetry, this model also includes the four major performance statistics of interest. The results are somewhat contrary to their predicted values thereby strengthening the claim that franchises may be overconfident in their ability to forecast performance into the future. The number of games played is significant and positive in compensation as expected. The winning percentage variable is negative in compensation in the ex-post model. While this result is contrary to its predicted direction, the variable is difficult to interpret in its current form. A better variable would be the difference in winning percentage this season from last season. In such a form, the realized gain from paying more for a quarterback can be better estimated. The quarterback rating variable, while positive, is insignificant suggesting that franchises do not realize the benefits of rating. Finally, Pro Bowl with its positive relationship acts as predicted and is significant.

Table 8. Ex-Ante versus Ex-Post

The dependent variable is the log of the adjusted salary. For the Ex-ante model, each observation represents either the player-year or career-aggregated value of that variable at a given point in time conditional on a player playing the current season. For the Ex-post model, each observation represents the current player-year value for that variable. Observations for the both regressions include only players who were on a roster after 1992 and have more than 5 years of experience. * indicates significant at the $p < .05$, ** indicates significant at $p < .01$.

Variable	(13)	(14)
exp_lag	-0.011 (0.13)	
exp_sq_lag	-0.003 (-0.67)	
s_gp_lag	0.078 (7.69)**	
c_gp_lag2	0.001 (0.67)	
s_team_wimper_lag	0.190 (0.73)	
c_team_wimper_lag2	1.304 (2.97)*	
s_rating_lag	0.003 (1.18)	
c_rating_lag2	0.004 (0.99)	
s_bowl_lag	0.349 (2.27)*	
c_per_bowl_lag2	1.472 (4.59)**	
exp		0.113 (1.15)
exp_sq		-0.005 (-1.23)
s_gp		0.092 (8.29)**
team_wimper		-0.428 (-1.44)
s_rating		0.000 (0.14)
s_bowl		0.506 (2.62)**
Constant	12.406 (22.84)**	13.156 (22.24)**
Observations	264	276
Pseudo R-squared	0.25	0.11

Two observations are particularly interesting when comparing the ex-ante and ex-post regressions. The first result worth noting is the differences in the direction and significance of the coefficients between the ex-ante and ex-post models. If NFL franchises were able to perfectly calibrate their hiring decisions, those variables on which teams pay (ex-ante) would be the same variables that are evident when determining on-field production the following season (ex-post). While NFL teams pay for games played, passer rating, winning percentage, and Pro Bowls, only games played and Pro Bowl behave in the predicted manner.

The reasoning behind the emerging disconnection between past and future performance is either the function of structural causes within the NFL or implies that past performance is a poorer indicator for future performance than expected. It is not surprising that a team gets increasing playing time from a player with a higher wage. At the highest level, it is the franchise itself that controls which players get playing time. The front office can exert pressure on the coaching staff and push for the insertion of any player into the lineup. Thus, the results of the ex-ante/ex-post comparison of the role of games played is as expected; teams with higher salaried quarterbacks on their roster have a greater propensity to put that player on the field over other less paid ones.

On the face of it, the results of the above models with regard to Pro Bowl appearances contradict the claim that past performance is indeed a poor predictor of future performance. However, an argument similar to the one made for games played can explain why the occurrence of Pro Bowl persists over time. While franchises themselves have no real power in ensuring that any of their players return to the Pro Bowl in a given year, the nature of the selection process lends itself to repeat all-stars. Given the infrequency of Pro Bowl quarterbacks relative to all others at the position, the “label” value of being named to the Pro Bowl is substantial. It has already been determined that the impact of past Pro Bowls is significant on salary several seasons later. Thus, while the statistical performance of a quarterback may not necessarily warrant his selection to the Pro Bowl, the fact that he was selected last year works in a player’s favor. The ex-post model corroborates such an explanation. In theory, Pro Bowl selection should be linked to a player’s passer rating and the winning percentage of his team. Such theory is confounded by the fact that while past Pro Bowls may predict future Pro Bowls, past QB rating and winning percentage have no such discernible effect.

The second key observation from the ex-ante/ex-post model comparison is the disparity in the pseudo R-squared values. While the ex-ante model is far from ideal, the inability of the ex-post model to explain even 15 percent of the variation supports the idea that franchises are overconfident in their predictive abilities. The forecasting of on field performance is much like that of mutual fund returns. The primary finding from this section is that teams continue to pay quarterbacks as though past performance is an adequate measure for future performance despite the empirical evidence that suggests that it is not.¹⁸

¹⁸ The divergence between the ex-ante and ex-post models is much more apparent when the two are run using OLS. While the coefficients and significance levels of all the variables remain similar to those in the Tobit models, the R-squared values separate a great deal more (R-squared = 0.55 for ex-ante; R-squared = 0.23 for ex-post).

4.1 – A Discussion of the Results

The purpose of this section is to discuss the analysis' results. However, in order to systematically discuss the implications of the models developed in this study, this discussion will be divided into the three main objectives of this paper, and proceed in the respective order of their presentation.

The first major result of this study is the value gained from the methodological enhancements implemented throughout the model building process. The dual improvements of incorporating decay into the model as well as the usage of a Tobit regression to avoid selection problems proved to be beneficial. Future investigations into related markets should encapsulate both these modifications into their analysis. Indeed, the nature of the hiring process and its dependence upon past performance raises questions about the memory capacity of hiring executives. Whether or not last season's performance statistics are given more weight than statistics from earlier on is a fundamental inquiry into the nature of the NFL quarterback market. Moreover, the importance of uncoupling the most recent season's performance from career performance was magnified when the Tobit model was employed.

Tobit models were utilized in this study to eliminate the significant selection problems that pervade the NFL labor landscape. Given the existence of a salary cap and minimum salary in the NFL along with the potential for players to leave the league and return, the observed data are only part of the entire consideration set. By employing the Tobit model in place of typical OLS, a great deal of importance was placed on whether or not a quarterback was in the league the year before. Not surprisingly, teams significantly devalued those quarterbacks who were left of the regular season roster in a past season the following year. As a result of this devaluation, the usage of the Tobit model also gave greater relative weight to recent performance over career performance.

The second major set of results of this study is the implications it has for free agency in the NFL. Conventional wisdom and basic economic theory would suggest that increased labor mobility would enhance the wage prospects for players who reach the open market. Based upon the analysis conducted in this investigation, free agency may indeed enhance wage prospects in the NFL. However, the impact is indirect rather than direct with those players who never reach the free agent benefiting much more than those who do. Such a provocative result occurs primarily because those players who reach free agency perform more poorly than those who re-negotiate their contracts before expiration. Thus, while the conditional effects of free agency may be mixed, the unconditional impact is negative. The moment a player becomes a free agent, his value is decreased because his current team decided not to re-sign him before he gained access to the market.

The third implication of this study revolves about the predictive capacity with which NFL franchises make their hiring decisions. A comparison between ex-ante and ex-post models suggests that NFL executives are overly reliant upon past performance in their

prediction of future performance. In the two areas in which predictive accuracy was found - games played and Pro Bowl - actual on-field performance is secondary to other considerations. The franchise itself controls the number of games in which a player plays. In order to recoup their lost wages, a franchise would seek to maximize the overall production potential by providing as many opportunities as possible to a highly paid QB. Furthermore, the selection process of the Pro Bowl is highly based upon past Pro Bowl selections and independent of actual performance.¹⁹

The inability of NFL franchises to accept and incorporate the lack of predictive power past performance has on future performance can lead to an implicitly flawed compensation system based upon unrealizable future productivity. However, such a phenomenon is not limited to the hiring process in the NFL and can be generalized to many labor markets. While hiring decisions are often based upon past performance and a candidate's skill set, organizations would be better served to understand the large degree of uncertainty that is associated with future output and modify their compensation frameworks accordingly.

5.1 – Concluding Remarks

As stated in the introduction of this paper, the aim of this investigation was two-fold. The first aim was to provide a working blueprint for future analysis into the NFL and related markets. This study should represent one of many urging executives and front office managers to utilize an empirically driven hiring system. The second aim was to address three major objectives pertaining to and going beyond the NFL. Through the improvement of past compensation methodology, market specific inquiry into the nature of free agency, and ex-ante/ex-post performance analysis, this investigation opens up several avenues for future research.

Further research into the nature of the NFL QB market may take into account any of the following four refinements. These refinements may serve as separate inquiries in their own right but also would make the current analysis more robust. The first modification revolves about the methodological change from OLS to Tobit regressions to correct for the selection problems prevalent in the league. Future analysis using Tobit models should better classify those players who did not play the previous season. As seen in the models developed in this paper, the inclusion of such a variable precludes the usage of some season statistics. It is advocated that a system that distinguishes between players who were not on a roster and those who simply did not play during the season be created.

A second improvement mentioned in the analytical portion of this paper involves distinguishing between team winning percentage and a quarterback's winning percentage. A tangential modification is the creation of a variable that reflects the change in team winning percentage from year to year. The inclusion of both these items would greatly enhance the interpretation of results from the compensation models constructed. Given the great importance sports franchises place on winning, the optimal hire would lead a team to an improved record the following season. Moreover, researchers would be able to answer

¹⁹ This potential for unwarranted autocorrelation in Pro Bowl selections lends itself to further research.

whether winning quarterbacks are paid more or if winning franchises pay more for their quarterbacks.

A third enhancement to the framework presented in this study would be to include a categorical analysis of compensation. It is probable that NFL executives do not always depend upon performance statistics in their hiring decisions. A more comprehensible compensation system would categorize players into different bins and determine the value of each of these bins to a franchise. For example, quarterbacks may be classified as either Pro Bowlers, above average starters, below average starters, backups, or deep backups. By calculating the value a team has for each of those categories and then examining a player's historical probabilities of falling into each of the categories, his overall value can be ascertained. The difficulty with using categorical variables arises when attempting to include decay in the regression. Weighting the importance of categorical performance over time is an econometrically difficult task but is doable.

A fourth and final consideration for studying the NFL labor market is conditioning the value added of a single player on the performance of his teammates. Football is the ultimate team game and a quarterback's performance is highly dependent upon the quality of his offensive line, wide receivers, and running backs. Numerous immensely talented quarterbacks have performed poorly when surrounded by sub par talent. Conversely, many marginal QBs have made a career by simply fulfilling a certain role within an offense loaded with "weapons." While an accurate determination of the expected role of a player is nearly impossible to determine econometrically, conditioning upon the individual performance of a QBs teammates may be a step in the right direction.

Author's Note:

Comments and suggestions on this investigation are welcome. The author can be reached at wb3@duke.edu.

Works Consulted

- Ahlburg, D., & Dworkin, J. (1991). Player Compensation in the National Football League. In Paul D. Staudohar and James A. Mangan (Ed.), *The Business of Professional Sports*. University of Illinois Press, 1991.
- Epple, D. (1987, February). Hedonic Prices and Implicit Markets [Electronic version]. *The Journal of Political Economy*, 95(1), 59-80.
- Hendricks, W., DeBrock, L., & Koenker, R. (2003). Uncertainty, Hiring, and Subsequent Performance: The NFL Draft [Electronic version]. *Journal of Labor Economics*, 21(4), 857-886.
- Jones, J.C.H., & Walsh, W.D. (1988, July). Salary Determination in the National Hockey League: The Effects of Skills, Franchise Characteristics, and Discrimination [Electronic version]. *Industrial and Labor Relations Review*, 41(4), 592-604.
- Kahn, L.M. (1992, January). The Effects of Race on Professional Football Players' Compensation [Electronic version]. *Industrial and Labor Relations Review*, 45(2), 295-310.
- Kahneman, D., & Tversky, A. (1973). On the Psychology of Prediction [Electronic version]. *Psychological Review*, 80, 237-251.
- Martin, E.J. (1999) A Bilateral Labor Market: Salary Determinants of National Football League Quarterbacks [Electronic version]. Working Paper. Virginia Polytechnic Institute and State University.
- Melvin, P.C., & McCormick, R.E. (2001, February) On the Demand and Supply of Golf [Electronic version]. Working Paper.
- Rosen, S. (1986). The Theory of Equalizing Differences [Electronic version]. *Handbook of Labor Economics*, 1, Chapter 12.
- Scahill, E. (1985, March). The Determinants of Average Salaries in Professional Football. *Atlantic Economic Journal*, 13, 103.
- Scott Jr., F.A., Long, J.E., & Somppi, K. (1983). Free Agency, Owner Incentives, and the National Football League Players Association. *Journal of Labor Research*, 4(3), 257-264.
- Scully, G.W. (1974). Pay and Performance in Major League Baseball [Electronic version]. *American Economic Review*, 66(6), 915-930.