How Valuable Are College Football Players?

Rebeka Czobor

Supervisor: Prof. Alessandro De Chiara

Central European University, Department of Economics

In partial fulfilment of the requirements for the degree of Masters of Arts

Budapest, Hungary

2018

Abstract

There has been an ongoing debate on the employment status of college football players in the past years. Evidence shows a great amount of students living below the poverty line, while generating huge profit for their schools. In this thesis I estimate how valuable these players really are by quantifying quarterbacks' marginal revenue product in two steps: using weekly game level data from 2005-2017 I estimated the contribution of players to team win numbers by running logit models. In the second step I calculated an additional win's impact on team revenues using fixed effect models. I concluded that elite quarterback's belonging to the 95th percentile of players generated an additional $1.1 million for their school team compared to the average athletes.

Table of contents

List of Tables and Figures

Introduction

There has been a long ongoing debate regarding the employment status of college football players. Although they are currently not allowed to accept payment from their schools, they are extremely good at generating profit for them with a rigorous training routine, which can take up 60 hours of their time a week. Some argue, that regarding the above mentioned conditions these students should in fact be considered as employees by the National Collegiate Athletic Association (NCAA), in order to get higher benefits and some legal protection, to which they are currently not subject to. Currently an injured player can lose his/her full scholarship, without any compensation whatsoever. On the other hand, there would also be non-negligible questions arising with the employment of college athletes, such as: how much should they be paid? Would schools offer different wages, which would trigger the bidding war (and on the other hand, is this bidding war a big enough issue to let a cartel-like system to operate)? Would less profitable (or non-profitable) college sports be funded?

When talking about the questions mentioned above, it might be better to examine the facts on how much money the schools actually earn on a player, and what are the benefits that a student gets. It is true, that a per se payment is not given, but college athletes receive no little amount of benefits. With a full tuition waiver, accommodation, meal plans, tutoring and money covering the cost of books and miscellaneous expenses, it is estimated that an NCAA football athlete is actually subject to an average funding ranging from $20,000-40,000[1], depending on the college. These are fairly straightforward figures, which can be easily calculated (although of course, illegal gifts and benefits are not included) but measuring the schools' earnings on

[1] Average scholarships from: http://www.scholarshipstats.com/football.html while average annual cost of attendace is estimated to be $20,770 by Statista: https://www.statista.com/statistics/235651/us-university-attendance-cost/

athletes is a much harder task especially in the case of team sports, where individual contribution is hardly separable from others'.

As the debate escalated upon the issue of college football and it is the biggest revenue-generating division in college sports with the most people involved in it, I decided to figure out, what really is the gap between individual contributions to school revenues and individual gains for the subset of football player athletes. To do so, I estimated players' marginal revenue product (MRP). I only focused on gains and contributions in the course of the athletes' studies – I did not take into account the possibility of future NFL contracts, which have an enormous effect on lifetime earnings even when discounted, nor have I taken into account merchandise sold by the schools after graduation, etc. I only focused on a special subset of football players: quarterbacks. Quarterbacks' contribution to games is the most easily quantified and also, they are the ones making the most money later in the NFL, which also indicates that they have the biggest added value during the game (this can be also seen from the fact that the Heismenn Trophy and the Most Valuable Player title is usually won by quarterbacks). Since quarterbacks' have the biggest revenue-generating power my estimations for player MRP will in fact be an upper bound measure.

In order to estimate the marginal revenue product of a player I needed to estimate their quality, and through that their contribution to winning probability. In the next step I measured wins' contribution to school revenue. In this multi-step estimation, I could grab what is an individual players' contribution to school revenues. For my estimations I used various data sources, such as ESPN metrics on player quality, data from Equity in Athletics Disclosure Act (EADA) on college revenues. According to my estimations an elite player has a seasonal contribution to winning an additional 3.8 games compared to the mean of players, and this means that they generate an additional $1.1 million in revenues to their college team compared to an average player.

The structure of the thesis is as it follows: in Chapter I. I give a short introduction to the theoretical background of this thesis' framework as well as to the current situation of college football players and review the previous literature on the topic. In Chapter II. I present my data sources, data cleaning methods, and I further continue to present the econometric framework used. In Chapter III. I present a methodology involving a multi-step analysis for college quarterbacks' MRP estimation. In Chapter 4. I conclude and explore what further analysis and research could be done.

1. Theoretical background and literature

In every research paper the goal is to answer a relevant question, or at least, take a step forward in getting a relevant answer to an interesting topic. The motive of this thesis is the unsolved problem of college athletes, who have been struggling for years for employment status in the USA. Although this issue has many aspects - mostly legal, technical or ideological - I decided to demonstrate whether the points of students can be valid from an economists' standpoint of view. It also reflects the question from a different angle, and it gives an interesting insight to the economics of college sports.

College football is a great deal in the USA and should be treated differently than other NCAA athletic programs – in fact an average Division I (FBS) school generated $31.9 million in football revenue in 2016, which is slightly more than what the next 35 sport generated combined[2]. According to the USA Today's report, in the school year of 2015-2016 there were 28 institutions generating a revenue in NCAA sports above $100 million. The Texas Longhorn football team alone made more than $121 million in 2014[3]. Above all that, some institutions are also heavily subsidised by the state. Besides being a billion dollar market, it is also true that college football involves a great number of students and in this regard as well it is the biggest class in NCAA. In 2017 a total of 73,063 students were part of a football team, while only 34,980 in baseball, 18,712 in men's and 16,532 in women's basketball[3]. Despite the enormous revenues it is also true that expenditures are not of a negligible size either – tuition, marketing, team travel, coaches and organizing games costs a lot. Still, the total market profit is around $3 billion. But what do these expenses mean, when it comes to tuition? What is the indirect payment students get for devoting usually 60 hours a week for training and taking part in games?

[2] From Business Insider, who accessed data from the Departement of Education

[3] Based on NCAA report: http://www.ncaa.org/sites/default/files/2016-17NCAA-0472_ParticRatesReport-FINAL_20171120.pdf

In the debate regarding college athletes the ultimate reason against employee status is that students do already get paid – even if indirectly. This mean a full or partial tuition waiver, accommodation, meal plans, books which is not a little amount each year. Despite the undeniable value of a college diploma – or even so, an Ivy League diploma – it is also worth noting that not all players are able to graduate or get jobs in their field upon graduating since studies are obviously secondary next to a week fully loaded with heavy training. A study of 2011 showed that "the compensation FBS athletes who are on "full scholarship" receive for living expenses situates the vast majority at or below the poverty level" and that for the academic year of 2009-2010 the out of pocket expenses of an average FBS player were $3,222 (Huma et al., 2011). It might also be a valid argument that future revenues in the National Football League (NFL) could offset these opportunity costs of those getting to the professional league. However, in 2017 the probability of transferring from NCAA to a Major Professional team was around 1.6% meaning realistically around 1200 players nationwide at a given academic year[4].

Theoretically there exists an "appropriate" wage for players if we look at them as employees just as in any other market. In neoclassical economics in a profit-maximizing firm employees will be hired until the marginal product of labour (MRP_L) equals the wage rate. This MRP is the increase in revenues which is due to the increase in output produced by the last hired employee. Technically, it is the product of the marginal product (increase in output by employing the last worker) and the marginal revenue (increase in revenues from the increment to output). Calculating these is though not an easy task – what is the marginal product of a player? Without any player, the team cannot participate in a game, and thus it is not able to generate revenue. There is also not a single, concrete output as such, which contributes to

[4] 4 Based on NCAA report

revenues – wins of course matter, but there are also marketing rights, merchandise and other goods which help teams stay profitable.

Although there has been a huge overall development in statistics, econometrics and data collecting in the past two decades, sport statistics has seen an even bigger improvement. As real-time betting, fantasy football and data based player performance measurement gained more and more relevance and the money invested in these field kept increasing, players' performance – especially in football, baseball and basketball – are better described with numbers than ever before. Data on every movement, every cent spent by teams or every cent (legally) paid to athletes is tracked and readily available online. It is straightforward to try to proxy player MRP by measuring their performance and through that their "quality's" contribution to winning probabilities, which have an effect on school revenue.

One of the first works on athletes' marginal revenue product was done by Scully in 1974: he investigated the case of Major League Baseball exploitation (Scully, 1974). In his paper he uses a multi-step regression model to measure player performance's effect on team revenue – a similar approach to what I use in Chapter 3. He argues that in general fans attend a game to see the end result (namely, whether their team wins or loses) not the hitting or pitching performance in itself. This way it is enough to see how player performance affects team winning probability and through winning, indirectly team revenues as well. While many may argue that this is not the case, I generally agreed with this approach and translated it to football – although with more data and metrics available, bigger sample size, and in a generally different setting. According to his calculations, empirical analysis pointed out a player exploitation of significant magnitude. He also argued that existing regulations introduced monopsony power and player economic rents.

More work has been done regarding college basketball players, as performance is a bit more easily measured. However, when it comes to the question of MRP, the difficulty lies in the lack of data for a big chunk of students, who are not starters, rarely or never play. Intuitively it cannot be argued, that secondary players do not contribute to team winnings or team performance – they provide replacement in case of a starter's injury, they provide scrimmage players for the starters and they too, devote just as much of their time for practice and games as the ones we have all the data on. Lane, Nigel and Netz proposed alternative measurements of college basketball athletes' MRPs, and also trying to capture the "exploitation magnitude" for the second-line players. By using school fixed effects, using the distribution of the National Basketball Association's (NBA) salaries to estimate MRPs for players without playing statistics and also using information on players later drafted into NBA, they concluded that around 60% of players' MRPs are greater than the scholarships caps for all basketball players, including those with little or no available game statistics.

In the case of college football, the earliest works have been done by Brown, using ex-post observations on draft status – in this case meaning that he used data from those players who later were drafted to the professional league, the NFL as a proxy – and a TSLS model. He later updated his model more times, using bigger sample size and more sophisticated metrics. In his latest paper on premium player value in 2010, he has found that "adding" a premium player (by premium meaning those ones who later are drafted to NFL) to a team means a ceteris paribus $1.1 million increase in team revenue in a given year. In his latest work in 2012, he estimated that although these player added values to team revenues are significantly higher than player compensation during college years, "between 33 and 38 percent of this sample players (active and inactive) will earn NFL incomes sufficient to offset their monopsony-lost college earnings" (Brown, 2012). Although these findings might be accurate, I also claim that this means around 0.61% of total yearly players based on the statistics mentioned above. This represents around 450 players in 2016, when 73,000 college football athletes played in one of the schools' teams. Because of this, I feel that MRP is more meaningful when it is measured not only for those later drafted in the NFL, making truly monopsony-rent offsetting wages, but for all the players aiming to go professional later on.

In football it is very hard to quantify player performance, so when estimating player MRP for all players (not only those later drafted) one is faced with a difficult task – weighting certain statistics as yards passed or tackles is a bit ad-hoc, and there are fundamental differences between offensive and defensive players. I decided to focus on quarterback players as quarterbacks' have a special position even among athletes and football players. In the NFL quarterbacks receive historically on average the highest payment and are considered as the key players of the game. Moreover, their performance is more easily quantifiable than any other position's – ESPN among many other metrics, has developed Total Quarterback Rating (QBR) which measures player "quality" on a 0-100 scale. Since these athletes' performance is measurable, I run my regressions on their subset, and since they are the ones expectedly with the highest revenue generating ability, I also state that this will be an upper estimation of college football player MRP. I further discuss QBR and other metrics used in Chapter 2.

QBR has been used before to proxy for player MRP. Hunsberger and Gitter concluded in a 2014 study that "a one standard deviation increase in QBR adds about 3 wins per season, and each additional win increases a school's football revenue roughly $740,000 compared to the average quarterback" (Hunsberger and Gitter, 2014). They also used a multi-step approach similar to Scully, taking into consideration player performance's effect to win probability and win's effect on marginal revenue. However, the QBR metric has been updated since to control for important factors, such as opponent ranking and teammates. It is also important how individual MRP changed in the past years as market revenues continued to rise in college football. In the light of all these, I chose a similar methodology as Hunsberger and Gitter in their paper, with using updated metrics and advanced data collection methods.

2. Data and econometric framework

In the previous pages I already mentioned some of my data sources and gave some idea on the econometric framework which is used in this thesis, but in the following sections I explore these areas in a more detailed and coherent way. All of the data used in the following sections is publicly available from the internet, although in some cases I did use a program called Python to scrape pages in order to access the data faster. For the analysis part, I used both Python and Stata, whichever seemed more convenient[5].

2.1 Data

Sport teams especially in the USA use a tremendous amount of data and advanced econometrics and this matter they have help. Some statisticians love to devote time to sports analytics, since it is interesting and not less importantly, they can make money on the betting market if they are sophisticated enough. In other cases sport channels and websites try to demonstrate their expertise by predicting certain outcomes.

Presumably in the same spirit ESPN, the biggest sports channel in the USA developed a special metric for quarterback ranking in 2011, called quarterback rating (QBR). While the computation of QBR might be hard to grasp for an average football fan, it can be explained in simple words. Based on play-by-play data they estimated how much a quarterback contributed to a win, but does so by using expected points as the basis of the evaluation (e.g. it does not treat a win of 24-0 the same as 24-23). The metric takes into account every relevant aspect in which the player contributes to the points, rushes, passes, yards, while also accounts for "exogenous" circumstances such as opponent team defence strength. Since it is based on play level it also accounts for the team compositions. I retrieved college level QBR by scraping ESPN's website containing the weekly leaders for each week in the season. I chose this option instead of just getting data on the season leaders, because on one hand this increased my sample

[5] All Python codes, stata files or datasets can be accessed by requesting them from the author at czobor.rebeka@gmail.com

size, and on the other, this contained information on the games played each week, and the points for each game, rather than just an aggregated seasonal outcome. Although NCAA covers more than 2000 institutions, I restricted my sample size to the 230 schools taking part in some of the big conferences (as football may not be a relevant sport division in other schools, and ESPN does not have extensive metrics for players in small teams). Moreover, out of the 230 schools covered, ESPN had only measurements for players out of 130 colleges – this is due to the fact that usually 70-120 weekly leaders are listed. In order to be able to qualify as a weekly leader, a player must have at least 20 action plays a week (which is around 5 minutes of game time on average). It is quite intuitive that the dataset is unbalanced: some schools have more games per season and qualify more weekly leaders than others. Eventually I had financial and player level data for 121 colleges.

The dataset has 14,879 observations on 1,318 players. This means approximately 8,000 matches – some are observed from losing and winning side as well (if both the losing and winning quarterbacks were qualified among the weekly leaders) and some are not. I did not count these as duplicated values, as the QBR of the winner and the loser are not linearly dependent as the measures take into consideration many more aspects than just the opponent ranking. I also excluded those games where the opponent was out of the 124 schools (in the other 100 for which I do not have observations for QBR). Not very surprisingly I have found that QBR and winning a game is highly correlated. In my subsample of 14,879 observations de mean QBR was 54.73 which is slightly higher than the calibrated 50 by ESPN. In my view this is due to the fact that I observed the weekly leader board, which misses out students who play too little amount (less than 5 minutes per season) who also are more likely to have lower ratings. Also it is important to mention, that there were no observations for the same game from the same team's view (e.g. there were no cases where two quarterbacks landed on the leader board from the same match). Otherwise, it would be hard to disentangle each players' contribution to the win.

Win	Mean	Std. Dev.	Freq.
0	43.27	23.30	7,511
1	66.41	22.99	7,368
Total	54.73	25.88	14,879

Table 1. Summary statistics of QBR

The summary statistics - with respect to the outcome of the game - for QBR can be found in Table 1. For winning games the average QBR of quarterbacks was 66.41, while for losing teams this was 43.27. The standard deviation of the variable – although it is high in both cases - was also lower in case the Win dummy was one (meaning the team won the match). A bit more surprising feature of Table 1. Is that there are slightly more observations for lost games in my sample (7,511 out of 14,879). Although this is not significantly higher than 50% of the sample, it indicates that the weekly leaders' board has more observations for quarterbacks whose team lost. It is worth noting at this point, that although the dataset is called weekly leaders, weak QBRs are also reported if there were games played that week by a certain team. QBR thus in my sample ranges from 0 to 99.9. Figure 1 represents the distribution of QBR over the Win dummy.

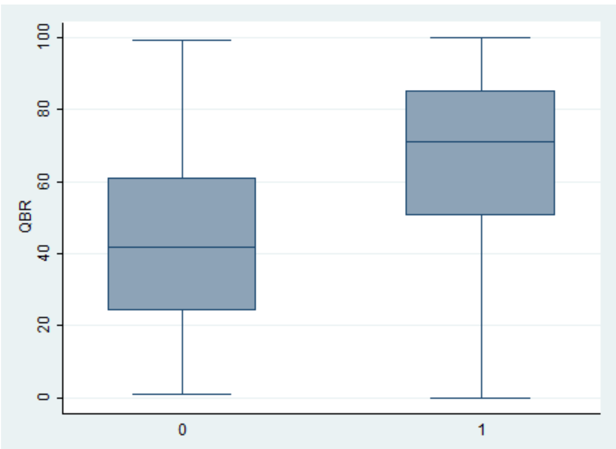

Figure 1. The distribution of QBR with respect to winning

I also needed to collect data on school level football expenditures and revenues. Luckily enough, these are all available on the EADA website (see data sources on the last page) and not just as an aggregate, but divided into subcategories such as ticket, merchandise, subsidy revenues or tuition, travel and other expenses. An average school from my sample had a team revenue of more than $2 million, but the range goes from $595,000 to $141 million. There has been no significant correlation between team or school size and revenues, however, the correlation coefficient was 0.89 between expenses and revenues. I assume this last finding is not particularly surprising: better teams have more money to spend on coaches, equipment, travel and food for players, which means more wins, more coverage, sold merchandise and tickets. Some basic measures on college revenues and attendees are represented in Table 2, while correlations are in Table 3.

Variable	**Obs.**	**Mean**	**Std. Dev.**	**Min**	**Max**
Revenues	1318	2.30e+07	2.16e+07	595,126	1.41e+08
Participants	1319	115.93	11.64	82	167
Expenses	1,319	1.37e+07	8,468,228	1,449,862	6.23e+07
Total stud	1,318	17,630.75	8085.91	1,247	50,394

Table 2. Summary statistics for Colleges

	Expenses	**Revenues**	**Participants**	**Total stud.**
Expenses	1.00			
Revenues	0.88	1.00		
Participants	0.23	0.27	1.00	
Total stud.	0.42	0.52	0.18	1.00

Table 3. Correlation coefficients among variables

Other statistical data was collected from various sources. Scholarship amounts were retrieved from Scholarshipstats, aggregated yearly player numbers or participation probabilities from the NCAA's own report.

2.2 Econometric framework

The framework I am using in the following chapter is similar to the one presented by Brown and by Hunsberger and Gitter, meaning that I chose the same three-step approach to estimate player MRP; in the first step I estimate player contribution to win probability (the marginal product), in the second I estimate the value of a marginal win (the marginal revenue), and based on the first two steps, I calculate player MRP by the well-known equation from microeconomics:

$$MRP = MR * MP$$

I will compare these results to the benefits players get: a close approximation of the value of the grant-in-aid, for which there is school-level data available. Player-level scholarships are of course not discussed, and although there is heterogeneity in the value of scholarships between colleges (as tuition fees also vary) there is not much deviation between schools compared to revenues or professional athletes' salaries. The average scholarship varies between $1,500 and $50,000 dollars, which is usually covering the tuition fees (ranging between $5,000 and $50,000). Full ride scholarships also cover food and accommodation for students, so based on living costs these benefits under full scholarships can make up $50,000-$100,000 a year.

In step one I run a logit regression on winning probability, where QBR is the main explanatory variable. I use logit as I have no underlying normality distribution assumption, but it is far easier to interpret. Nevertheless, I did run both regression types, but I have found no significant difference in marginal effects. As I have mentioned before, quarterback ratings control for the opponents' defence strength. In American football a team is composed of offensive and defensive players but always only one type is on the field – e.g. we can look at this as two separate teams participating at every game. Whenever the defensive line is on the field, the offensives sit on the bench. Quarterbacks are offensive players, so QBR has little to do with other team's offensive line. However, since opponents' defence strength affects QBR and also the win probability, leaving out opponent offense rating would lead to endogeneity.

Luckily, these ratings are also easily available through web scraping. In the win probability regression thus I included opponent ranking (Defencerank variable) and whether the game was played at the team's home (Home dummy) stadium, since historically it is an important factor. Defence ranking is computed by using a substantial amount of additional measures, such as successful tackles, defence touchdowns and so on. Since the defence team's points calculated this way say a bit too little to an average reader (200 points vs 100 mean not so much if standard deviations vary each year and the measure is not normalised) so I only ranked the teams based on points. Including ranking is much more straightforward in my opinion, which eases interpretation. Just to clarify this, the probability of winning the j^{th} game in season t is:

(1) $$P(Win_{jt}=1|\mathbf{X\alpha}) = \alpha_0+\alpha_1 Home_{jt}+\alpha_2 Defenserank_{jt}+\alpha_3 QBR_{jt}+u_{jt}$$

In step two I estimate an additional win's contribution to school revenues. This is harder than it seems at first sight: it would be convenient to assume that revenues are linearly dependent on win numbers, however, it is quite intuitively not the case. An additional win in pre-season does not hold the same value as a championship final win (as usually monetary prizes are tied to these titles). The linearity assumption is something which in general can be made as a simplifying assumption, since we know that wins in any state positively correlate with revenues, and on average we might capture a similar win effect. One should bear in mind nevertheless, that the coefficient of winning in the following regressions might be upward or downward biased.

For the second step the dependent variable is obviously the school revenue, whereas on the right hand side of the equation I use the number of wins per season for a given team (Winnumber), the number of games played at the home stadium per season (Homenumber) as the biggest revenue generating part is the ticket sales. Table 4 gives a quick summary on these two variables.

Variable	Obs	Mean	Std. Dev.	Min	Max
Winnumber	1264	4.929	2.778	0	17
Homenumber	1264	4.917	1.486	0	12

Table 4. Summary statistics of regression variables

I also included year dummies to account for changes in price levels and other factors. In order to have a more precise estimation I included team fixed effects. This way I estimated that in time *t* for team *i* the total revenue is given by the following equation:

$$\text{(2)} \qquad \text{Revenue}_{jt}=\beta_0+\beta_1\text{Winnumber}+\beta_2\text{Homenumber}+\beta_3\text{Year}_t+\omega_{it}$$

In step three I utilise the coefficients found in step one and step two to get marginal revenue product of players as a product of marginal revenue and marginal product. First I examine how much QBR contributes to wins, by taking the average marginal effects from the logit model. Then, from seeing how much additional wins this means in a season I will also estimate QBR's effect on team revenues.

3. Discussion and findings

In the following section I present the output of the regressions mentioned in the previous chapter. As a first step, Table 5 contains the coefficients and standard errors for the logit regression capturing the effect of quarterback rating on winning probability. The coefficients reported in the table are as raw Stata outputs, which are hardly interpreted – however, fortunately logit estimations are linear in the odds ratio. Taking the coefficients in the form of e^{α} gives us a linear effect of the variables on $\frac{P_{win}}{1-P_{win}}$. To have more precise estimation, I also included the average marginal effects in Table 5.

Win	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
Defencerank	.0083	.00	16.61	0.000	.007	.009
QBR	.042	.00	50.70	0.000	.040	.043
Home	.78	.038	20.51	0.000	.71	.85
_cons	-3.25	.067	-48.67	0.000	-3.38	-3.12
R^2	0.19					

	dy/dx	Std. Err.	z	P>|z|	[95% Conf. Interval]	
Defencerank	.002	.0001	17.15	0.000	.0014	.0018
QBR	.008	.0001	80.75	0.000	.0077	.0081
Home	.15	.007	21.55	0.000	.1346	.1615

Table 5. Logit regression outputs, with average marginal effects

As it can be seen, all the variables are statistically significant – although not so much necessarily in an economic way. Ceteris paribus raising QBR by one points means that the odds ratio will raise to its multiplication of 1.043. Although this does not seem much, let’s not forget that the range of QBR is from 0 to 100. So all in all a one point increase may not lead to significant differences in the odds ratio, a raise by 25.9 (or one standard deviation) very much matters: taking the average marginal effect, increasing QBR by one standard deviation above the mean means a 0.21 increase in win probability per game. At an average of 11 played games this means an average of 2.31 additional wins per season.

This is also true for defence ranking, which ranges between 1 and 121. Not so surprisingly, the Home dummy is statistically and economically significant: in case the game is played in the home stadium it raises the winning odds ratio more than twofold. The overall R^2 of the estimation is 0.19 which is not particularly high, but my aim was not mostly to estimate the winning probability the most accurately, but to capture QBR's effect on winning probability. I believe that there are no significant factors left out the estimation that correlate with QBR and win probability, and there is no significant endogeneity in the model. The captured coefficient of QBR by this logic, can be used to estimate MRP.

Table 6 summarizes the results from the revenue regression, with the explanatory variables number of wins, number of games played in the home stadium and year dummies. I also included team fixed effects for all teams. The number of wins is statistically and economically significant: ceteris paribus an increase in the yearly win numbers raises team revenues by $287,334. Year dummies are also significant, and their sign is as expected – since the baseline year is 2005, it could be predicted that revenues will increase later on. At a certain team with the same number of wins and hosted games revenues ceteris paribus increased by almost $2 million in 2016.

The only surprising coefficient might be the one of the number of hosted games. As I mentioned in the previous section, hosting games is generally beneficial for revenues as ticket sales increase. It is true that costs are also higher but I did not account for those in this paper. On the other hand, from the logistic regression it can be easily shown, that games played on the home field are won by much greater percentage and probability. Since there is a very high correlation between the variables Win and Home, I can imagine that it is hard to disentangle each effect, which causes the coefficient of Homenumber to be not so intuitively negative.

However, I could also argue that this variable is not even statistically significant, and even less so if I leave out the Winnumber variable. Also, by leaving out either variable would lead to endogeneity issues in my opinion. Leaving out Homenumber slightly decreases the coefficient of Winnumber and also its t-value, but does not change it drastically.

I included team fixed effects as I assumed that there are some team specific attributes that I cannot control for but affect revenues, such as alumni donations, state subsidies or long term sponsorships. However, accounting for 121 teams while having 1320 observations also means a lot of lost degrees of freedom. The unbalanced panel data can also cause some issues, both with and without fixed effects. The right column of the table contains simple OLS regression estimates without team effects – as it can be seen the coefficients and significances are very different. I concluded, that economically it makes more sense to include team fixed effects: bigger teams can generate a huge amount of revenues even after a bad season, simply because they have a whole infrastructure built around them. These team-specific trait for which I cannot control for are not negligible in my opinion, and the fixed effect estimation is better at grasping the additional win contribution to revenues. Since I estimated a linear OLS regression, the coefficients can be interpreted in an easy way. I already discussed the effects in the paragraphs above.

In the further step I estimate the marginal revenue product of players based on the findings of step one logit regression and the coefficient of Winnumber from the OLS regression using team fixed effects. Based on this we can calculate the MRP of college quarterbacks: as I mentioned before. In the first step I estimated one point of additional QBR to have a 0.008 average marginal effect on win probability, whereas an additional win has an impact of $287,334 on team revenues.

Revenues	**With team fixed effects.**	**Without team FE**
Winnumber	287,333.5*** (85,079.55)	1,929,333*** (213,469.3)
Homenumber	-262,751.2 (166,591.7)	4,335,693*** (404,554.6)
Year		
2006	1,980,883** (846,198.7)	-1,468,067 (2,642,534)
2007	4,118,986*** (854,495.4)	-1,664,679 (2,647,678)
2008	5,324,034*** (842,079.2)	-381,586.5 (2,606,099)
2009	6,437,668*** (846,162.8)	908,032.1 (2,611,847)
2010	7,680,663*** (845,528.1)	5,149,381* (2,626,145)
2011	9,622,826*** (851,896.7)	3,973,860 (2,629,499)
2012	1.12e+07*** (849,757.9)	4,587,971* (2,605,587)
2013	1.30e+07*** (857,481.9)	5,345,104* (2,634,227)
2014	1.53e+07*** (849,828.9)	8,353,756*** (2,612,880)
2015	1.74e+07*** 852,010.7	1.03e+07*** (2,614,960)
2016	1.94e+07*** (849,294.3)	1.30e+07*** (2,611,180)
_cons	1.43e+07*** (910,234.1)	-1.10e+07*** (2,403,615)
R^2	0.501	0.26
F-test	87.63***	34.63***
sigma_u	20,333,253	-
sigma_e	5,957,658.7	-
rho	.921	-

Table 6. Revenue regression outputs with and without team fixed effects[6]

[6]*** - significant on the 1% level
** - significant on the 5% level
* - significant on the 10% level

Let's consider players belonging to the 95th percentile of the distribution being "elite" players. This means an approximate QBR of 93.6, which is 1.65 standard deviations from the mean. Elite players on average generate an additional 3.81 win per season compared to the mean, meaning that each season they generate and additional revenue of $1.1 million for their team. In case we take the maximum scholarship and other benefits they get from their college being $100,000 this means that teams have an approximately $1 million profit on elite players, disregarding other financial benefits students might bring. In case we consider a QBR of 0 being the baseline, the effect is $2.4 million – this assumption however, makes a bit less sense than comparing to the mean, but reflects well the revenue-generating ability of high-ranked players.

So based on these equations, what QBR generates a profit of 0 for schools? Estimating a maximum expenditure on the quarterback being 100,000 a player should "generate" an additional 0.35 win per season, which corresponds to a QBR of 43.5 which is the 35th percentile of players. Of course, this is an imperfect measure of players' contribution to revenues as I mentioned before: even those quarterbacks who do not play or do not contribute directly much to win numbers have an impact on team performance. During practice they help, they provide substitutes for elite players (so in some cases, even participation would not be possible without them) they promote the school, and possibly even contribute to team revenues through merchandise or alumni contribution. The primary source of revenues for teams is through ticket sales and sponsorship though, which is hard to acquire when the performance is low.

Although this measure is just an approximation to player marginal revenue product, it can be clearly seen that even with the more modest estimations most of the players generate profit for schools, and some of them have a marginal revenue product of more than $1 million compared to the mean of players, and even $2.4 million if compared to a baseline QBR of 0. College football players are not considered as employees yet, however, according to multiple estimations this might not be such a huge step away.

4. Conclusions

In this thesis I showed how valuable college football players are – at least, a subset of them – by using weekly data on college quarterbacks from the last 13 years. I estimated that elite quarterbacks can on average mean 3.8 additional wins per season by using a logit model, and based on school revenue data I also concluded that an additional win means an approximately $300,000 plus in team revenues ceteris paribus. Based on this results I showed that elite quarterbacks can earn an additional of $1.1 million to their college compared to the mean of players.

There are certain caveats to this study of course, which need attention. On one hand I only computed the marginal revenue product of certain players, as delicate measurement on performance is missing for other posts. An extension of this investigation could highlight other results for other students, since quarterbacks are the most valuable players in the teams generally. Another week point can be that although players do contribute to school revenues through wins, and I truest the results of this thesis, but this is not the only channel of revenue-generation for players. One could also try to estimate what additional values are brought for schools by players, who e.g. only sit on the bench, or are so popular, that sponsors are fighting to support their team. This is a hard task, and it is out of the scope of this short essay. Lastly, another caveat which I mentioned above, is the wins' contribution to revenues – namely, that not every one of them is worth the same. Controlling for each individual wins' impact although is nearly impossible, and I believe that the linearity assumption of my OLS estimation is generally defendable. A further improvement could be made by generating win groups, in case one has the data for that kind of estimation.

5. References

Brown, R. W. (1993). An estimate of the rent generated by a premium college football player. *Economic Inquiry*, 31, 671-684.

Brown, R. W. (2012). Do NFL player earnings compensate for monopsony exploitation in college? *Journal of Sports Economics*, 13(4), 393-405.

Huma, Ramogi; Staurowsky, Ellen J., The Price of Poverty in Big Time College Sport (2011), NCPA study, available at: http://assets.usw.org/ncpa/The-Price-of-Poverty-in-Big-Time-College-Sport.pdf

Hunsberger, Peter K., Gitter, Seth R. (2015), What is a Blue Chip Recruit Worth? Estimating the Marginal Revenue Product of College Football Qarterbacks. *Journal of Sports Economics*, 16 (6), 664-690.

Wolfson, Julian; Addona, Vittorio and Schmicker, R. J. (2011), The quarterback prediction problem: Forecasting the performance of college quarterbacks selected in the NFL draft, *Journal of Quantitative Analysis in Sports*, Vol.7, Iss.3.

Lane, Erin B.; Nagel, Juan C. and Netz, Janet S., Alternative Approaches to Measuring MRP: Are All Men's College Basketball Players Exploited? (March 21, 2012). Available at SSRN: https://ssrn.com/abstract=1532886

Scully, Gerald W. (1974), "Pay and Performance in Major League Baseball", *The American Economic Review*, Vol. 64 (6), pp. 915-930.

Data Sources

QBR: Collected from ESPN Weekly leaders boards (scraped through Python). Can be found at: http://www.espn.com/ncf/qbr/_/type/player-week/week/1)

College revenues and expenditures from EADA website: https://ope.ed.gov/athletics/#/

Defence rankings and points: from NCAA website. Can be found at: https://international.ncaa.com/