DEPARTMENT OF ECONOMICS
JOHANNES KEPLER UNIVERSITY OF LINZ

And most of us go Pro in something other than Sports - Hiring Preferences and their Effect on the Labor Market for Collegiate Football Players

by

Mario Lackner

Working Paper No. 1010
August 2010

Johannes Kepler University of Linz
Department of Economics
Altenberger Strasse 69
A-4040 Linz - Auhof, Austria
www.econ.jku.at

mario.lackner@jku.at
phone +43 (0)732 2468 -8246, -9821 (fax)

And most of us go Pro in something other than Sports - Hiring Preferences and their Effect on the Labor Market for Collegiate Football Players.

MARIO LACKNER*

Department of Economics, JKU Linz

Revised version, August 27, 2010

Abstract

This paper analyzes the labor market for collegiate football players and argues that professional football teams have discriminating preferences when making their hiring decisions. An empirical analysis of panel data of 32 NFL teams in recent seasons is carried out to test the effects of such preferences on the performance of teams. The results provide strong evidence that certain criteria, which do have a high influence on a player's chances to start a career in Professional Football, have actually little influence on team-efficiency whatsoever. Consequently, this implies that discrimination in the form of hiring preferences create a sub-optimal result in terms of building a team, as well as for the overall labor market in Professional Football.

JEL Classification: J7, J2, L83, C23.

Keywords: Labor market in sports, discrimination in hiring, production efficiency, stochastic production frontier.

*Address for correspondence: Johannes Kepler University of Linz, Department of Economics, Altenbergerstr. 69, 4040 Linz, Austria. Phone:+43 70 2468 8246, Fax: +43 70 2468 9821, email: mario.lackner@jku.at. Helpful comments from discussants at the First European Conference in Sports Economics, René Böheim, Franz Hackl, Martin Halla, Gerald Pruckner and Rudolf Winter-Ebmer, as well as welcome support in data collecting by Julia Sageder, are gratefully acknowledged. The usual disclaimer applies.

1 Introduction

When reading one of the countless Draft evaluation magazines in the month leading up to the draft of the National Football League (NFL) quite regularly terms and phrases like *'a bit undersized '* , *'has a smallish frame '* , *'has exceptionally short arms'* , *'has too small hands to get a firm grip of the football '* ,*'lacks ideal bulk '* and rarely even *'is too tall '* are attributed to evaluated players. As the hiring of unproven amateur football players out of the NCAA collegiate football system is certainly a highly complex matter, it is safe to assume that there are certain signals that are of considerably high importance during the evaluation and scouting process. One of this easy accessible signals available as public information is the physical appearance and collegiate background of players in the draft pool. The interesting question is: How relevant is this signal and is it actually an efficient indicator for talent and potential to make a career in Professional Football? This is certainly of high relevance, especially from a player's point of view, as there are hundreds of college athletes losing eligibility for playing in the college sports system (NCAA). They basically end up being jobless if they do not get picked up by an NFL team in the annual draft process. If it should turn out that teams show a tendency towards relying on certain personal characteristics as signals, it is of high interest if these characteristics do actually have an influence on the production process of the teams. Should this not be the case or if there is only evidence for minimal influence, one can argue that the draft and in particular the player evaluation process contains major inefficiencies. Not only will they have severe implications on hundreds of careers of young athletes who often attain low levels of education because they dedicate their lives to playing semi-professional Football in the NCAA, basically without any compensation[1]. To sum up, the intention of this paper is to test if hiring decisions in the NFL are, at least partly, based on irrational reasoning. This is highly important because of the strong consequences for the labor market and and the severe implications for job perspectives of a large number of young college athletes.

Similar to the growing literature on discrimination in labor markets, where issues like discrimination for gender[2], color of skin (Goldsmith et al., 2006), sexual orientation (Weichselbaumer, 2003) , ethnical origins of applicants' names (Bertrand and Mullainathan, 2004) and even physical appearance are investigated empirically, this paper interprets hiring preferences for certain player characteristics as a form of labor market discrimination. Numerous papers investigate labor market discrimination in sports (Kahn,

[1]The semi-professional character of NCAA IA football is often questioned as NCAA colleges dedicate a huge amount of effort and resources to running their football programs. A number of papers argue that compensation for College players should reflect the actual rent they generate for the Colleges they play for. Brown (1993) discusses in detail how much rent a premium football player generates during his football career in the NCAA. Consult Meggyesy (2000) for a critical discussion of the NCAA system. Eckard (1998) and Kahn (2007) provide an excellent overview.

[2]Discussed by numerous recent contributions. Consult Blau and Kahn (1992) for an overview on the early literature.

1991), but little focus has been on discrimination for physical characteristics.

From the seminal work of Scully (1974), focusing on payment in the Major League Baseball (MLB) by introducing a production function concept, to recent work on European Soccer by Espitia-Escuer and GarcIa-CebriAn (2004), the National Basketball Association (NBA) (Hofler and Payne, 1997; Lee and Berri, 2008) and the National Hockey League (NHL)(Kahane, 2005; Heyne, Fenn and Brook, 2006) respectively, productivity and efficiency analysis in Sports Economics has used different approaches and technical methodologies to evaluate the production process of teams in numerous papers. Little, though, has been done on the efficiency of NFL teams, which is most likely a result of the highly complex nature and multilayered structure of the game of Football[3](Borland, 2006). While Hendricks, DeBrock and Koenker (2003) focus on individual player performances based on when they were drafted this paper relates draft decisions to team performance. As a common tool fitting the task of this paper *Stochastic Frontier Analysis* is employed to analyze the effect of hiring preferences on production efficiency of 32 NFL teams over the time span of 8 years. The analysis follows a methodology introduced by Huang and Liu (1994) and later refined in Battese and Coelli (1995) and Battese and Coelli (1996). Using this methodology allows to estimate team (in)efficiency by specifying a *Production Function* and testing for player specific personal characteristics as influences on production efficiency. While Battese and Coelli (1996) uses stochastic frontier analysis to identify farmer specific efficiency effects in an analysis of Australian agriculture productivity, numerous papers have adapted the methodology to measure production efficiency in many different areas. Dawson et al. (2000) employs stochastic frontier analysis to measure managerial efficiency in English soccer, while Scully (1994) focuses on managing efficiency in football, basketball and baseball employing survival analysis.

The paper is organized in the following way: Section 2 gives a short overview on the institutional settings of the NFL, the NCAA and the annual NFL Draft process. Section 3 introduces a simple model of team decisions based on maximization of output and presents two central propositions, section 4 empirically analyzes hiring preferences in the NFL draft and tests one of the propositions postulated in sectio 3. Section 5 gives a short overview on data and introduces the econometric approach as well as methodology to test the second proposition. Section 6 presents the main results, while section 7 concludes.

[3]If one breaks down an American Football team, each (NFL)-team is divided into three sub-units who operate on a high degree of independence to produce the final output. These three sub-groups are further divided into several different positions, each requiring different physical characteristics and abilities.

2 A short introduction to the NFL Draft

Hiring new players in the National Football League is mainly done in the draft system. This means that the majority of players are hired by selecting them directly out of College. So the main recruitment pool for young player, i.e. new employees, is the amount of players in the NCAA football system. The current NFL rules for draft eligibility state that a collegiate athlete has to be at least three years removed from high school leaving age. In the standard case, this means that a collegiate football player has to stay in College until he has finished his '*Junior year*' until he can declare to enter the NFL draft process. The NCAA standard rules, however, state that the maximum tenure of playing in a College, participating in the NCAA system, is four years. Athletes are forced to leave the team after their '*Senior year*'.[4] This rather tight window involves some crucial career decisions for young athletes: They have to decide whether to leave early after their third year, in case their value is high enough to make it to the pro-level, or to stay the maximum tenure and risk losing value instead of increasing it[5].

The actual technical procedure of the NFL Draft is rather simple: Prior to the process, at the end of the regular season, teams are ranked in reverse order with respect to their success in the most recent season. This ranking represents the order in which all teams are allowed to make their selections on the draft weekend, which is usually held on a Sunday in late April. All teams can potentially pick out of a huge pool of players meeting the eligibility conditions. The process is organized in a certain time schedule, with the first two rounds being done on Saturday and the remaining 5 rounds on Sunday. The time a team is 'On the clock ' to make a decision has previously been set at 20 minutes and was recently reduced to 10 minutes in 2008. Trades within the draft procedure are allowed and can be quite frequent. So with every round of the draft and with every single pick the assumption is that, neglecting teams specific needs for certain positions, the overall talent pool becomes smaller. The whole process continues until, after the final pick in round 7, all players, who are considered good enough for the league, are taken and therefore have a job[6]. The overall draft process is the most important opportunity for teams to hire new players and build their rosters, as *Free Agents*, who are veteran players free to be signed by any team, come at a much

[4]Consult `www.ncaa.org` for more details on NCAA regulations and eligibility rules. For reasons of simplification we assume the standard case here as the status quo.

[5]One possible and frequent reason for this is the possibility to sustain injuries during or before a player finishes his Senior year. Such an injury certainly has the potential to dramatically reduce a players value or even end a career.

[6]For the porpoise of simplicity Rookie Free Agents (FAs), i. e. players who get picked up after the draft are not taken into account here. This is certainly justifiable as Rookie FAs usually face huge obstacles in their path to making an Opening-day roster than drafted players do. In addition to this very few un-drafted players become significant contributors on their team. In some way a large amount of 'productive ' Rookie FAs would certainly contradict the idea that only the most talented players make it to the league. This would actually provide an even stronger argument in favor of the theory that there are inefficiencies in the NFL hiring process and in the way teams evaluate players. It is also true that being drafted does not guarantee a job for a whole NFL season, but it increases the chance of being employed dramatically.

higher price than rookies. The reason for this is the existence of minimum salaries, which are higher due to seniority rules built into NFL contract regulations. Another fact highlighting the strong implications of the draft is that players who do not get picked and spend some time out of football have little chance to make it to the league later on. The NCAA is basically acknowledging the fact that many leftovers are there at the Football labor market every year. The title of this paper is a reference to a series of official NCAA television spots propagating the ideas and principles of NCAA collegiate sports by pointing out the gains of students from participating in sports in college. The fact that DIV IA football is basically resembles fulltime professional sports, leaving few time for studying, however, is not mentioned.

3 Hiring Preferences: A simple model

Before we can build a simply model to illustrate draft decisions in the NFL we need to make three crucial assumptions. All three of them are necessary to sufficiently simplify the formal approach.

Assumption 1. *Teams are win maximizers. Consequently all decision makers influencing an NFL team organization are optimizing success of sporting operations.*

Assumption 2. *Teams produce output (wins) by employing work force. The more talented their players are, the higher is the level of output.*

Assumption 3. *The main hiring pool of employees for NFL teams is the annual rookie draft.*

While assumptions 2 and 3 are rather plausible, assumption 1 needs a closer look: Although sports economists have assessed that teams in major US sports tend to be profit maximizers one can certainly make an argument in favor of NFL teams as being win optimizers. The regulations of the NFL enforce a salary cap on all teams in the league and restrict the maximum payroll they can assign to salaries. A brief look at team payrolls in the NFL reveals that all teams are operating more or less exactly at the salary cap level. Consequently, if we only look at the sports related operations of teams we can certainly state that profit maximization cannot be the objective of any team in the NFL. From this perspective teams have to be win maximizers, as no other strategy would be reflected in the way teams are managed concerning their expenditures on players.

According to the specific design of the NFL draft process as described section 2 one can assume a certain form of output that teams produce by employing players the acquire through the draft out of college. The additional assumption would be that team officials consider individual characteristics of players to influence their productivity. Consequently it is possible to formulate the relation how a team's productions process depends on drafting players in each round of the draft as

$$Y_i = f(R_i, \quad \Phi_i, \quad Q), \tag{1}$$

where R_i is team i's draft rank in each round from 1 to 7. Φ is a vector including factors like the rank of other teams and team-individual personal targets. Equation 1 states that the sooner a team picks, the more it will benefit from picking. This benefit for the production process is decreasing with the teams draft rank, which means that that $\frac{\partial Y_i}{\partial R} < 0$, and $\frac{\partial^2 Y_i}{\partial R} > 0$ holds. Overall benefit decreases with every round. This assumption reflects the approach of teams to rank players and preassign them to an estimated draft round. The overall quality level of the draft is measured by Q which indicates the average potential of all players who are eligible and, consequently, the focus of this analysis. So we can define this overall talent measure Q as

$$Q = f(\sum_{n=1}^{N} X_i N^{-1}, \quad \sum_{i=1}^{N} Z_i N^{-1}). \tag{2}$$

In equation 2 the quality depth of a draft depends on two factors, where X_i defines player i's talent level and Z_i denotes a player's individual publicly known physical, as well as other characteristics. So Z is defined as a vector of player i's publicly known personal characteristics. Both sets of indicators, X and Z, are assumed to be the observable manifestations of a latent variable Γ, describing the actual overall talent level of a player and defining the success he is able to attain. The relation is given as

$$\Gamma_i = f(X_i, Z_i, \Theta_i), \tag{3}$$

which means that the observable college production of a certain player is related to his actual unobservable talent level X_i, his physical endowments Z_i and some stochastic term Θ reflecting effort, injuries or college team specific effects.[7] It is important to note that teams only have limited information on Θ and limited direct information on Γ. Consequently, the assumption is that they tend to evaluate players relying mostly on the signals X and Z. As team managers and scouts are assumed to believe that X_i is partly defined and interrelated with Z_i, it is plausible to assume that teams will foremost rely on the personal characteristics Z as signals for player potential. Especially given the fact that Z is free public

[7] These effects could be related to different coaching philosophies, as football is a highly diversified game in terms of different tactics and approaches.

information. Any team official (i.e. General Manager[8], GM) now has the intention to maximize overall team production by forming a roster. This will be influenced by fact that the sum of spending for contracts is limited by the previously mentioned existence of a league wide *salary cap*, denoted by $\overline{C}$, in the form as it currently exist for the NFL. This requires the crucial assumption that NFL team officials are performance maximizing. It is not implausible as profit maximization should also lead to win maximization (Fort, 2003). Any GM's maximization problem can then be formulates as:

$$\begin{aligned} max\, L = Y(Y^D(\overline{X}^D, \overline{Z}^D), Y^O(\overline{X}^O, \overline{Z}^O)) \\ s.t.: \sum_{i=1}^{N} X_{it} p_t^X + \sum_{i=1}^{N} Z_{it} p_t^Z) \leq \overline{C} \\ X \neq 0 \\ Z \neq 0 \end{aligned} \tag{4}$$

where Z and X with higher case ^{O} or ^{D} represent inputs on offense and defense. Y^D denotes *production from Defense*, while Y^O indicates *production from Offense*. From this the Lagrangian for the GM's maximization problem can be formulated as

$$\mathscr{L} = Y(Y^D(\overline{X}^D, \overline{Z}^D), Y^O(\overline{X}^O, \overline{Z}^O))\xi - \lambda[(\sum_{i=1}^{N} X_{it} p_t^X c^s + \sum_{i=1}^{N} Z_{it} p_t^Z) \leq \overline{C}]. \tag{5}$$

ξ is a stochastic parameter defined as $0 < \xi \leq 1$, that describes how big the effort level of team i is, given players' decisions on their effort contribution and the coaches' ability to maximize overall effort of the whole team. Here the assumption is that the only way a GM can influence effort is by changing the coaches. p_t^Z and p_t^X denote the prizes of one 'unit' of the respected talent indicators. These prizes are a function of demand and supply, meaning that both prize levels can be defined as

$$\begin{aligned} p_t^X = f(X_t^d, \overline{X}_t) \\ p_t^Z = f(Z_t^d, \overline{Z}_t), \end{aligned} \tag{6}$$

where X_t^d and Z_t^d are the available units of Z and Z in the draft of year t and $\overline{X}_t$ and $\overline{Z}_t$ the overall in-league levels of Z and X. Following the equation above, the only significant way a GM can directly influence production of his team is to decide which indicator of talent he relies on, given prizes for these

[8]Here it is important to acknowledge that the vast majority of NFL teams are governed by two positions: The Headcoach and the General Manager. Usually the General Manager is primarily responsible for personnel decisions for both staff and players. He is in charge to hire and fire players and his decisions do not necessarily have to be in accordance with the Headcoach.

talent indicators. If there are strong hiring preferences regarding player features in Z, it is a reasonable assumption to state that the prize will be rather high as large manifestations of Z will not be available in large quantities. This will result in a situation where $p_t^Z > p_t^X$. Z is public information available to everybody at zero cost, while X can only be assessed by studying college achievements og players, carrying out scouting and having a good general knowledge of a their abilities and potential. So, i.e., it is more convenient and, at first sight, cheaper [9] to rely on Z as the main, albeit not sole, indicator for player potential instead of X. The downside, however, is the higher risk that is involved and, of course, the higher price resulting from high overall league-demand for high manifestations of Z. Solving the maximization problem formulated in equation 5 yields the following first-order conditions for Z and X:

$$\frac{\partial Y(Z_{it}, X_{it})\xi}{\partial Z} - \lambda p_t^Z = 0 \ \ and \ \ \frac{\partial Y(Z_{it}, X_{it})\xi}{\partial X} - \lambda p_t^X = 0. \tag{7}$$

Rearranging and solving for λ yields

$$\frac{\partial Y(Z_{it}, G_{it})\xi / \partial Z}{p_t^Z} = \frac{\partial Y(Z_{it}, G_{it})\xi / \partial G}{p_t^X} \tag{8}$$

In order to capture the cost resulting from evaluating X, i.e. the cost of scouting, we introduce a cost parameter c^s. This parameter is a multiplicative cost parameter, which has a special character: These costs are not of fixed-cost nature but extra unit costs. The notion is very simple: The more and the better a team scouts the higher are the costs associated with it. c^s is assumed to have the following structure:

$$c^s(x) = \begin{cases} 1, & \text{if } x < \bar{x} \\ e^n, & \text{if } x > \bar{x} \end{cases} \tag{9}$$

The intuition of costs c^s is straight forward: The existence of a threshold value $\bar{x}$ indicates that information on players is available at zero cost up to this threshold level of $\bar{x} \in X$, where costs of scouting information on player talent starts to rise exponentially to infinity. This indicates that in reality the whole set of indicators technically can never fully discovered. Here one could argue that gaining and processing information on X is always associated with some fixed cost like costs resulting from running a scouting

[9] It is cheaper in the process because running a big scouting department and arranging a large number of player meetings is a rather costly procedure.

department. Including these fixed cost in would not be difficult, but it would certainly not change the results derived later. So here we define X as a set $X = \{x_1, x_2 ... \bar{x}, ... x_j\}$ and c^s as a function of X.

By including this cost parameter c^s equation 8 can be written

$$\frac{\partial Y(Z_{it}, G_{it})\xi/\partial Z}{p_t^Z} = \frac{\partial Y(Z_{it}, G_{it})\xi/\partial G}{p_t^X c^s} \tag{10}$$

The implications of equation 10 are rather trivial: If the GM wants to maximize team output by deciding on the set of players he drafts, he has to compare the ratios of benefits from X and Z, respectively, as well as the prices of these factors. His decision will depend on the cost of evaluating X, the price levels p_t^Z and p_t^X and the marginal product of X and Z. If we assume a situation where $\frac{\partial U}{\partial X} = \frac{\partial U}{\partial Z}$ and $p_t^Z = p_t^X$, the decision will be in favor of relying mostly on Z, as there are no costs associated with this and the GM has complete information on Z. If one wants to create a more realistic setup and include c^s as the cost of scouting, i.e. deviate from the basic assumption and return to a situation where $p_t^Z > p_t^X c^s$ holds, then the GM's approach could be different. As compared to a situation with equal marginal returns and equality in prices, which is described by $p_t^Z < p_t^X c^s$ a situation where the price for one unit of Z is greater than the price for one unit of X, X should be preferred as the main signal for talent until the additional costs from scouting raise the overall price of X over p_t^Z. Consequently, in a situation where $p_t^X c^s < p_t^Z$ indeed holds, teams should prefer to rely heavily on evaluating X as the signal for talent to build their rosters. From this model we can now derive the following two propositions:

Proposition 1. *In a situation with no hiring preferences for characteristics in Z exist a situation with* $p_t^Z = p_t^X$ *should be observable. If hiring preferences for Z exist prizes for Z should increase as we observe fixed supply of Z in the draft system. This should lead to a situation where* $p_t^Z > p_t^X c^s$ *holds and hiring preferences push the prize for Z above the prize of X including cost for scouting* C^s*.*

Proposition 2. *If it is the case that* $\frac{\partial Y}{\partial X} = \frac{\partial Y}{\partial Z}$ *either X or Z are good indicators for talent. If* $\frac{\partial Y}{\partial Z}$ *is equal to* 0*, or sufficiently close to* 0*, one can conclude that Z is not an efficient indicator for player talent.*

It will be the objective of following sections to find evidence in favor of Propositions 1 and 2. If it turns out that characteristics captured in Z are merely without a positive efficiency effect in the production process concerning technical efficiency, it is disadvantageous for a team to go for high values of Z in the draft and overpay for an indicator that is no good signal for talent. I. e., we should not observe a tendency of teams to draft players with high Zs earlier because it will lead to higher costs and work against the restriction of the salary cap. Instead of overpaying by coveting Z, teams should further increase their scouting expenditures to a level of c^s, where $p_t^Z = p_t^X c^s$ holds. Or they should simply make their scouting more efficient. If they invest heavily in scouting and find a talented player with lower physical measurements they will still benefit despite the increased costs because high investments into scouting are relatively cheap if high demand for Z raises p_t^Z.

Concerning the other focus, the labor market for players, the consequences are even stronger: If teams mainly base their hiring decisions on characteristics in Z, there is an increased possibility that this hiring bias will lead to adverse selection in the labor market of collegiate football players. This will result in a situation where talented players are overlooked because they do not meet certain 'minimum-requirements'in values of Z that are defined and coveted by teams. It would lead to inefficiencies in the labor market that would essentially leave many players unemployed and waste a certain amount of human capital because teams simply refuse to scout player below minimum characteristics. Even if players with low values of Z are indeed drafted they will not get the initial salary they deserve.

Another aspect is that these inefficient hiring preferences will provide a great incentive for cheating in the form of using prohibited substances to improve certain physical characteristics. Numerous recent medical studies (see for example Dickinson, Goldberg, Elliot, Spratt, Rogol and Fish (2005); Buzzini (2007)) point out that the usage of human growth hormones among college athletes has increased dramatically. Hiring preferences in the form as described above provide a perfect explanation for this: Imagine a young Highschool athlete who plans to take on college football and plays a position on the *Offensive line* in Highschool football. It is most likely that at the time he leaves Highschool and gets recruited by a college program he may not have reached his full body size. Still, at today's competitive environment in NCAA football, he will feel pressure to stay committed to his initial playing position of choice, even if it turns out that he will never quite reach the minimum measurements defined by NFL teams for Offense Line players[10]. One could argue that this will ultimately lead to a selection process where only those who have the adequate measures will succeed. While this is certainly a good argument, there will be a strong incentive to use growth enhancing substances if possibilities are present. But not only preferences for physical attributes in Z will lead to questionable situations. If NFL teams also show preferences for players out of prolific football colleges, which is certainly a component of Z as well, the recruitment process for major football colleges in Division IA colleges will get more competitive. This will consequently increase the likelihood of overcompetitive reactions from both players and colleges[11].

[10]Shifting positions is rather rare due to the complex structure of the game of Football. Although some players make successful transitions from one position to another position switching is a rather exceptional thing in the NFL. This is definitely true for the type of position categorization as it is done for the underlying data.

[11]This can explain incidents like the fake recruitment decision of Mike Hart, who announced at a press conference that he

4 Preferences in the Draft.

The implications of the presented model are rather trivial: If there is high demand for Z it is inefficient to make draft decisions based on Z as it will raise the cost for teams. Instead they should scout for a similarly talented (with an equal value of Γ) player, i.e. invest in the increase of knowledge of X, with lower values of Z. This player will be cheaper available later in the draft as long as the scouting costs do not push the price over the level as it would be if high levels of Z were targeted ($p_t^X c^s < p_t^Z$). In the special framework of the NFL and its salary cap this implication is even more convincing: As scouting costs would not enter a player's contract they would not count against the salary cap and therefore create an advantage for a team that invests in scouting instead of following the trend and count on Z. For a critical discussion of the strong emphasis on body measurements in the NFL draft consult Oates and Durham (2004)

So far little research has been done on the NFL draft and the teams' ability to evaluate and find talent [12]. In this section it is the major objective to clarify whether there actually are strong preferences for certain physical characteristics in the NFL player draft or not. To do this one can either argue with the countless expert analysis of college players hoping to become professional football players, or, in order to follow a more scientific way, do a basic econometric analysis of draft data. This section will follow both ways in an attempt to argue in favor of the existence of hiring preferences.

4.1 Experts

One of the most detailed Expert[13] analysis of college players published annually by Lindy's *Lindy's Pro Football Draft* (2009) lists several so called *minimum characteristics* of NFL players concerning their size and weight[14]. In addition to this, they also state the NFL optimum level of size and weight for each position. Taking a closer look at these optimum values of size it is striking that those characteristics can be perceived as well above average human measurements and even above average values for NFL players. If we compare them to the average values for physical characteristics of players we see that the optimal characteristics according to experts still are relatively high. If we look at Offense Line players, for example, we see that the Lindy's suggests an optimal size for them at 6 feet 6 inches and a minimal

had chosen the university of California (Berkeley) over the University of Oregon to play College Football there. Later it turned out that neither college football program had actually made an attempt to recruit him.

[12] Spurr (2000) analyzes the ability of baseball teams to find talent in the MLB draft, but the NFL draft has not been the focus of research yet.

[13] Bigler and Jeffries (2008) focuses on the evaluation of Black Quarterbacks draft prospects by experts.

[14] These optimal and minimal characteristics did not vary in the three most recent editions of Lindy's Pro Football Draft magazine

size of 6 feet 4 for an Offensive Tackle, 6 - 4 for a Guard and 6 - 3 for a Center. Table 3 shows that the average size of a running backs optimal size is defined as 6 - 2, which is two inches taller than the league average for starters at the Running Back position. Optimal quarterback size is assessed at 6 - 4, which is also significantly over the league average for starting Quarterbacks. An optimal wide receiver is said to be 6 - 4, while average starting NFL receiver only stood 6 - 2 from 2000 to 2007. Table 3 compares optimal levels for various important positions to actual averages in size. It is easy to see that from 2000 to 2007 this was exactly the situation in the NFL for Wide Receivers, Quarterbacks, Running Backs, Linebackers and Defensive Backs.

It is obvious that for most positions the average size is close to the size experts and analysts consider to be the minimal value. This certainly suggests a strong tendency to look for taller players in the draft if we assume hiring decisions follow expert opinions that taller means better. As there seemingly are few players of optimal size there has to be a high prize for these players. In the context of the NFL draft this means that they will be drafted early and cost their teams plenty of salary cap room because they will demand big contracts. Consequently, if we observe a situation where the average player is below the optimal level of size, this will create a downward push on the value of smaller players and an upward push on [15] the prize of taller players.

Table 1: Experts vs. averages.[a]

	Experts optimum	Experts minimum	Average Size
QB	6 - 4	6 - 0	appr. 6 - 3
RB	6 - 2	5 - 9	appr. 5 - 11
WR	6 - 4	5 - 10	appr. 6 - 2
LB[b]	6 - 3, 6 - 2	6 - 0, 5 - 11	appr. 6 - 2
DB[c]	6 - 1, 6 - 2	5 - 9, 5 - 11	appr. 6 - 0

[a] Size measured in *feet - inches*. Expert values taken from Lindy's 2009 NFL Draft magazine. *Linemen (Offense and Defense)* were omitted due to high degree of differentiation. [b] First value is for *Outside Linebackers*, second for *Inside Linebackers*. [c] First value is for *Cornerbacks*, second for *Safeties*.

4.2 More elaborate data analysis

As the experts assessments are basically an outsider opinion which have no scientific incidence and do not necessarily mean that the people in charge of teams, the General Managers, do actually decide along

[15] Or players measuring exactly the optimal size as assessed by experts.

these preferences published by analysts. Moreover, one could argue that a sports periodical is far from being an adequate source for a proof of a scientific proposition. In order to have a stronger argument an econometric analysis was done to thoroughly test the theory that there are strong preferences for players with rather extreme physical attributes[16]. In order to do this NFL roster data acquired from `http://www.pro-football-reference.com/` consisting of over 4,200 individual players were analyzed to show if physical characteristics actually did improve draft status of players in the past. Tables 4 and 5 presents the results from an ordered logit model testing the hypothesis of hiring preferences.

From the results one can see that a player's size and weight during the period from 2000 to 2007 had a significant positive[17] influence on when he was drafted. Simply put: The taller a player was, the earlier he was drafted and, consequently, the higher was his salary. The results also indicate that weight had a positive influence as well if we look at the coefficient of a the weight of a player. While the statistically significant effect of weight disappears when the sample of players was reduced to drafted players (see table 5), the influence of size turns out to be robust. Overall the differences between an analysis including undrafted players and the one excluding them are minimal. Only weight and top 10 players turn out to be insignificant in the sample without undrafted players.

Table 2 presents the predicted probabilities to be drafted in the first round of the NFL draft for three major positions in football, based on the ordered probit results presented in tabel 4. It is obvious that a player who stands 190 centimeters faces a substantially higher probability of being a first round draft pick than a player with the size of 170 centimeters. The table also distinguishes probabilities for players from a top 10 college program and for those from lower ranked programs. Again we can observe that probabilities to get drafted in round 1 are much higher for both measurements of size when a player is from a top 10 college. It is also an interesting result that the difference in probabilities concerning sizes 170 and 190 centimeters is larger when players are from a top college. Thus it seems that in the case of players from these colleges there is even more emphasis put on size than for those who are not. This is consistent with the fact that players from top ranked colleges are more represented in the media and players from fringe college programs are sometimes well kept secrets that need more scouting initially simply to find them. The biggest difference was found for Quarterbacks out of top 10 colleges, where QBs measuring 190 cm face a 5.7 % higher probability to get drafted than those who stand 170 cm. A QB from a top 10 college with size of 190 cm is drafted in round 1 with a probability of over 15 % which is a very high probability. A Quarterback who stands 170 cm faces only about half the probability as one who stands 190 and is coming out of a top 10 college

[16]In the context of the model presented in section 3 our Z.

[17]Indicated by a negative influence in the Ordered Probit Results

Table 2: Predicted probabilities for being drafted in round 1^a.

Wide Receivers			
Top 10 college	**Predicted probability**	**Predicted probability**	**Difference**
	Size: 170 cm	*Size: 190 cm*	
No	0.0762	0.1235	0.0473
Yes	0.0923	0.1478	0.0555

Running Backs			
Top 10 college	**Predicted probability**	**Predicted probability**	**Difference**
	Size: 170 cm	*Size: 190 cm*	
No	0.0589	0.0965	0.0376
Yes	0.0716	0.1163	0.0447

Quarterbacks			
Top 10 college	**Predicted probability**	**Predicted probability**	**Difference**
	Size: 170 cm	*Size: 190 cm*	
No	0.0787	0.1272	0.0485
Yes	0.0951	0.1521	0.057

Predicted probabilities based on mean ordered logit results including undrafted players (8 draft rounds).

After reviewing the simple, yet effective, methodology how preferences for high levels of certain characteristics and physical measurements were analyzed, one could argue that those results are merely a result of omitted variable bias, because the estimation models do not include any measurement of production in college as a proxy for talent. As Hendricks, DeBrock and Koenker (2003) point out, NFL administrative sources and media do have an enormous amount of data on hundreds of college football players including highly detailed information on intelligence evaluation tests, records of behavior and many other details including results from work outs like the NFL combine[18]. The argument in favor of such a simple methodology is the following: As only in-league players are subject of the analysis one can make the assumption that, on average, the talent level of these players should be rather close together. All players were at one point evaluated as having enough talent for the league and so leaving out all the possible talent-signal variables would not spoil the results obtained by the multi-nominal models discussed above. This should at least be true for Xs that are scouted at no or very low costs.[19] While control variables measuring productivity on the collegiate level as well as workout statistics indicating physical ability would be of great value results might actually not differ much qualitatively.

One possibility to control talent is to include career length, which should at least partly control for talent and consequently productiveness in College Football. This is supported by the fact that the variable

[18]The NFL Combine is an annual pre-draft event where invited college athletes work out in front of NFL coaches, scouts and analysts.

[19]One possible area of future research on hiring preference would be to integrate college production into an analysis similar to the one carried out in this paper. So far data availability has provided major setbacks to this.

career length exhibits a negative coefficient in model II of table 6, meaning that players with longer careers were drafted in early rounds. In addition to this, it appears crucial to control for high ranked college players by including a dummy which takes the value 1 if the player played on a top 20 ranked college in the NCAA. Finally one has to disentangle the complex nature of the game and its demand for players of various body types by including position dummies at a reasonable level of differentiation of 9 position groups [20].

If we look at drafting a player to be essentially an investment, we can apply some theoretical implication of the theory of investment under uncertainty[21] (see Dixit and Pindyck (1994)) to the findings of the regression analysis of this section it can be seen that the results are perfectly fitting the theory's implications: Given the assumption that all players should have a combination of talent indicators that are within a tight band, it is a reasonable strategy for teams to go for the one signal that might tell them the most about future potential because it is rather invariant, i. e. factors like size, weight or the NCAA football rank of college. The fact that players are drafted later if their measurements are lower, i.e. perceived as suboptimal, or they are from a low ranked college football program fits the idea of investment theory that teams will want to give those players some time on their roster to prove themselves because releasing them after only a short time is less expensive[22]. So if one Offensive line player who stands $6-2$ has the same college stats as a second Offensive line player who stands $6-7$ teams will pursue the taller player first not only because they think his superior size is an indicator of more talent, but also due to the fact that a failure of the smaller player will simply be cheaper as he will receive a lower contract involving less guaranteed salary if he is drafted later. Should size turn out to be the right signal and the smaller player is the weaker one or even fails to make the roster this will cost the team less. So teams will look for the size advantage in any case, even if size is not the adequate talent signal.

5 Data and Methodology

The data for this paper were collected from the *Pro-Football References* web-page[23]. All variables present per-year averages created from NFL roster files of at least 60 players. In order to capture the fact that only inputs that are actually a substantial part of the production process only players who started 7 or more games were included. For more details on data refer to the Data Appendix.

[20]These groups are: Running Back, Quarterback, Fullbacks, Wide Receiver, Defensive Linemen, Offensive Linemen, Kickers, Defensive Backs and Linebackers.

[21]A drafted player can certainly be seen as an investment and there is sufficient uncertainty involved in the draft process.

[22]This is simply a result of contracting rules in the NFL. Releasing a player with a large contract is more expansive than firing a player with a contract rather close to the minimum.

[23]To be found online at `http://www.pro-football-reference.com/`.

In order to test the effect of individual player characteristics on overall team success the most obvious and straight forward approach would be a model explaining some success measure, like the amount of total wins per season, by some explanatory variables and the set of characteristics. While this approach might look appealing at first sight, it does not really capture the complex nature of the game of American Football as it is played in the NFL. In the recent literature on sport economics and team productivity efficiency analysis has become an often employed methodology to model the complex nature of productivity in sports econometrically and get a better notion of how success in sports is really achieved.

This paper will adapt a methodology often applied in agricultural analysis. I will estimate a stochastic production frontier in the context of American Football. The model will take the general form of

$$Y_t = \beta X_t - U_t + V_t, \qquad t = 1, ..., T \tag{11}$$

where X_t is a vector of inputs to produce output Y_t, U_t is a vector of inefficiency parameters and V_t is an ideosyncratic error term with $iidN(0, \sigma_v^2)$. Following Battese and Coelli (1995), the inefficiency term is modeled as linear dependent of a set of explanatory variables, in this context the player characteristics in the $1 \times K$ vector Z_t:

$$U_t = \gamma_0 + \gamma Z_t, \qquad t = 1, ..., T. \tag{12}$$

The general formulation of the model would be

$$Y_{it} = \sum_{j=1}^{J} \beta_j x_{jit} - u_{it} + v_{it}, \qquad i = 1...K\ t = 1, ..., T, \tag{13}$$

where the subscript i denominates the team, t the season and j is an index for the different inputs. The inefficiency terms u_{it} are modeled as

$$u_{it} = \gamma_0 + \sum_{k=1}^{K} \gamma_k z_{kit} \qquad i = 1...K\ t = 1, ..., T. \tag{14}$$

where z_{kit} is the personal characteristic k of team i in the season t. The model presented in equation 14 was also used in Battese and Coelli (1995) to measure agricultural efficiency with the consideration of farm-specific productivity effects, as well as in numerous other papers. Here this notion is transferred to the world of sport by basically respecting the presence of team specific efficiency effects resulting from heterogeneous players as important factors in the production process. This heterogeneity is induced by the differences in teams' hiring preferences and in their way of building and managing their work force, i.e. the team roster. The main focus of the analysis is to test the influence of the way a team is built concerning certain before-mentioned player characteristics on production efficient. If the results show that the same characteristics that influence the hiring process in the NFL draft have little or even no effect on team efficiency then the hiring behavior, or hiring preferences, are in itself inefficient.

One of the most challenging tasks in efficiency analysis, and specifically in stochastic frontier analysis, is the specification of a suitable form of *production function*. This is even more difficult if such a highly complex production process as in sports is the matter of interest. While it may already be troublesome to formulate such a production function for sports like Basketball, Ice Hockey or Soccer, it is even more challenging to capture the elaborate multi-leveled nature of American Football. As an American Football team is divided into three subdivisions we have at least three separate units of production with completely different sets of input factors. As already mentioned before, one of the three subdivisions, namely special teams, will be completely left out due to reasons of simplification. So we are left with *Offense* and *Defense* as the two main components of any football team. One of the simplest approaches to the production process in Football is the addressing of the main target of the whole game: the gain or defense of space on the playing field, measured in yards. While it is the intention of the offense squad to gain yards on every down, the Defense will try to avoid giving them up. This statistical measurement is, on the one hand, the true core of how the game works, on the other hand it is the most rudimentary indicator of performance, effort and ability. Still it does not fully predict the final product, which will be measured in win percentage as the ratio of wins over total games played. As a conclusion the two main inputs in the underlying production function will be *yards per play on offense* and *yards per play on defense*. In order to capture other performance related effects an additional variable will be included as a control variable that represent a form of *performance shock* as a result of errors, lack of concentration or motivation, by both teams competing. This will be the ratio of *interceptions made to interceptions committed*. The *ratio of fumbles forced to fumbles made*[24] would also be available, but the results of earlier regressions came up with somewhat puzzling results for this second control variable, hinting there could be a measurement problem or specification issues. Therefore only the interception ratio was included to

[24]In American Football the term Interception is used to describe the event when a forward pass is intercepted by a defender and the ball is therefore turned over to the other team. A fumble describes a turnover as well, with the only difference that it is a result of merely dropping the ball and having an opposing defender pick it up.

capture shocks within the production function. This leaves us with a production function of the following form:

$$Y = Y^{O\beta_1} Y^{D\beta_2}. \tag{15}$$

As mentioned before, the final product or output Y will be measured in the amount of total wins per season over the maximum number of wins possible, which is 16. This form of production function uses two inputs: production on offense, i.e. yards gained, and production on defense, i.e. yards allowed. Linearizing this Cobb-Douglas type production function above and including the error shock ratios as additional control variables, we can define the core of the frontier model to be estimated as

$$ln(Y)_{it} = \beta_1 ln(y_{it}^O) + \beta_2 ln(y_{it}^D) + \beta_3 I_{it}^{ratio} + \varepsilon_{it} u_{it} + v_{it}, \tag{16}$$

where the error term ε_{it} is decomposed into

$$\varepsilon_{it} = v_{it} - u_{it}. \tag{17}$$

The conditional mean error term in the frontier model employed in this section regresses the mean of the inefficiency term u on a set of exogenous variables. For the purpose of this paper two different sets are created to analyze their effect on production efficiency of NFL teams. Set 1 (non physical characteristics as well as control variables) includes the average round all players with more than 6 starts per season were drafted, their average career as well as a career squared term allowing for a cubic relationship, and a variable that counts the number of players drafted out of a top 10 as well as 20 college program. Set 2 is a set of average values of physical measurements for players with 7 or more stats per season. Both sets are somewhat related, as it was possible to show that certain physical attributes influence the round a player is drafted in. The empirical analysis also consists of a model including both sets of variables and one additionally controlling for coaching changes[25] Fizel and D'Itri (1999) gives a good overview on how efficiency and changes in leadership (i.e. management or in our case coaching) are influencing each other.

[25]It is critical to notice that usually it is a difficult task to include coaching changes in regressions explaining success in sports, as there has to be a high degree of endogeneity. This is due to the fact that coaches influence success but success will also have an influence on the job safety of the coach. In our analysis, however, we can safely neglect this endogeneity problem because coaching is only allowed to have an influence on the efficiency term and not on the dependent variable. Here we are only interested in the influence of having changed the Headcoach of a team on the team's production efficiency.

6 Results

The results derived from stochastic frontier analysis introduced in section 5 are presented in Table 6[26]. In the basic model specification (I), where only the simple production function is estimated, all signs and magnitudes of coefficients are as expected. This is also the case for all other model specifications. The result supports the specification of the basic production function. The only surprising result is the negative coefficient on the *Fumble ratio*. As an increase of this variable measures an improvement in the ratio of fumbles forced to fumbles committed, one would certainly expect a positive influence of this control variable on the win ratio. One explanation for this puzzling result could be that both variables from which the ratio was constructed do not necessarily indicate turnovers as they are not net-values: They represent only the overall number of fumbles without indication whether possession of the ball changed or was retained.

6.1 Physical characteristics

The results in Table 6 concerning physical characteristics of players tell a clear story: There is no indication that physical characteristics do have much of an influence on technical efficiency. The only position that seems to have some influence is the position of Wide Receiver. Here two model specifications (III and V) yield a statistically significant influence on efficiency. While a Wide Receiver's weight seems to reduces inefficiency, there is some evidence that his size actually increases it. In addition to this the results show that the average weight of a team's starting running backs increases technical inefficiency. An explanation for this could be the fact that a higher weight for a running back might indicate less speed or agility which could be the reason for the negative influence on efficiency. All other physical characteristics, however, show no significant influence on technical efficiency and they are unable to deliver an explanation for the presence of efficiency differences.

One notable result is that in model (V) all before mentioned statistically significant influences of average physical characteristics of starting players are still present if we include a binary variable taking the value 1 if the year is one after the team has changed coaches. This coaching change dummy has a highly significant positive influence on inefficiency and should take care for most team heterogeneity. Only the significant coefficient for average Wide Receiver weight turns out insignificant.

[26]All three models were estimated using the *FRONTIER* command in STATA.

6.2 Other control variables

In addition to pure physical characteristics it is important to control for additional player characteristics that could influence his draft status. One important control variable is the status of the college program a player originates from. One might argue that the more players from top-ranked colleges a team has, the more efficient it should be. Another important influence on efficiency could be a player's experience, which will be measured as average career length of starting players on the roster in a particular season. It is reasonable to include a quadratic term because one can assume a non-linear relationship between efficiency and experience. One might even go further and assume a positive but declining relation between experience and efficiency. This could be due to declining physical ability or the increasing number of minor or major injuries over the time span of a career. As an addition to this it is also important to include one of the seemingly appropriate quality measures of a team, namely the average draft pick of starters, into the empirical analysis.

The results in table 6 show that the the whole set of control variables and non-physical characteristic explaining the systematic inefficiency term u_{it} in Model specifications (II) through (V), have very little influence on technical efficiency. The average round players were drafted turns out to have no statistically significant influence on technical efficiency at all. Only specification (II) yields a significant negative, but small, influence of the average draft pick of a team on production efficiency. But as soon as physical characteristics are included into the model of the conditional efficiency term this positive effect on efficiency disappears. Neither the experience of a player (as well as the quadratic term), nor the distinction if he played college football at a top 10 or 20 ranked College seems to have any influence on technical efficiency. So the reasonable assumption that more experience players from 'better 'colleges are improving a team's efficiency turns out to be wrong. This could be interpreted the following way: Teams with more experienced and higher drafted players do not dominate other teams using the same basic inputs in the production function, because they are not playing more efficiently.

In general, the results even raise doubt on the relation the draft position of teams and their production efficiency: There is little evidence that having more players on the roster that were drafted in early rounds increases technical efficiency. While this is surprising, it might reflect that the criteria that influence draft choices are partly irrelevant and leading to undesired outcomes. This becomes evident as the coefficient of the variable measuring the average player's draft round is not significant in any model specification presented in Table 6. Another clear indication of draft inefficiencies is that there is no evidence to be found supporting a positive influence of the number of players from top colleges on efficiency: The results of model specification (II) show an actually negative influence of the number of players from elite colleges[27] on the efficiency parameter. Simply put, having many players from the top-level football colleges on the roster increases technical inefficiency. However, this positive effect on the inefficiency parameter disappears if we control for physical characteristics in specifications (IV) and (V) when we control for coaching changes. In addition to this the number of players out of the 20 top colleges has no statistically significant influence on efficiency.

7 Conclusion

After it was possible to find evidence in favor of proposition 1, i.e. that there is a number player characteristics that clearly influence the hiring decisions of NFL teams (see section 4), it was possible to show that most of these characteristics are not systematically determining efficiency. This is clearly to be interpreted as a strong argument supporting major aspects of proposition 2. The main conclusion from the results of the frontier analysis carried out is that the player characteristics, that are one of the main targets for teams in the NFL draft, do actually have little overall influence on technical efficiency. Although measuring technical efficiency does only cover certain aspects of the successful production process of a professional football team, it is certainly possible to conclude that the criteria for draft decisions are partly irrelevant and leading to suboptimal results. Looking at it in more detail, we can argue that the drafting preferences of teams create inefficiencies via two different channels: First, there is within-team inefficiency because the demand for players with high measurements causes the price for players offering these measurements to rise relative to others who do not have these features. This leads to a higher team salaries which will, in the presence of a binding salary cap, leave the team endowed with a lower overall talent level at the same maximum overall team salary restricted by the salary cap[28]. The second channel of inefficiency which has even more severe implications and is the focus of this paper is operating through the labor market of college players. If teams prefer certain characteristics in NCAA players they will tend to overlook players who do not fit their model of an ideal player and still may have a sufficiently high level of talent. This will lead to a situation where adverse selection dominates and a lot of human capital or talent is left unutilized in the labor market. It might even be the case that certain players completely fall off the table and end up unemployed after their eligibility to play in the NCAA ends. On the one hand teams will suffer from this as they miss out on relatively cheap resources of employees. On the

[27] Here ‘elite ’meaning one of the most successful colleges in the highest level of collegiate competition in NCAA history.

[28] This argument even extends to the point where we would consequently observe a team salary structure which is highly unequally structured with some highly paid players taking the major share of the salary cap room. As Borghesi (2008) finds this inequality in payments has a negative influence on team success.

other hand the most problematic consequences arise for players who might have enough talent to play in the NFL, but are either too short, have not put on enough weight or played collegiate football at a low ranked college. Keeping in mind that a Div.IA NCAA Football player has had little or basically no time to do anything else than dedicate his full attention to football at basically zero monetary compensation, we can conclude that this has serious consequences on hundreds of young players who end up being jobless and have not graduated due to lack of time to focus on studying. It is certainly hard to justify why players actually end up being left-overs due to lacking certain characteristics if some important criteria for draft decisions based on these characteristics actually turn out to be more or less irrelevant for the later production process. While one might argue that one major compensation for NCAA collegiate football players is the opportunity to earn a university degree for free, this argument is somewhat weak in the presence of only 64 percent graduation success rate in NCAA Division I football in 2005 [29].

From an organizational perspective one strategy for an NFL team to avoid hiring inefficiencies based on the overvaluation of irrelevant characteristics would be to trade down in the draft, invest highly in scouting and predominantly hire less coveted players in later rounds who do not display optimal physical characteristics. This will enable the team to end up with a set of players that will sign rather cheap contracts and still might be as talented as others. This strategy will work as other teams will simply ignore these players in early rounds and they will still be available later in the draft. The results of this paper at least suggest that production efficiency will not suffer from such a strategy. As the NFL is a league with very strong salary cap regulations saving contract volume is a necessity for teams to stay competitive over a longer period of time. Is is certainly possible to conclude that physical characteristics size and weight, as well as the ranking of college a player originates from are weak signals for talent. From the labor market perspective one could assess that the NCAA as well as the NFL should have an interest to optimize the labor market for pro football players. Although there is no clear policy implication how to achieve this, it is certainly true that hiring graduation rates as well as building up some kind of league financed social security system for collegiate players who fail to make it to Professional Football. Not to mention that the current NCAA policy of zero monetary compensation and the cartel power of colleges are highly questionable.

Summing up, this paper argues that there is little evidence that certain player characteristics have any influence on team efficiency, although there are convincing indications that certain preferences for them do exist. The analysis only finds some influence for two physical parameters of two positions in

[29] Data from the 2005 NCAA Division I Graduation Success Rate (GSR) Data report available at `www.ncaa.com`. Values for 2006 and 2007 are 65% and 67%. While there is certainly an increase in graduation rates it is still the case that one third of all collegiate football players do not graduate. And this value could still be biased from colleges who have a strict pro graduation policy like Notre Dame.

football, namely Wide Receivers and Running Backs but no statistically significant influence of other non-physical characteristics. There, however, is some evidence that coaching changes have a significant negative influence on efficiency. In the context of labor markets this paper demonstrates how irrational hiring preferences can drive up the cost of certain types of workers and discriminate others to an extent where they might end up being unemployed.

References

Battese, BatteseGeorge E. and Timothy J. Coelli (1996), ‘Identification of Factors which influence the Technical Inefficiency of Indian Farmers’, *Australian Journal of Agricultural Economics* **40 (2)**, pp. 103 – 128.

Battese, George E. and Timothy J. Coelli (1995), ‘A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data.’, *Empirical Economics* **20 (2)**, pp. 325 – 332.

Bertrand, Marianne and Sendhil Mullainathan (2004), ‘Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination’, *The American Economic Review* **94**(4), 991–1013.

Bigler, Matthew and Judson L. Jeffries (2008), ‘Evaluations of Black Quarterbacks’, *Journal of African American Studies* **12**(2), 120–141.

Blau, Francine D. and Lawrence M. Kahn (1992), ‘The Gender Earnings Gap: Learning from International Comparisons’, *The American Economic Review* **82**(2), 533–538.

Borghesi, Richard (2008), ‘Allocation of scarce resources: Insight from the NFL salary cap’, *Journal of Economics and Business* **60**(6), 536 – 550.

Borland, Jeff (2006), *Handbook on the economics of sport*, Edward Elgar Publishing Limited, chapter The Production of Professional Team Sports, pp. 22–26.

Brown, Robert W. (1993), ‘An Estimate of the Rent Generated by a Premium College Football Player’, *Economic Inquiry* **31 (4)**, pp. 671 – 684.

Buzzini, Sergio R.R. (2007), ‘Abuse of Growth Hormone Among Young Athletes’, *Pediatric Clinics of North America* **54**(4), 823 – 843. Peformance Enhancing Drugs.

Dawson, Peter, Stephen Dobson and Bill Gerrard (2000), ‘Stochastic Frontiers and the Temporal Structure of Managerial Efficiency in English Soccer’, *Journal of Sports Economics* **1 (4)**(4), 341–362.

Dickinson, Barry, Linn Goldberg, Diane Elliot, Daniel Spratt, Alan D. Rogol and Lisa H. Fish (2005), ‘Hormone abuse in adolescents and adults: A review of current knowledge’, *The Endocrinologist* **15**(2), 115–125.

Dixit, Avinash K. and Robert S. Pindyck (1994), *Investment under Uncertainty*, Princeton University Press.

Eckard, E. Woodrow (1998), 'The NCAA Cartel and Competitive Balance in College Football', *Review of Industrial Organization* **13**(3), pp. 347–369.

Espitia-Escuer, Manuel and LucIa Isabel GarcIa-CebriAn (2004), 'Measuring the Efficiency of Spanish First-Division Soccer Teams', *Journal of Sports Economics* **5**(4), 329–346.

Fizel, John L. and Michael P. D'Itri (1999), 'Firing and Hiring of Managers: Does Efficiency Matter?', *Journal of Management* **25**(4), 567–585.

Fort, Rodney (2003), 'European and North American Sports Differences(?)', *Scottish Journal of Political Economy* **47**(4), 431 – 455.

Goldsmith, Arthur H., Darrick Hamilton and William Darity (2006), 'Shades of discrimination: Skin tone and wages', *American Economic Review* **96**(2), 242–245.

Hendricks, Wallace, Lawrence DeBrock and Roger Koenker (2003), 'Uncertainty, Hiring, and Subsequent Performance: The NFL Draft', *Journal of Labor Economics* **21**(4), pp. 857–886.

Heyne, John J., Aju J. Fenn and Stacey Brook (2006), 'NHL Team Production', *Colorado College Working Paper* **No. 2006-07**, –.

Hofler, Richard A. and James E. Payne (1997), 'Measuring efficiency in the National Basketball Association', *Economics Letters* **55**(2), 293 – 299.

Huang, Clief J. and Jin-Tan Liu (1994), 'Estimation of a Non-Neutral Stochastic Frontier Production Function', *Journal of Productivity Analysis* **5 (2)**, pp. 171–180.

Kahane, Leo H. (2005), 'Production Efficiency and Discriminatory Hiring Practices in the National Hockey League: A Stochastic Frontier Approach Zeitschrift', *Review of Industrial Organization* **27 (1)**, pp. 47–71.

Kahn, Lawrence M. (1991), 'Discrimination in Professional Sports: a Survey of the Literature', *Industrial and Labor Relations Review* **44**(3), 395–418.

Kahn, Lawrence M. (2007), 'Markets: Cartel Behavior and Amateurism in College Sports', *Journal of Economic Perspectives* **21**(1), pp. 209–226.

Lee, Young Hoon and David Berri (2008), ‘A Re-Examination of Production Functions and Efficiency Estimates for the National Basketball Association’, *Scottish Journal of Political Economy* **55 (1)**, pp. 51–66.

Lindy’s Pro Football Draft (2009), J. Lindy Davis Jr.

Meggyesy, David (2000), ‘Athletes in Big-Time College Sport’, *Society* **37 (3)**, pp. 24–28.

Oates, Thomas P. and Meenakshi Gigi Durham (2004), ‘The mismeasure of masculinity: the male body, ’race’ and power in the enumerative discourses of the NFL Draft’, *Patterns of Prejudice* **38**(3), 301–320.

Scully, Gerald W. (1974), ‘Pay and performance in major league baseball’, *The American Economic Review* **64**(6), 915–930.

Scully, Gerald W. (1994), ‘Managerial Efficiency and Survivability in Professional Team Sports’, *Managerial and Decision Economics* **Volume 15**(5), pp. 403 – 411.

Spurr, Stephen J. (2000), ‘The Baseball Draft’, *Journal of Sports Economics* **1**(1), 66–85.

Weichselbaumer, Doris (2003), ‘Sexual orientation discrimination in hiring’, *Labor Economics* **10**(6), 629–642.

8 Data appendix

NFL Roster data: Data on players were collected from `http://www.pro-football-reference.com/` in May 2008. This data was rearranged, checked and formed into a panel of 32 NFL teams and their players from 2000 to 2007. In order to capture the desired efficiency effects only starting players were include. A starter was defined as a player who started at least 6 games in a season. `http://www.pro-football-reference.com/`. gives information about games participated and games started, so it was easy to form a panel only consisting of starters. Although there is no clear definition on which player can be called a starter in the NFL it is arguably a safe assumption to say that 7 starts qualify for the category starter as 7 games are just below 50 percent of all season games. Including also players with lower levels of contribution might lead to a situation where player features and characteristics are included which are not sufficiently part of the production process because they are not on the field of play. Players who were drafted before 1987 were omitted in the analysis of section 4.

All *physical characteristics variables* were created from the overall roster panel by collapsing the individual values into team by year averages. All size values were transformed to cm and weight is measured in US pounds.

The coaching change variable is based on data on coaching tenures for all 32 teams in the NFl also available at `http://www.pro-football-reference.com/`. The binary variable takes the value 1 in a year where a new coach has his first season or part of it on a new team, else it takes the value 0.

Team statistics were collected from `http://www.pro-football-reference.com/`. and several variables were constructed from the multitude of team stats available. The dependent variable in the efficiency model, win ratio, is a variable measuring the ratio of number of games won against games played in a season. This essentially boils down to $\frac{wins}{16}$ because every team in the NFL played exactly 16 regular season games in the 7 seasons from 2000 to 2007.

The binary variable indicating if a player was drafted out of a top 10 or top 20 college program was generated based on multiple all-time college rankings available online. All of these rankings are mostly based on a college's performance over it's total existence.

9 Appendix

Table 3: Descriptive Statistics - NFL Roster Data.

Variable	Observ.	Mean	Std. Dev.	Min	Max
Win ratio	253	0.48	0.19	0.06	1.19
Yards per play Offense	253	5.12	0.52	3.8	7
Yards per play Defense	253	5.13	0.39	4.2	6.3
PLAYERS' ATTRIBUTES					
Average Draft Pick	253	38.23	24.05	1	115.5
Career	253	4.85	1.4	1.6	9.2
Career Squared	253	25.53	14.56	2.56	84.64
Numb. of Top 10 Coll. Players	253	4.61	1.86	0	13
Numb. of Top 20 Coll. Players	253	8.48	2.46	2	16
PLAYERS' PHYSICAL ATTRIBUTES (Averages over all 32 teams from 2000 to 2007)					
Avg. OL Size	253	194.21	1.61	190.5	197.61
Avg. OL Weight	253	312.09	8.43	286	339
Avg. DL Size	253	191.84	2.1	186.69	198.12
Avg.DL Weight	253	292.34	9.64	268.8	325
Avg. WR Size	253	188.6	1	185.53	191.45
Avg. WR Weight	253	254.29	4.25	243.7	266.67
Avg. QB Size	251	190.58	3.74	177.8	198.12
Avg. QB Weight	251	223.5	12.13	180	264
Avg. LB Size	253	187.36	2.01	181.19	191.35
Avg. LB Weight	253	242.25	6	227.67	257
Avg. RB Size	252	180.95	3.97	172.72	193.04
Avg. RB Weight	252	221.1	14.19	180	290.5
Avg. DB Size	253	182.15	1.97	175.26	189.23
Avg. DB Weight	253	199.88	4.95	186.5	217.25

Table 4: Ordered logit model to trace draft preferences (part 1)[a].

	Model 1	**Marginal effects for outcomes** 1 - 8[b]							
		1	2	3	4	5	6	7	8
Size	-0.027***	0.002	0.002	0.001	0.001	0.000	-0.000	-0.001	-0.006
	(0.007)								
Weight	-0.010***	0.001	0.001	0.001	0.000	0.000	-0.000	-0.000	-0.002
	(0.002)								
Career length	-0.368***	0.029	0.026	0.020	0.013	0.004	-0.003	-0.010	-0.079
	(0.033)								
Career length sq.	0.017***	-0.001	-0.001	-0.001	-0.001	-0.000	0.000	0.000	0.004
	(0.003)								
Top 10 College	-0.208**	0.018	0.015	0.011	0.006	0.001	-0.002	-0.006	-0.043
	(0.090)								
Top 20 College	-0.833***	0.077	0.062	0.043	0.021	0.001	-0.012	-0.026	-0.165
	(0.070)								
Offense lineman	-0.089	0.007	0.006	0.005	0.003	0.001	-0.001	-0.003	-0.019
	(0.211)								
Fullback	-0.384	0.036	0.029	0.020	0.010	0.000	-0.006	-0.013	-0.076
	(0.250)								
Defense lineman	-0.405**	0.036	0.030	0.022	0.011	0.001	-0.006	-0.013	-0.082
	(0.181)								
Wide receiver	-1.433***	0.178	0.106	0.050	0.007	-0.020	-0.036	-0.051	-0.235
	(0.177)								
Tight end	-0.311*	0.028	0.023	0.017	0.009	0.001	-0.004	-0.010	-0.063
	(0.166)								
Quarterback	-1.467***	0.198	0.107	0.043	-0.002	-0.027	-0.040	-0.054	-0.224
	(0.238)								
Linebacker	-0.798***	0.081	0.061	0.039	0.016	-0.003	-0.016	-0.028	-0.150
	(0.169)								
Defensive back	-1.524***	0.180	0.112	0.057	0.013	-0.016	-0.034	-0.052	-0.260
	(0.167)								
Running back	-1.157***	0.136	0.088	0.046	0.010	-0.014	-0.028	-0.042	-0.197
	(0.171)								
Team fixed effects	*Yes*								
Year fixed effects	*Yes*								
Number of observations	4262								

Robust standard errors in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Cut-point estimates in ordered logit model were omitted. [a] Including players with missing draft rounds assumed to be undrafted and coded as drafted in round 8. [b] All outcomes represent respective draft rounds in the NFL draft.

Table 5: Ordered logit model to trace draft preferences (part 2)[a].

	Model 2	**Marginal effects for outcomes** 1 - 7[b]						
		1	2	3	4	5	6	7
Size	-0.041*** (0.009)	0.005	0.004	0.001	-0.001	-0.002	-0.003	-0.004
Weight	-0.003 (0.002)	0.000	0.000	0.000	-0.000	-0.000	-0.000	-0.000
Career length	-0.210*** (0.034)	0.027	0.018	0.008	-0.004	-0.011	-0.016	-0.021
Career length sq.	0.009*** (0.002)	-0.001	-0.001	-0.000	0.000	0.000	0.001	0.001
Top 10 college	0.006 (0.129)	-0.001	-0.001	-0.000	0.000	0.000	0.000	0.001
Top 20 college	-0.730*** (0.096)	0.100	0.061	0.020	-0.018	-0.040	-0.054	-0.068
Offense lineman	0.093 (0.273)	-0.011	-0.008	-0.004	0.001	0.005	0.007	0.010
Fullback	0.432 (0.288)	-0.047	-0.037	-0.021	0.001	0.020	0.034	0.051
Defense lineman	-0.156 (0.260)	0.020	0.013	0.005	-0.003	-0.009	-0.012	-0.015
Wide receiver	-0.857*** (0.202)	0.135	0.065	0.008	-0.033	-0.050	-0.058	-0.067
Tight end	0.095 (0.244)	-0.012	-0.008	-0.004	0.001	0.005	0.007	0.010
Quarterback	-0.622** (0.280)	0.095	0.049	0.009	-0.023	-0.037	-0.043	-0.050
Linebacker	-0.382 (0.240)	0.053	0.032	0.010	-0.011	-0.022	-0.028	-0.034
Defensive back	-0.899*** (0.226)	0.137	0.070	0.013	-0.031	-0.052	-0.062	-0.074
Running back	-1.026*** (0.227)	0.172	0.072	0.001	-0.044	-0.060	-0.066	-0.074
Team fixed effects	*Yes*							
Year fixed effects	*Yes*							
Number of observations	2836							

Robust standard errors in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Cut-point estimates in ordered logit model were omitted. [a] Without any undrafted players. [b] All outcomes represent respective draft rounds in the NFL draft.

Table 6: Efficiency analysis results.[a]

	(I)	(II)	(III)	(IV)	(V)
Log of yards on offense	2.041***	1.988***	1.915***	1.875***	2.084***
	(0.227)	(0.227)	(0.237)	(0.238)	(0.246)
Log of yards on defense	-2.272***	-2.216***	-2.162***	-2.121***	-1.918***
	(0.204)	(0.204)	(0.211)	(0.212)	(0.309)
Interception ratio	0.339***	0.336***	0.353***	0.350***	0.317***
	(0.051)	(0.051)	(0.051)	(0.052)	(0.042)
INEFFICIENCY					
SET 1					
Draft Round		0.070		0.075	0.081
		(0.062)		(0.063)	(0.058)
Draft pick		-0.002		-0.002	-0.002
		(0.002)		(0.002)	(0.002)
Career		-0.153		-0.242	-0.199
		(0.292)		(0.304)	(0.276)
$Career^2$		0.019		0.028	0.023
		(0.029)		(0.031)	(0.028)
Out of top 10 college		0.031**		0.024	0.019
		(0.015)		(0.016)	(0.014)
Out of top 20 college		-0.004		0.001	0.002
		(0.012)		(0.012)	(0.011)
SET 2					
Average RB weight			0.004*	0.005**	0.004*
			(0.002)	(0.002)	(0.002)
Average RB size			-0.003	-0.006	-0.005
			(0.007)	(0.007)	(0.007)
Average WR weight			-0.006*	-0.004	-0.005
			(0.003)	(0.003)	(0.003)
Average WR size			0.018*	0.013	0.016*
			(0.009)	(0.010)	(0.009)
Average OL weight			0.004	0.003	0.003
			(0.003)	(0.003)	(0.003)
Average OL size			0.014	0.018	0.017
			(0.015)	(0.016)	(0.014)
Average DL weight			0.002	0.001	0.001
			(0.003)	(0.003)	(0.002)
Average DL size			0.013	0.011	0.010
			(0.012)	(0.012)	(0.011)
Average LB weight			-0.001	-0.001	-0.001
			(0.005)	(0.005)	(0.004)
Average LB size			-0.019	-0.019	-0.018
			(0.013)	(0.013)	(0.012)
Average QB weight			-0.004*	-0.003	-0.002
			(0.002)	(0.002)	(0.002)
Average QB size			0.003	-0.001	-0.001
			(0.007)	(0.008)	(0.007)
Average DB weight			0.007	0.004	0.005
			(0.006)	(0.006)	(0.005)
Average DB size			-0.003	0.002	-0.001
			(0.014)	(0.015)	(0.013)
Year after coaching change					0.102*
					(0.055)
Constant	0.772***	0.855	-5.217	-3.784	-3.102
	(0.047)	(0.739)	(4.883)	(5.275)	(4.763)
Number of Observations	250	250	250	250	250

Standard Errors in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. [a] The dependent variable is a logarithmic transformation of the win ratio. All size variables are measured in centimeter. Weight is measured in US pounds. Stochastic frontier parameters were omitted.