

Catching a draft: on the process of selecting quarterbacks in the National Football League amateur draft

David J. Berri · Rob Simmons

Published online: 18 September 2009
© Springer Science+Business Media, LLC 2009

Abstract The reverse order college draft gives the worst teams in the National Football League (NFL) the opportunity to hire the best amateur talent. For it to work effectively, teams must be able to identify the “best” talent. Our study of NFL quarterbacks highlights problems with the draft process. We find only a weak correlation between teams’ evaluations on draft day and subsequent quarterback performance in the NFL. Moreover, many of the factors that enhance a quarterback’s draft position are unrelated to future NFL performance. Our analysis highlights the difficulties in evaluating workers in the uncertain environment of professional sports.

Keywords Quarterback · College · Draft · Performance

JEL Classification L83

1 Introduction

In 1935 the Brooklyn Dodgers and Philadelphia Eagles of the National Football League (NFL) (ProFootballReference.com) entered into a bidding war for the services of fullback Stan Kostkas, a fullback with the University of Minnesota.¹ When the bidding was done, Kostkas had

agreed to a \$5,000 contract with the Dodgers. This contract rivaled the pay of Bronco Nagurski, the player many considered the best player in the NFL.

This bidding war led the owner of the Eagles—Bert Bell—to propose the reverse-order draft. Specifically, Bell argued that NFL teams should not compete for the services of college talent. Bell proposed that NFL teams should choose college players, and the worst teams from the previous season should get to choose first.

Such a structure clearly benefitted Bell. Bell’s Eagles were the worst team in 1935, and consequently when the draft was instituted in 1936, Bell got to choose first. Despite Bell’s obvious self-interest, defenders of the reverse-order draft have seen this institution as key to the economic health of a professional team sport. Specifically, the reverse-order draft is considered a mechanism to enhance a league’s competitive balance. By funneling the best amateur talent to the worst teams, the worst teams are given an opportunity to improve. Hence the differences between the best and worst—at least theoretically—should lessen over time.

Although research on the impact a draft has on competitive balance is hardly encouraging,² the draft has been shown to have a clear benefit to the league’s owners. Studies have shown that the draft does depress salaries. Evidence from North American major sports leagues reveals that drafted players suffer monopsonistic exploitation by team owners, with pay set below their contributions to team revenues (Scully 1974; Krautmann 1999; Kahn

D. J. Berri
Department of Economics and Finance, Southern Utah
University, School of Business Building #318, 351 West
University Boulevard, Cedar City, UT 84720, USA
e-mail: berri@suu.edu

R. Simmons (✉)
Department of Economics, The Management School, Lancaster
University, Lancaster LA1 4YX, UK
e-mail: r.simmons@lancaster.ac.uk

¹ The story of the birth of the NFL Draft is reported in Quirk and Fort (1992, pp. 187–188). This story was also noted in Leeds and Von Allmen (2008, p. 163), Fort (2006, p. 258), and Quinn (2008).

² Quinn (2008) also reviewed research on the impact the draft has had on competitive balance. This research indicates there is little relationship between a reverse order draft and the level of competitive balance.

2000; Krautmann et al. 2009; Lehn 1982); see Quinn (2008) for a good summary of the operation and effects of player draft systems in the four major North American sports leagues. This implies that, if a team does acquire a very productive player, a team should be able to acquire that production at a discount.

The focus of our research will be the NFL. Given that a draft theoretically should allow a team to acquire production at a discount, it is natural to wonder how good NFL teams are at finding productive players.

2 Two prior studies

We are not the first to ask such a question. Both Cade Massey and Richard Thaler (2005) and Hendricks et al. (2003) offered examinations of the NFL draft. Before moving on to our approach, we need to review the analysis offered in these papers.

2.1 Massey and Thaler (2005)

The Massey and Thaler paper is essentially an analysis of decision-making in the NFL. Specifically, these authors considered the surplus value of a draft pick, or the difference between the projected economic value of a pick and the compensation cost of the player. This research indicated that surplus value peaked in the second round of the draft. In other words, the top picks in the draft were overvalued by decision-makers in the NFL.

Massey and Thaler argued that the overvaluation of first round picks was due to a combination of non-rational expectations by team owners and mis-pricing of players. At the root of player over-valuation was an inability on the part of team managers to successfully predict the performance of players in the NFL. Again, overvaluation is estimated as the ‘surplus value’ of each draft pick. Surplus value is specifically determined by comparing the monetary value of each veteran (free agent) in the NFL labor market to the actual compensation of drafted players. Utilizing data from 1991 to 2002, Massey and Thaler find that “surplus value increases at the top of the order, rising to its maximum of \$750,000 in the top half of the second round before declining through the rest of the draft. Consequently, as noted, the treasured first pick in the draft is, according to this analysis, actually the least valuable pick in the first round! To be clear, the player taken with the first pick does have the highest expected performance (that is, the performance curve is monotonically decreasing), but he also has the highest salary, and in terms of performance per dollar, is less valuable than players taken in the second round.” (Massey and Thaler 2005, p. 25).

Given Massey–Thaler’s findings, one might expect teams that are picking at the top of the first round to do everything they can to trade down. But teams often take the opposite approach. A premium is often paid by teams to move up from the first round. An example of this practice is cited by Massey and Thaler. In 2004, the New York Giants and the San Diego Chargers completed a trade where the Giants were given the first pick in the draft and the rights to sign quarterback Eli Manning (who later became Most Valuable Player in the 2008 Super Bowl). In exchange, the Chargers acquired the fourth pick, or the rights to quarterback Philip Rivers. In addition to the rights to Rivers, the Chargers were also given a third round pick in 2004 and one-first round and one-fifth round pick in 2005. Hence, in order to move up just three spots in the draft, the Giants surrendered three additional picks.

The Massey–Thaler study is certainly impressive, yet not without its flaws. The authors essentially compare two evaluations of NFL players. The first evaluation they consider is before the draft and is essentially revealed by where a player is taken on draft day. The second evaluation they make is after the draft, and consists of games played, games started and Pro Bowl appearances in the NFL. Massey and Thaler find that these evaluations are not consistent.

There is a problem, though, with comparing the two evaluations. The formal empirical analysis offered by Massey and Thaler uses surplus value of a draft pick as the dependent variable in their regression model. This conflates the expectation of player performance with monetary compensation for that performance. Teams are essentially predicting the success of a drafted player in the NFL and pricing that player’s expected services over the duration of initial contract.

To understand the problem with this approach one needs to consider the empirical link between the draft and player salary. Specifically, salary models for NFL positions are considered more formally in a pair of companion papers by the authors (Berri and Simmons 2009; Simmons and Berri 2009). Those papers analysed the reported pay of NFL quarterbacks and running backs, respectively, and found evidence of substantial positive salary premia, after controlling for player experience and productivity, for players drafted in rounds one and two, but not thereafter. These premia persisted over a player’s career up to qualification for free agency or up to a trade from the drafting team, whichever was earlier. In sum, where a player is drafted impacts his future pay. And therefore it is hard to use data on future pay to evaluate the quality of the initial draft choice.

Beyond this issue, we would also note that the Massey–Thaler paper attempted to look at all positions on the field. Consequently they had to look at measures that would apply to all positions. As noted above, these authors considered

games played, games started, and the probability of making a roster or making the Pro Bowl. The problem with each of these metrics is that they are really statements from decision-makers about who is better or worse. So the authors are essentially using one statement from decision-makers to evaluate the rationality of another statement.

2.2 Hendricks et al. (2003)

This same criticism would apply to a similar paper by Hendricks et al. (2003). Their paper, though, had a different theoretical orientation to Massey and Thaler. Hendricks et al. argued that draft choices reflect different kinds of uncertainty about a drafted players' future productivity. In the college system, schools are divided into different divisions. At the top is a collection of schools—historically referred to as Division IA—labeled the Bowl Subdivision. Within the Bowl Subdivision we find schools with significantly greater resources to spend on coaching and training facilities.

Beneath this grouping is the Championship Subdivision, as well as Division II and Division III. These schools are less homogenous, and hence performance of players is harder to evaluate. There are two main consequences of this separation in the pool of talent. First, players from the non-Bowl Subdivision schools may be disadvantaged in the draft process since there is greater uncertainty over the reliability of their college performance measures. This is an example of statistical discrimination. On the other hand, some teams may be willing to take members of the minority non-Bowl Subdivision group as risky propositions in the hope of finding stars that will eventually deliver a competitive edge over NFL opponents. This is an example of 'option value', following Lazear (1986).

Intriguingly, Hendricks et al. find both aspects of uncertainty to be valid in the hiring process for NFL players at entry level. If teams are choosing between two (predicted) star players out of college in the early rounds of the draft, they tend to be risk-averse and opt for athletes from the more visible Bowl Subdivision program. But in later draft rounds, non-Bowl Subdivision players are overvalued relative to top programs, supporting the option value explanation. Note, though, that athletes from non-Bowl Subdivision schools do not, according to Hendricks et al., have longer or stronger NFL careers than drafted players from superior programs. This reflects the inherent uncertainty surrounding draft choices.

Like Massey and Thaler, Hendricks et al. sought to examine all positions. Consequently Hendricks et al. employed such variables as years played in the NFL, a dummy variable indicating that a drafted player actually appeared in the NFL, and percentage of player's active NFL years that he appeared in the Pro Bowl as dependent variables in their analysis of the NFL draft. None of these

variables, though, are direct measures of player performance and as we previously noted, lead the authors to use one evaluation by decision-makers to access the validity of another evaluation.

2.3 Our approach

We wish to take a different approach to the above-cited papers on NFL draft. Specifically, we seek to offer an evaluation of decision-making that is less dependent on actual decisions. In other words, we wish to use explicit player performance measures.

Such an approach in football, though, is problematic. Performance metrics vary across positions. Consequently, to properly address decision-making one would have to adjust one's analysis to each position examined. Given this reality, we will focus our study on just one position, the NFL quarterback.

The quarterback is the only position in football that is credited by observers of the game directly with wins and losses. As a consequence, the player who starts in this position is often thought of as the face of the franchise. The importance of this position suggests that teams would devote most of their decision-making resources to get the choice of quarterback "right."

To assess whether or not this is true, we will investigate the following questions:

- What is the relationship between a quarterback's draft position and his subsequent performance in the NFL?
- What factors do NFL teams consider in drafting a quarterback?
- How do the factors the NFL teams consider in drafting a quarterback related to subsequent performance?

The answer to these questions will begin with a discussion of how performance of quarterbacks is measured. We will then examine the link between draft position and NFL performance. This will be followed by a model designed to predict draft position and an additional model linking what is known on draft day to performance in the NFL. Concluding observations will close the paper.

3 Measuring the performance of quarterbacks

To study quarterbacks in the NFL one must first have a measure of player performance. And the most commonly cited statistic is the NFL's quarterback rating measure. But as the following equation reveals, it is hardly a simple or intuitive metric.³

³ ESPN.com, as well as other web sites, reports the equation for the NFL's quarterback's rating.

$$\left(\frac{\text{COMP}}{\text{PASSATT}} - 0.3 + \frac{\text{PASSYDS}}{\text{PASSATT}} - 3 + \frac{\text{PASSTD}}{\text{PASSATT}} + \frac{0.095 - \frac{\text{INT}}{\text{PASSATT}}}{0.04} \right) \frac{100}{6} \quad (1)$$

where COMP = Completions, PASSYDS = Yards passing, PASSTD = Touchdown passes thrown, INT = Interceptions thrown, PASSATT = Passing attempts.

This rating was devised by Don Smith in 1971, at the behest of NFL commissioner Pete Rozelle. Smith wanted to derive a rating for quarterbacks that was independent of how other quarterbacks performed. To do this, Smith began with the four available measures of quarterback passing performance: pass completion rate, pass yards, touchdowns and interceptions. He then assumed that average performance in each measure would score one point, spectacular performance would earn two points while poor performance would earn zero points. Smith then experimented with a set of weights of the four measures until he eventually arrived at the above formula. This formula—despite its opaqueness—remains prominent in telecasts and other media coverage to this day. This is primarily because the NFL has given official approval to the measure.

Although the calculation of this measure lacks intuition, the quarterback rating does have some meaningful content. The four selected passing indicators are all likely to be correlated (positively with the exception of interceptions per attempt) with team wins. But there are two major flaws with the rating measure. First, the weights are arbitrarily imposed and may be inappropriate. Second, quarterbacks do occasionally run (scramble) with the football and this dimension of performance is disregarded altogether in the rating measure.

Given these issues, we will also employ a metric initially detailed in Berri et al. (2006) and Berri (2007). These works began with empirical models designed to explain both points scored and points surrendered in the NFL. Such models were employed to assess a quarterback's marginal physical product, i.e., the contribution of a quarterback to team wins. Such an approach follows in the tradition of Blass (1992), who assessed the value of baseball hitters to their teams by regressing runs scored by hitters on a set of batter statistics (including singles, doubles, stolen bases, walks, etc.). Two models were employed to measure team offense and team defense in the NFL. For offense, the team performance is split into four stages of the game: offensive ball acquisition, offensive ball movement, offensive ball retention and offensive scoring. Each stage comprises a set of independent variables. Evaluation of defensive performance follows in similar manner. It is important to note that the imputed effects of underlying covariates on the two quarterback metrics (net points and quarterback score) are

derived from regression coefficients rather than just imposed as Don Smith did for his quarterback rating.

The models of Berri et al. (2006) and Berri (2007) were used to derive the value in terms of net points and wins, of passing yards, rushing yards, passing attempts, rushing attempts, sacks, interceptions, and fumbles lost. The value of each factor—in terms of net points and wins—is reported in Table 1. These values can be used to measure a quarterback's contribution to both scoring—or Net Points—and Wins Produced. The calculation of each measure simply involves multiplying the values—associated with either Net Points or Wins Produced—in Table 1 by each quarterback's production of each statistic.

Another metric, QB Score, is even easier to calculate.

$$\text{QB Score} = \text{All Yards} - 3 \times \text{All Plays} - 30 \times \text{All Turnovers} \quad (2)$$

where

$$\text{All yards} = \text{Passing yards} + \text{Rushing yards} - \text{Yards lost from sacks}$$

$$\text{All Plays} = \text{Passing attempts} + \text{Rushing attempts} + \text{Sacks}$$

$$\text{All Turnovers} = \text{Interceptions} + \text{Fumbles lost}$$

(Data on fumbles lost is only tabulated by Yahoo.com from 1994 to the present. So for this study, only interceptions are used to calculate QB Score, Net Points, and Wins Produced.)

The simpler measure is derived from normalizing the value of plays and turnovers around one yard. For example, as seen in Table 1, each play is worth about three yards and each turnover costs about 30 yards.⁴ As noted in Berri (2007), the correlation between QB Score per play (or QB Score divided by all plays) and Net Points per play (or Net Points divided by All Plays) is 0.98.

All of the measures of quarterback performance considered so far are infected by the performance of teammates. For example, fumbles lost will impact QB Score (but not quarterback rating). The fumble could be attributable to the center who delivers (snaps) the ball to the quarterback, to a running back who fails to take the ball from the quarterback, or the quarterback himself; or even some combination of the three. Similarly, the responsibility of interceptions will be shared between quarterback and receiver. Sacks, where opponents hit the quarterback behind line of scrimmage before release of the football

⁴ Specifically, from Table One we see that a play that does not produce any yards will cost a team -2.7 points. An interception will cost a team 34.5 yards while the cost of losing a fumble is 36.4 yards. As noted in Berri (2007), the value of 30 for a turnover is chosen for simplicity.

Table 1 Value of various quarterback statistics in terms of net points and wins

Variable	Value of each variable in terms of net points	Value of each variable in terms of wins
Yards (rushing yards and passing yards)	0.08	0.002
Plays (rushing attempts, passing attempts, and sacks)	−0.21	−0.006
Interceptions	−2.7	−0.078
Fumbles lost	−2.9	−0.082

Source: Berri et al. (2006) and Berri (2007)

causing loss of yards and down, will be influenced by the offensive line and opposition defense. All of these interactions indicate that it’s not possible to convincingly isolate quarterback performance from other team members in a simple performance metric. This point needs to be remembered as we progress in our analysis.

With our metrics in hand we can now turn to the evaluation of the NFL draft. And this evaluation starts with a simple question: Do quarterbacks taken higher in the draft out-perform those taken later?

4 Draft position and NFL performance

We begin with season data on quarterbacks that met the following criteria⁵:

- the quarterback was drafted from 1970 to 2007
- the quarterback was chosen between picks 1 to 250
- the quarterback played in at least one game in the season

In all we have 1,943 season observations drawn from 331 different quarterbacks.⁶ This sample was divided into

⁵ The NFL and AFL merged before the start of the 1970 season. According to ProFootballReference.com, from 1970 to 1976 the NFL’s draft consisted of 17 rounds and at least 442 picks. From 1978 to 1992 the draft was only 12 rounds (and from 330 to 336 picks). For the 1993 season the NFL draft was eight rounds and 224 picks. After the 1993 season the draft was only seven rounds. In 1994 there were only 222 picks. But after 1994 the number of picks exceeded 250 (but never exceeding 262 picks) in all but three seasons. Consequently we settled on a cut-off of pick 250 for our study. Quarterbacks chosen after 250 were not considered for this examination.

⁶ Our NFL performance data on quarterbacks was taken from sports.yahoo.com. (<http://sports.yahoo.com/nfl/stats/byposition?pos=QB>). The NFL has changed quite a bit since 1970. Consequently, to compare quarterbacks across this time period one has to adjust the numbers. Specifically, we calculated a quarterback’s relative performance with respect to each statistic. This calculation began by calculating the average performance in each statistic from 1970 to 2007. Then in each year we subtracted the average in that statistic from that season from each quarterback’s performance in that statistical measure. We then added the average performance across the entire period. For example, in 1975 Terry Bradshaw’s net points per play was 0.162. The average quarterback in 1975 posted a net points per play mark of 0.088 while the average mark from 1970 to 2007 was 0.144. Given these numbers, Bradshaw’s relative net points per play in 1975 was

five, roughly equal, segments. For these five segments we measured the quarterback’s performance with respect to QB Score, Net Points, Wins Produced, and the NFL’s QB Rating (as well as the elements of this rating).

The results, reported in Table 2, indicate that with respect to the aggregate measures, where you are chosen impacts the results. The lower a quarterback’s draft status (with the number one pick being the lowest), the better aggregate and per game numbers we see.

When we look at per-play numbers, though, the story changes. On a per play basis, quarterbacks chosen with picks 11–50, as well as picks 51–90, outperform quarterbacks chosen in the top 10. Such a result suggests that top 10 quarterbacks really don’t offer more, they just get to play more. And the story doesn’t change when we look at the NFL’s QB Rating measure (which is a per pass attempt metric). Whether we look at the aggregate QB Rating measure, or the elements of the QB Rating metric, the same story is told. Quarterbacks taken in the top 10 are outperformed by signal callers taken from 11 to 90.

Although the numbers reported in Table 2 suggest that quarterbacks taken at the top of the draft are not any better than those taken later, there is another possibility. This is a reverse order draft, so those chosen in the top 10 are going to relatively poor teams. Perhaps the quality of teams hiring the top ten picks is lowering their numbers.

To examine this possibility we took two different views of the relationship between a quarterback’s performance at different points of their career and the quarterback’s draft position. First we looked at how a quarterback’s performance at each year of experience. For each year, only quarterbacks who logged at least 100 plays were considered. As Table 3 indicates, the strongest correlation remains between pick and plays. The per play metrics, as well as the per pass attempt measures, are all very weakly correlated with where a player was chosen.

In Table 4 we take a slightly different approach. Instead of looking at a quarterback’s performance at each year of experience, we considered his aggregate performance after

Footnote 6 continued
0.220, or $[(0.162 - 0.088) + 0.144]$. It is these relative numbers that were used in our analysis of quarterbacks from 1970 to 2007.

Table 2 Performance of NFL quarterbacks chosen at different points in the NFL draft years: 1970–2007

Picks	Observations	Games	Plays	QB score	Net points	Wins
Picks 1–10	396	4,370	131,965	217,399	19,004	485.2
Picks 11–50	400	3,993	108,765	185,866	16,204	414.4
Picks 51–90	372	3,190	72,958	122,239	10,659	272.2
Picks 91–150	413	3,298	68,689	103,575	9,073	230.0
Picks 151–250	362	2,887	54,293	86,734	7,567	192.5
Picks	Observations	QB score per game		Net points per game	Wins per game	
Picks 1–10	396	49.7		4.3	0.111	
Picks 11–50	400	46.5		4.1	0.104	
Picks 51–90	372	38.3		3.3	0.085	
Picks 91–150	413	31.4		2.8	0.070	
Picks 151–250	362	30.0		2.6	0.067	
Picks	Observations	QB score per play		Net points per play	Wins per play	
Picks 1–10	396	1.647		0.144	0.368	
Picks 11–50	400	1.709		0.149	0.381	
Picks 51–90	372	1.675		0.146	0.373	
Picks 91–150	413	1.508		0.132	0.335	
Picks 151–250	362	1.598		0.139	0.354	
Picks	Observations	Completion percentage (%)	Passing yards per pass attempt	Touchdowns per pass attempt	Interceptions per pass attempt	QB rating
Picks 1–10	396	56.09	6.78	0.0396	0.0381	74.4
Picks 11–50	400	56.73	6.86	0.0419	0.0375	76.3
Picks 51–90	372	56.43	6.86	0.0404	0.0392	74.8
Picks 91–150	413	55.26	6.73	0.0386	0.0403	72.3
Picks 151–250	362	55.79	6.77	0.0385	0.0402	72.9

each year of his career. Again only quarterbacks with at least 100 plays to that point of their career were considered. And again, we do not find much of a relationship between per play, or per pass attempt measures, and where a quarterback was chosen. We do find a much stronger relationship between career plays and draft status.

All of this suggests that where a quarterback is chosen impacts how much he plays. But it does not appear strongly related to how well a quarterback plays.

Of course one wonders why draft status and performance are so weakly related. To further address this issue we next consider the factors that impact where a player is chosen in the draft.

5 Determining draft position

Every quarterback chosen in the draft played college football. At the time of the draft we know what a quarterback did on the college football field. In other words, we can measure QB Score, Net Points, Wins Produced, and the

NFL's QB rating for a quarterback's college career. And of course we also have data on all the elements of QB rating (completion percentage, passing yards per attempts, etc....).

Beyond performance, we also have data from the NFL's scouting combine. The NFL Scouting Combine (specifically called the National Invitational Camp) began in 1982 in Tampa, Florida. Since 1987 it has been held in Indianapolis, Indiana. At the combine players take medical exams, as well as both physical and psychological tests.⁷ From these tests we learn a quarterback's height, weight, how fast he runs (in the 40 yard, 20 yard, and 10 yard dash), his vertical jump, his broad jump, and how fast he runs the shuttle and cones.⁸

Quarterbacks also take the Wonderlic test. The Wonderlic test—according to Wonderlic.com—was developed by industrial psychologist Eldon F. Wonderlic in 1937. The principle purpose of this test is to assess mental agility. NFL

⁷ The history of the NFL's National Invitational Camp can be found at (<http://www.nflcombine.net/?q=node/9>).

⁸ Combine data from 1999 to 2008 can be found at nfldraftscout.com.

Table 3 Correlation between draft position and performance at different levels of experience years: 1970–2007 minimum 100 plays in year examined

Experience	QB score	Net points	Wins	QB score per play	Net points per play	Wins per play
1	-0.14	-0.15	-0.14	-0.03	-0.03	-0.03
2	-0.10	-0.11	-0.10	-0.04	-0.05	-0.05
3	-0.15	-0.16	-0.16	-0.05	-0.06	-0.06
4	-0.17	-0.17	-0.17	-0.04	-0.05	-0.05
5	-0.19	-0.20	-0.19	-0.03	-0.04	-0.04
6	-0.10	-0.11	-0.11	0.04	0.03	0.03
7	-0.08	-0.08	-0.08	-0.02	-0.03	-0.03
8	-0.22	-0.21	-0.21	-0.18	-0.18	-0.18
9	-0.20	-0.20	-0.20	-0.05	-0.06	-0.06
10	-0.16	-0.16	-0.16	-0.12	-0.13	-0.13

Experience	Plays	Completion percentage	Passing yards per pass attempt	Touchdowns per pass attempt	Interceptions per pass attempt	QB rating
1	-0.26	0.03	-0.05	0.07	0.03	0.01
2	-0.21	-0.07	0.04	0.16	0.12	0.00
3	-0.29	-0.09	-0.02	-0.14	0.09	-0.12
4	-0.25	-0.07	-0.03	-0.19	0.10	-0.14
5	-0.32	-0.04	0.07	0.01	0.11	-0.04
6	-0.23	0.02	0.09	-0.05	0.09	-0.03
7	-0.14	0.02	0.02	-0.12	0.11	-0.09
8	-0.22	-0.20	-0.19	-0.15	0.05	-0.20
9	-0.31	-0.02	0.01	0.02	0.13	-0.07
10	-0.19	0.05	-0.06	-0.08	0.27	-0.16

Table 4 Correlation between draft position and career performance at different levels of experience years: 1970–2007 minimum 100 plays in year examined

Experience	QB score	Net points	Wins	QB score per play	Net points per play	Wins per play
2	-0.23	-0.24	-0.23	-0.10	-0.10	-0.10
3	-0.22	-0.23	-0.23	-0.01	-0.01	0.00
4	-0.25	-0.26	-0.26	-0.05	-0.06	-0.06
5	-0.26	-0.27	-0.26	-0.01	-0.02	-0.02
6	-0.29	-0.30	-0.30	-0.06	-0.06	-0.07
7	-0.29	-0.30	-0.29	-0.01	-0.01	-0.01
8	-0.29	-0.30	-0.30	-0.03	-0.04	-0.04

Experience	Plays	Completion percentage	Passing yards per pass attempt	Touchdowns per pass attempt	Interceptions per pass attempt	QB rating
2	-0.38	-0.05	-0.10	0.06	0.05	-0.03
3	-0.40	-0.06	-0.04	-0.04	0.09	-0.05
4	-0.40	-0.01	-0.01	-0.08	0.02	-0.08
5	-0.43	-0.01	-0.02	-0.12	0.00	-0.06
6	-0.46	0.00	0.00	-0.11	0.08	-0.10
7	-0.46	0.07	0.01	-0.12	0.02	-0.04
8	-0.46	0.05	-0.04	-0.14	0.02	-0.06

Equation (3)

quarterbacks need to read opposition defense formations and tactics in a period of just a few seconds. They also need to assess whether the original play that was called by the coaches is feasible and, if not, need to be prepared to improvise. These tasks require considerable mental capabilities. The test utilized by the NFL consists of 50 questions and must be answered in 12 min. The average score of all people who take the test (it is not just taken by NFL prospects) is 21.⁹ For the NFL quarterbacks in our data set,¹⁰ the average score was 26.1, with a range from 10 to 42.

From 1999 to 2008 there were 132 quarterbacks selected in the NFL draft. For all of these we have data on draft position, height, and weight. But the other elements of our data set were not consistently available. Specifically, for one quarterback we could not find a 40 yard dash time. For another three quarterbacks we could not locate college performance statistics.¹¹ And for eight quarterbacks there was no report of a Wonderlic test. In all, we only had complete data for 121 quarterbacks.¹²

With data in hand we first wish to see how the combine data related to performance. Specifically we regressed a quarterback's Wins Produced from his last year in college on his height, his body mass index (BMI),¹³ BMI squared, Wonderlic score, and time in the 40 yard dash. The results indicate that none of these factors are related to a

quarterback's college performance.¹⁴ Such results indicate that the combine measures are not able to capture key attributes of the quarterback.¹⁵

We then turned to the relationship between where a quarterback was chosen in the draft and both his performance his senior year as well as his combine numbers. Specifically we estimated the following model:

$$\begin{aligned} \text{PICK} = & a_0 + a_1 \times \text{Height} + a_2 \times \text{BMI} + a_3 \times \text{BMI}^2 \\ & + a_4 \times \text{Wonderlic} + a_5 \times \text{40 Yard Dash Time} \\ & + a_6 \times \text{Dummy for non-Division I-A quarterbacks} \\ & + a_7 \times \text{PERFORMANCE} + e_t \end{aligned} \quad (3)$$

where Pick = where a quarterback is chosen in the draft, Height = quarterback's height in inches, BMI = Body Mass Index, Wonderlic = Score on the Wonderlic test.

Performance is measured as

Wins Produced
Net Points
QB Score
Career Plays and Wins Produced per play
Career Plays and Net Points per play
Career Plays and QB Score per play
Career Plays and QB Rating
Career Plays, Completion Percentage, Interceptions per Attempt, and Passing Yards per Attempt

Equation (3) was estimated with data from 1999 to 2008. If we did not include any performance data our sample consisted of 124 quarterbacks. When performance is measured with Wins Produced, Net Points, and QB Score our sample falls to 121. And when we include the per play measures and Career Plays,¹⁶ our sample falls to 98. The results are reported in Table 5.

⁹ Information on the number of test questions, the time given for the test, and average score in the population was taken from an article published in *The USA Today* by Chappell (2006).

¹⁰ The Wonderlic scores we utilized were taken from NFL Quarterback Wonderlic Scores (<http://www.macmirabile.com/wonderlic.htm>). This is a website maintained by Mac Mirabile. As Mirabile notes, "... these results represent research and generally come from reliable sources, i.e., Notes from NFL scouts, newspaper articles. It is important to understand that scores cannot be "verified" since they are not released by the NFL, but rather leaked by teams or scouts."

¹¹ Data on college quarterbacks since 2000 was taken from the NCAA's website reporting Division I Football Statistics (<http://web1.ncaa.org/d1mfb/mainpage.jsp?site=org>). College data for quarterbacks selected in the 1999 and 2000 drafts was taken from CNN.SI.com.

¹² Data on 20 yard and 10 yard dash times, vertical jump, broad jump, and the shuttle and cone test was reported for fewer than 90 quarterbacks in our sample. Consequently these variables were not included in our study.

¹³ According to the Centers for Disease Control and Prevention (cdc.gov), the Body Mass Index is calculated by first dividing Weight (in pounds) by height (in inches) squared. This number is then multiplied by 703. A score of 18.5 indicates that a person's weight is below normal. A score between 18.5 and 24.9 is considered normal. A BMI from 25.0 to 29.9 indicates a person is overweight. And scores above 30.0 are indicative of an obese person. In our sample of NFL quarterbacks the average BMI score was 27.8, with a range from 24.4 to 31.5. The CDC notes that "highly trained athletes may have a high BMI because of increased muscularity rather than increased body fatness." [http://www.cdc.gov/nccdphp/dnpa/healthyweight/assessing/bmi/adult_BMI/about_adult_BMI.htm#Interpreted].

¹⁴ This is true whether we measure performance with QB Score, Net Points, Wins Produced, or the NFL's QB Rating. These results are available from the authors upon request. When we turn to per play measures of QB Score, Net Points, and Wins Produced, we do find that faster times in the 40 yard dash lead to reduced levels of per play performance (at the 10% level of significance). The adjusted *R*-squared from these regressions, though, is in the negative range and the *F*-statistic is statistically insignificant. Such results indicate that there is little relationship between the combine statistics and per play performance.

¹⁵ These results might also indicate that our Wins Produced measure of college performance is imperfect.

¹⁶ Career Plays is the number of plays a quarterback participated in throughout his college career. We only were able to collect career numbers on 105 quarterbacks taken from 2001 to the present. The inclusion of this variable was inspired by an article by David Lewin posted at ESPN.com [College Stats Don't Lie (April 17, 2008): <http://sports.espn.go.com/nfl/news/story?id=3350135>]. Lewin argued in this article that NFL performance was influenced by only two statistics, games started in college and completion percentage. Lewin's full results were not published, but he did indicate that his sample consisted of "highly drafted quarterbacks since 1996." We

Table 5 Estimation of Eq. (3) dependent variable is PICK white heteroskedasticity-consistent standard errors & covariance

Variable	1	2	3	4	5	6	7	8	9
Constant	4963.03	5069.79	5065.02	5069.76	4425.82	4446.79	4441.62	4213.52	4559.40
	3.04	3.06	3.05	3.05	2.27	2.29	2.29	2.19	2.45
Height	-19.55	-18.82	-18.85	-18.90	-15.31	-15.23	-15.26	-15.55	-14.87
	-4.24	-4.11	-4.12	-4.13	-2.51	-2.50	-2.51	-2.50	-2.39
BMI	-272.67	-277.33	-276.98	-277.06	-249.66	-251.20	-250.71	-233.46	-269.85
	-2.42	-2.40	-2.40	-2.40	-1.81	-1.83	-1.82	-1.71	-2.03
BMI Squared	4.68	4.76	4.75	4.75	4.24	4.27	4.26	3.96	4.59
	2.33	2.33	2.32	2.32	1.74	1.75	1.75	1.63	1.95
Wonderlic	-1.94	-2.09	-2.09	-2.09	-2.72	-2.71	-2.72	-2.63	-2.70
	-1.82	-2.00	-2.00	-2.00	-2.37	-2.37	-2.37	-2.26	-2.29
40 yard dash	128.81	119.65	120.07	119.94	134.91	134.57	134.67	138.43	153.21
	3.16	2.81	2.82	2.82	2.53	2.53	2.53	2.58	2.81
Division I-AA dummy	55.96	57.60	57.43	57.27	56.36	57.23	56.95	53.75	63.02
	3.31	3.23	3.22	3.21	2.30	2.34	2.33	2.15	2.64
Career plays	-	-	-	-	-0.03	-0.03	-0.03	-0.03	-0.03
	-	-	-	-	-1.98	-1.99	-1.98	-1.61	-1.95
Wins produced	-	-13.77	-	-	-	-	-	-	-
	-	-3.00	-	-	-	-	-	-	-
Net points	-	-	-0.36	-	-	-	-	-	-
	-	-	-2.95	-	-	-	-	-	-
QB score	-	-	-	-0.03	-	-	-	-	-
	-	-	-	-2.95	-	-	-	-	-
Wins produced per play	-	-	-	-	-15.09	-	-	-	-
	-	-	-	-	-2.46	-	-	-	-
Net points per play	-	-	-	-	-	-72.39	-	-	-
	-	-	-	-	-	-2.58	-	-	-
QB score per play	-	-	-	-	-	-	-191.75	-	-
	-	-	-	-	-	-	-2.54	-	-
QB rating	-	-	-	-	-	-	-	-0.79	-
	-	-	-	-	-	-	-	-1.89	-
Completion percentage	-	-	-	-	-	-	-	-	46.75
	-	-	-	-	-	-	-	-	0.35
Interceptions per attempt	-	-	-	-	-	-	-	-	1574.60
	-	-	-	-	-	-	-	-	2.62
Yards per attempt	-	-	-	-	-	-	-	-	-10.41
	-	-	-	-	-	-	-	-	-1.17
Adjusted R-squared	0.196	0.225	0.224	0.224	0.214	0.217	0.216	0.198	0.222
Number of observations	124	121	121	121	98	98	98	98	98

Table 5 reports that PICK is statistically related—at the 90, 95, or 99% level—in every formulation of the model to a quarterback’s Height, Wonderlic score, 40 yard dash time, and being a non-Division I-A player.¹⁷ Specifically

we find that taller, smarter, faster quarterbacks who play at Division I-A schools are likely to be picked higher in the draft, notwithstanding the lack of correlation of these

Footnote 16 continued
 did not have data on games started for all the quarterbacks selected since 1999, but we do think the number of career plays would be highly correlated with the number of games started in a quarterback’s career.

¹⁷ The NCAA groups teams into Division I-A (now called the Football Bowl Subdivision), Division I-AA (Football Championship Subdivision), Division II, and Division III. Of the 132 quarterbacks in our sample, only 16 did not come from a Division I-A school. Our results indicate that not playing in the Football Bowl Subdivision reduces your draft position by 56–63 slots, or nearly two rounds.

measures with our college performance measure. Additionally, body mass index and its square are significant in some formulations. For example, in column 2, we see that the relationship between draft pick and body mass index is U-shaped. As body mass index rises from its minimum value in our sample, college quarterbacks tend to be drafted earlier. But (from column 2), beyond the within-sample of 29.1 a further increase in body mass index is associated, for given height and other variables, with being picked later in the draft.

It is important to note that we also included a dummy variable for race (equal to one if the player was black) but this was not significant. Despite this result, there are variations in the combine data by race. Specifically, on the Wonderlic test, the 30 black quarterback's averaged only a 20.2 score while the white quarterback's scored a 27.7. White quarterbacks were also taller (75.3 inches vs. 74.3) but slower in the 40 yard dash (4.85 s vs. 4.68). When we turn to performance, blacks on average offered more. Blacks on average have a higher QB rating, QB Score per play, Net Points per play, and Wins Produced per 100 plays.¹⁸ Despite these differences, though, the average black quarterback was only selected six slots ahead of the average white signal caller (118.0 vs. 124.5).

When we look at the entire sample of black and white quarterbacks we see that nearly 20% of the variation in a quarterback's draft position is explained by just the combine factors. When we add a performance measure, our explanatory power rises less than 3%. In other words, the combine factors appear to be more important than the actual college performance of the quarterbacks.¹⁹ These results suggest that NFL scouts are more influenced by what they see when they meet the players at the combine than what they players actually did playing the game of football.²⁰ There are two possible explanations for this

¹⁸ With respect to QB Rating, blacks posted an average mark of 100.6 versus 92.6 for whites. For QB Score per play, Net points per play, and Wins Produced per 100 play the differences were 3.676 versus 3.021, 0.311 versus 0.257, and 0.818 versus 0.673, respectively.

¹⁹ One potential issue is that there is very little variation in college performance. After all, only the quarterbacks who are considered the best in college get a chance to play in the NFL. To address this issue we looked at quarterbacks from 1998 to 2007 that were both drafted and logged at least 100 plays in a single NFL regular season. In all we had 215 NFL season observations. The standard deviation in QB Score per play, Net Points per play, Wins Produced per 100 plays, and QB Rating in the NFL was 1.054, 0.086, 0.231, and 14.45, respectively. For these quarterbacks the standard deviation for these same stats in college was 1.001, 0.081, 0.219, and 15.47. In sum, with respect to QB Score, Net Points, and Wins Produced we see slightly less variation in the college numbers. For QB Rating, though, the college numbers have a greater level of variation.

²⁰ We also regressed PICK on just the college performance numbers. When we regress PICK on aggregate college performance numbers

result. First, the combine measures isolate the quarterback and are not influenced by other players on the team and the opposition. Although the lack of real game competition appears to be a disadvantage, it may be an advantage for scouts in helping them to assess quarterbacks as individuals. Second, the combine measures may actually be correlated with intangible quarterback attributes not revealed by college performance indicators. Nevertheless, it does not follow that the scouts' assessments of quarterbacks will translate effectively into successful NFL playing career performance and we proceed to consider this in Sect. 6.

6 Connecting draft data to future performance

We now turn to the question of how the factors known on draft day relate to future NFL performance. This is a somewhat difficult issue to address. We have data on what quarterbacks did in college. And we have information from the NFL combine for each quarterback. But what performance should we use as the dependent variable?

The first issue is an adequate sample of performance data. Specifically, we need enough data on NFL performance to develop a reasonable assessment of a quarterback's performance. Following the practice noted earlier, we will only look at quarterbacks who logged at least 100 plays. Of the quarterbacks drafted since 1998, 72 signal callers participated in at least 100 plays in a single season once. If we look at this by years of experience, we have 43 quarterbacks with 100 plays in their first year in the NFL. In second year we see 50 observations, while in year three and year four there were 32 and 30 observations, respectively. After year four, though, we have 21 or fewer observations. The scarcity of observations indicates that we can only focus on what a quarterback did his first four seasons.

With data sets in hand, we estimated the following model:

$$\begin{aligned} \text{NFL PERFORMANCE} = & b_0 + b_1 \times \text{Height} + b_2 \times \text{BMI} \\ & + b_3 \times \text{BMI}^2 + b_4 \times \text{Wonderlic} + b_5 \times 40 \text{ Yard Dash Time} \\ & + b_6 \times \text{Dummy for non-Division I-A quarterbacks} + b_7 \\ & \times \text{COLLEGE PERFORMANCE} + e_t \end{aligned} \quad (4)$$

where Performance is measured as

Wins Produced per 100 plays
QB Rating

Footnote 20 continued

(QB Score, Net Points, or Wins Produced) we are able to explain 7% of the variation in draft position. When we consider the per play measures and career plays, our explanatory power rises to 9%.

Completion Percentage

Interceptions per Attempt

Passing Yards per Attempt (Note: If NFL Performance was measured as Wins Produced per 100 plays, then college performance was also Wins Produced per 100 plays. The same story can be told for each of our performance metrics.)

With four levels of experience, and multiple definitions of performance, we were able to estimate Eq. (4) many times.²¹ In all of our formulations, we never found that the combine factors, or the college performance with respect to Wins Produced per 100 plays or QB rating, had a significant impact—of the expected sign—on NFL Wins Produced per play or NFL QB Rating at any level of experience in the NFL.²²

When we look at the single metrics (completion percentage, yards per attempt, touchdowns per attempt, interceptions per attempt), we do find that college completion percentage has a statistically significant and positive impact on completion percentage at each level of experience.²³ As for the other single metrics, we again did not find any statistically significant explanatory variables for yards per attempt or TD per attempt.²⁴ Interceptions per attempt were positively impacted by being a non-Division I-A player in the first year of a player's career. Also in the first year, a higher BMI was found to reduce interceptions per attempt. But no other factor had a statically significant impact on interceptions per attempt.²⁵

To summarize our results, it appears that completion percentage in college tells us something about completion percentage in the NFL. On the surface, this does suggest that passers who were accurate in college remain accurate

²¹ There did not appear a simple way to present all of these regressions in a table. The results, though, are available from the authors upon request.

²² The Wonderlic score was statistically significant for 1 year of experience, but when this happened the sign was negative. In other words, higher Wonderlic scores were associated with lower levels of performance.

²³ When we estimate Eq. (4) with completion percentage as the performance metric, we find that college completion percentage is statistically significant. But the model's adjusted *r*-squared is only 0.18. In other words, much of the variation in an NFL's quarterback's completion percentage is not related to what that player did in college. Furthermore, the estimated elasticity of NFL completion percentage relative to college completion percentage is only 0.34. So a 10% increase in a player's completion percentage in college only leads to a 3.4% improvement in what the player will do in the NFL. Such results suggest that although college completion percentage has statistical significance, the economic significance of this factor is quite small.

²⁴ We did find that Wonderlic had a negative impact on touchdowns per attempt during the first year of experience.

²⁵ We also considered NFL career performance after 3 years (to hold experience constant) as a dependent variable. We further experimented with college career plays as an explanatory variable. The results of estimation were little different.

in their professional football careers. One should note, though, that college completion percentage explains less than 20% of the completion percentage we observe in the NFL. Consequently, it does not follow from our analysis that quarterbacks fail to develop in terms of accuracy as they move up into the professional ranks. We believe this is because college and professional league competitions vary greatly in terms of overall levels and heterogeneity of talent available, the degree and quality of competition, and speed and intensity of play. In any case, our examination of the factors that determine draft position indicated that completion percentage in college was not considered important on draft day. The factors that do appear important—height, Wonderlic score, and 40 yard dash times—do not appear to have an impact on future NFL performance.

Of course, NFL scouts will be experienced and trained to recognize intangible features of quarterback play, such as leadership qualities and ability to perform against defensive pressure that may not be captured by either college performance statistics or combine data. We can check whether scouting data help explain eventual NFL performance by the following method. First, we regress draft pick on college performance (here selected as net points) and other variables shown in Eq. (3). We save the residuals; these represent everything that went into selecting the player that we did not have in the original model. We suggest that these residuals capture scouting data, although other unobservables will be present. Next we regress NFL performance on the same variables used in the draft pick model plus NFL experience. This will tell us whether the omitted scouting variables are really important or not. We find that the coefficient on residuals in the NFL net points equation is -0.182 with a *t*-statistic, computed using robust standard errors, of 2.49. The *R*-squared in the net points model rises from 0.23 without residuals to 0.27 with residuals. Overall, this does suggest that less tangible scouting data do matter for explaining the variation in NFL quarterback performance, although the explanatory power of scouting—as it has been measured—appears rather small.

As a final experiment, we estimated a two stage least squares model in which draft pick was determined by Eq. (3) above in the first stage. In the second stage we estimated:

$$\begin{aligned} \text{NFL PERFORMANCE} = & c_0 + c_1 \times \text{NFL experience} \\ & + c_2 \times \text{NFL experience}^2 + c_3 \times \text{Pick} + e_t \end{aligned} \quad (5)$$

The sample used for estimation was a set of 59 quarterbacks who had at least 100 plays in the NFL in a given season, with pooled observations over the period 1998–2007. Since we have multiple observations for most quarterbacks, the standard errors were clustered by quarterback. This corrects for interdependence of errors within quarterbacks, while preserving independence of errors across quarterbacks.

Table 6 ($N = 172$), alternative dependent variables for performance

Variable	QB rating	Total yards	QB score	Net points	Wins
Experience	6.433 (4.18)	626.9 (3.69)	287.3 (3.41)	24.35 (3.42)	0.638 (3.40)
Experience squared	-0.449 (2.34)	-58.2 (2.59)	-22.8 (1.92)	-1.95 (1.94)	-0.051 (1.91)
Pick	-0.019 (0.52)	-2.347 (0.72)	-0.915 (0.55)	-0.079 (0.57)	-0.002 (0.56)

Standard errors are clustered by quarterback

The results of estimation of Eq. (5) are reported in Table 6. Regardless of the selected NFL performance metric, experience determines performance independently of draft pick, with no significant role for draft choice. Put another way, comparing two quarterbacks with same NFL experience, the player selected earlier in the draft is not predicted to have significantly different NFL performance levels than a player picked later in the draft. Draft pick is not a significant predictor of NFL performance.

7 Concluding observations

We began with nearly four decades of data on NFL performance and where a quarterback was taken in the NFL draft. Our analysis revealed that there was a relationship between aggregate performance and where a player was chosen. But when we looked at per play performance, the relationship between production and draft position was quite weak. In contrast, a much stronger relationship existed between how many plays a quarterback ran and where he was selected. In sum, draft position can get a quarterback on the field. But quarterbacks taken higher do not appear to perform any better.

This finding led us to investigate the factors that determine where a player was chosen in the draft. Our study indicated that it was the combine factors that dominated the decision. College performance did impact where a quarterback was taken, but performance on the field was dominated by factors like height, Wonderlic score, and 40 yard dash times.

Our study of subsequent NFL performance—which was hampered by a lack of data—failed to find that the combine factors had much of an impact on future performance. In essence, NFL decision-makers can be impressed by taller, smarter, and faster signal callers. But there is no evidence that the extra inches, better test scores, or faster 40 yard dash times make any difference in subsequent NFL performance. Indeed, as Table 6 shows, overall, draft pick is not a significant predictor of NFL performance.

Of course one can argue that our model designed to predict draft position and NFL performance is incomplete and our discussion of intangible scouting data does suggest that there are omitted scouting variables in our model. NFL decision-makers are also able to interview the candidates

and look at factors such as arm strength and accuracy in passing drills. Also, we have focused on the position in NFL which is hardest to assess and where eventual NFL success is hardest to predict. Quarterbacks are multi-skilled and have complex mental and physical tasks to perform on the field. In contrast, running backs and offensive line players have roles that are more narrowly defined. It may be that, for example, it is easier to predict the eventual NFL career performance of college offensive lineman somewhat easier than for quarterbacks.

However, two key points emerge from our study. First, college football is very dissimilar as a competition to the NFL. Second, the combine puts an aspirant NFL player into an artificial training environment where league competition is absent. Both factors mean that it is extremely difficult for NFL franchises to assess player talent at point of draft and generate efficient matching of that talent to NFL rosters. We should also stress that when we look at nearly 40 years of data, we fail to find a relationship between where a quarterback is selected and his subsequent NFL performance on a per-play level. So although NFL decision-makers may consider more factors than we offer in our models, there is little evidence that these additional factors help NFL decision-makers make better forecasts of future performance.

References

- Berri DJ (2007) Back to back evaluation on the gridiron. In: Albert JH, Koning RH (eds) *Statistical thinking in sport*. Chapman & Hall/CRC, Boca Raton
- Berri DJ, Simmons R (2009) Race and the evaluation of signal callers in the National Football League. *J Sports Econ* 10(1):23–43. doi: [10.1177/1527002508327383](https://doi.org/10.1177/1527002508327383)
- Berri DJ, Schmidt MB, Brook SL (2006) *The wages of wins: taking measure of the many myths in modern sport*. Stanford University Press, Stanford CA
- Blass AA (1992) Does the baseball labor market contradict the human capital model of investment? *Rev Econ Stats* 74(2):261–268
- Chappell M (2006) Wondering about the Wonderlic? Try it. USA Today, 28 Feb. http://www.usatoday.com/sports/football/draft/2006-02-28-wonderlic_x.htm
- Fort RD (2006) *Sports economics*, 2nd edn. Prentice Hall, Upper Saddle River
- Hendricks W, Debrock L, Koenker R (2003) Uncertainty, hiring, and subsequent performance: the NFL draft. *J Labor Econ* 21(4):857–886. doi:[10.1086/377025](https://doi.org/10.1086/377025)

- Kahn LM (2000) The sports business as a labor market laboratory. *J Econ Perspect* 14(3):75–94
- Krautmann AC (1999) What's wrong with Scully-estimates of a player's marginal revenue product? *Econ Inq* 37(2):369–381
- Krautmann AC, von Allmen P, Berri DJ (2009) The underpayment of restricted players in North American sports leagues. *Int J Sport Fin* 4(3):161–175
- Lazear EP (1986) Incentive contracts. National Bureau of Economic Research Working Paper No. W1917
- Leeds MA, von Allmen P (2008) *The economics of sports*, 3rd edn. Addison Wesley, New York
- Lehn K (1982) Property rights, risk sharing and player disability in Major League Baseball. *J Law Econ* 25(2):343–366. doi:[10.1086/467020](https://doi.org/10.1086/467020)
- Massey C, Thaler R (2005) Overconfidence vs. market efficiency in the National Football League. NBER Working Paper No. 11270
- Quinn KG (2008) Player drafts in the major North American sports leagues. In: Humphreys BR, Howard D (eds) *The business of sport*, vol 3. Praeger, Westport
- Quirk J, Fort R (1992) *Pay dirt: the business of professional team sports*. Princeton University Press, Princeton
- Scully GW (1974) Pay and performance in Major League Baseball. *Am Econ Rev* 64(6):915–930
- Simmons R, Berri DJ (2009) Gains from specialization and free agency: the story from the gridiron. *Rev Ind Organ* 34(1):81–98. doi:[10.1007/s11151-009-9200-9](https://doi.org/10.1007/s11151-009-9200-9)