

What Does It Take? Deciphering Performance Indicators of NFL Running Backs through the Examination of Collegiate Performance and NFL Combine Results

Antonio DeMoss, Kevin Hindenach, Zach Vos, & Colin Weaver

University of San Francisco

February 2015

ABSTRACT

This research uses a linear regression model to investigate the relationship between prospective NFL running backs' NCAA FBS football statistics, NFL Combine measureables, and realized performance in the NFL as evaluated by Pro Football Focus. We observe 435 player-seasons from 2007-2014. The model suggests that collegiate conference affiliation, collegiate touchdowns, and NFL team passing strength are positively associated with NFL running back performance at statistically significant levels. Conference affiliation has the most substantial effect. NFL talent evaluators must appreciate that context is king when evaluating potential, and that pure stats are only a small piece of the puzzle.

INTRODUCTION

Talent is not distributed evenly among the elite athletes of the National Football League. There are countless sources of the differing performances of professional football players: work ethic, team chemistry, coaching, socioeconomic background, genetics, etc. All these factors may interact with one another to some degree. Amid this cornucopia of information—potentially worth millions of dollars to NFL decision makers—perhaps there is a signal of future performance. This study builds on previous findings and uses existing data to search for clues of future performance at the running back position.

Football is a complex sport, and a running back does not compete in a vacuum. The strategies employed and plays called by a running back's coaching staff surely impact his performance and are not equal across the league. The strength and weaknesses of his teammates also must play a role, either directly through the quality of the offensive line that blocks for him, or indirectly through the quarterback and wide receivers whose actions may focus the defense's attention elsewhere (or not). The quality of the opposition may also conflate a running back's perceived performance. A great defense can make a good running back look bad, just as a terrible defense can make a bad running back look good. Game conditions play a role. Running indoors on FieldTurf is not exactly the same as running outdoors on natural grass (or mud, or snow). In the real world of countless variables and effects, in a sport in which 22 players are on the field at once, quantifying individual performance is difficult.

This research investigates the above complexities to the fullest possible extent. While the true counterfactual remains ever elusive, the advent of player statistics from Pro Football Focus (PFF) equips analysts with tools to tease out a player's true, isolated performance. Up until now researches have yet to regress any PFF data on potential independent variables. This study

changes that. Our regression analysis indicates that context matters a great deal, both in college and in the NFL. Running backs from a National Collegiate Athletic Association “power conference” within Division-IA (Football Bowl Subdivision) perform better in the NFL than others. Additionally, the number of collegiate touchdowns a running back scores relates positively to later NFL performance, as does the quality of his NFL team’s passing offense.

LITERATURE REVIEW

One of the most important factors in the success of an NFL franchise is deciding which players to select in the NFL draft. McGee and Burkett suggest that “Performance measures could...be used as a means of predicting future athletic success in football” (2003, p. 6), but we found no reliable performance measure(s) utilized consistently in the literature. There are so many variables in the evaluative process that influence a team’s draft choice that it is impossible to know if a player will succeed or not until after a team has already invested a considerable amount of time and resources. We wished to delve further into those relationships using newly available variables from PFF. In order to develop a model that could aid NFL personnel decision makers in their process, we narrowed our scope to identify potential realistic, reliable, and usable relationships between past and future performance. We focused on the two categories of statistics available to team executives at the time of the NFL draft: a player’s collegiate performance and his NFL combine results.

We identified collegiate performance as an obvious critical variable for NFL draft prospects, and wanted to examine it in relation to a player’s future professional performance. Lyons et al. mentioned that “A common axiom in personnel selection is that past performance should be among the best predictors of future performance” (2011, p.158). This axiom makes sense in general. However, we needed to make sure the idea fit into the context of collegiate and

professional football. We examined similar topics in other sports and found evidence that in the NBA, collegiate performance may predict NBA success (Coates & Oguntimein, 2008, p.6). This is a good indication that collegiate football performance has a legitimate connection to professional football performance.

However, NCAA football players' past performance is extremely relative. Therefore, it was important to distinguish specifically which players would make the most sense to study. According to Hendricks et al., the collegiate performance of FBS players has much more certainty as a predictor of future NFL performance than that of non-FBS players (2003, p. 14). For this reason we decided only to focus on FBS players. Additionally, there is a gap between the strongest and weakest teams within the FBS segment. The schools that produce the most NFL players belong to only four of the 11 FBS conferences, and players from these conferences are often chosen higher in the draft due to the stronger competition they consistently face (Noel, 2012, p.3). It may be significant to look more closely at players from these conferences. If they perform well against the upper echelon of collegiate competition, they would theoretically be better prepared to perform at the professional level.

Spieler et al. found some strong correlations between a player's past and future performance, focusing on a player's mental makeup and background as a student (2007). Their study was only conducted on a small sample of players from a single non-FBS conference, making it difficult to determine if these findings were applicable for the FBS players we are interested in. Regardless, the next step for NFL prospects immediately following their collegiate career is the NFL combine.

The NFL hosts the combine specifically to help personnel professionals evaluate players through a series of standardized drills in an unbiased environment. A player's performance at the

combine has been shown to affect draft stock, especially among running backs, wide receivers and defensive backs (McGee & Burkett, 2003). However, we wanted to know if the NFL Combine was helpful in predicting actual NFL performance. According to Robbins, who studied the difference between normalized and raw data from the NFL combine in relation to future performance, there was a very weak correlation between the two on the whole—regardless of whether the data was normalized (2010). Kuzmitz & Adams offered an explanation as to why this would be the case, distinguishing skill from ability: “a skill narrowly focuses on a particular task, whereas ability more broadly relates to a multiple set of tasks, or competency” (2008, p.1726). The combine is only a test of individual skills (not ability) in a non-competitive setting.

Despite the tenuous validity of combine performance as a predictor of NFL success, we did find one significant correlation. The 40-yard dash sprint test has been a good predictor of running back performance in the NFL. In fact, it is the only combination of skill and position at the combine that has had a statistically strong correlation (Kuzmitz & Adams, 2008). For reasons like this, we decided to focus solely on running backs.

Another common theme in our research of collegiate and combine performance was the effect of intelligence testing and mental ability. There is currently some personality and aptitude testing at the combine. The usefulness of these tests is widely questioned and there have been suggestions for the use of more contemporary human resource practices to better determine fit and propensity for the next level (Kuzmitz & Adams, 2008, p.1727). General mental ability has been a proven forecaster of workplace performance when evaluated properly, but it is important to note that success in a football context is largely a function of physical ability as well. In football, mental ability is merely a piece of the puzzle and not reliable as a single forecaster of performance, unlike other professions (Lyons et al., 2009, p.5).

Since we wanted to help improve the draft process, it was important to check out previous research on draft selection and how exactly it relates to the evaluative process. It is extremely detrimental to a team when they select a player high in the draft and that player does not pan out, so one of the problems we aim to address is the fact that most NFL “teams overvalue the right to choose,” especially in the first round of the NFL draft (Massey & Thaler, 2005, p.5). Quinn et al. had similar findings, and added that the overvaluation of high draft choices is not only due to football evaluative error, but also is “...more consistent with the enduring decision-making biases suggested by psychology than by the economics of rational expectations and efficient markets” (2007, p.4). This speaks to the importance of being as objective as possible in determining the value of an NFL prospect, and not falling for such traps as the “winner’s curse”, the idea that with a high number of agents bidding on a good of uncertain value, the winner likely overpays (Massey & Thaler, 2005, p.8). This literature helped us decide that the draft itself was not something that we wanted to focus our research on, but reaffirmed our desire to come up with a way to help teams with the evaluative process. NFL teams face substantial uncertainty that can really harm their chances for success.

The last topic that we gathered background information about was how NFL performance and success are measured. One of our initial thoughts was to use a player’s documented accolades as a measure of performance, but as Berri & Simmons have noted, measures used to determine common productivity across all players (games started, pro bowl selections, etc.) are not adequate indicators of performance (2009, p.39). This told us that we needed to shift to a more statistical approach, and underscored the need to study one specific position so that we would have more data to directly compare across players. There are more football statistics and metrics available today than ever before, and we found a plethora of useful information from

PFF. Unfortunately, being founded in 2007, their data does not exist before that year. Yet after finding limited academic studies on how to evaluate NFL running back performance specifically, we felt a need to introduce a new statistical measure to the literature.

Researchers have spent considerable time assessing the evaluative process of collegiate athletes. We, however, identified an opportunity to combine a study that brings something from every part of the process together. It is our goal to identify indicators that will review the running back position like never before and provide an evaluative tool for professional personnel decision makers. We critically analyze the entire process by relating a running back's collegiate and combine performance to his subsequent performance in the NFL.

METHODOLOGY

Research Question

Among prospective NFL running backs who attended the NFL Combine from 2007-2014, what is the relationship between pre-draft performance—in college and at the NFL combine—and post-draft, NFL performance as measured by Pro Football Focus?

Research Design

Due to football's enormous popularity in the United States, a wealth of various statistics exists for almost all prospective NFL players. This research attempts to identify and integrate this secondary data into a comprehensive (if not all-encompassing) regression analysis. The analysis looks for relationships between a running back's contextualized, pre-NFL, on-field production and measurables and his NFL on-field performance from 2007-2014.

Population and Sample

This research examines the population of prospective NFL running backs, identified as such by their attendance of the NFL Combine in their draft year. Our sample includes all 220

NFL Combine attendees at the running back position from 2007-2014.

Data

College Football Reference (<http://www.sports-reference.com/cfb/>)

CFR provides objective, conventional, collegiate football individual player statistics. We observed a running back's nominal collegiate conference affiliation, his discrete career *collegiate-touches* (rushes plus receptions), and his discrete career *collegiate-touchdowns* (rushing touchdowns plus receiving touchdowns) for running back prospects entering the NFL from 2007-2014. Among running backs whose collegiate conference affiliation changed, we used their final conference affiliation before entering the NFL.

We recoded conference affiliations as a dummy variable, *power-conference*. The *power-conference* variable identifies running backs who either: played in a conference whose champion automatically qualified to one of the five Bowl Championship Series bowl games in FBS football; or played for Notre Dame, an independent FBS football team that, due to the "Notre Dame Rule," could also automatically qualify for one of the aforementioned bowl games upon meeting certain NCAA standards. From 2007-2012, the FBS "power" conferences with automatic qualification berths were the Atlantic Coast Conference, the Big 12 Conference, the Big East Conference, the Big Ten Conference, the Pacific-12 Conference (formerly the Pacific-10 Conference), and the Southeastern Conference. In 2013 the American Athletic Conference replaced the Big East Conference. We coded *power-conference* equal to one if a running back played in a power conference (as defined above) or for Notre Dame, and zero otherwise.

NFL Combine Results (<http://nflcombinerresults.com/>)

The NFL Combine is an annual event for invited prospects to showcase their objective attributes and skills in a highly controlled and regulated environment. We observed a prospective

running back's continuous *height* in inches, continuous *weight* in pounds, and continuous *forty-yard-dash* time in seconds, for all combine attendees from 2007-2014.

Pro Football Focus (<http://www.profootballfocus.com/>)

PFF produces subjective individual and team performance grades for all NFL players and teams, respectively. Multiple PFF experts examine every player on every play of every game to produce their grades, which NFL teams and evaluators themselves use. PFF's methods attempt to account for a play's context: a running back with bad blocking who rushes for the same number of yards as a running back with good blocking, all else equal, performs better and consequently receives a higher PFF grade for that play. An average performance receives a grade of zero. (For more about PFF grades, please see <https://www.profootballfocus.com/about/grading/>.)

We observed a professional running back's continuous *PFF Rushing and Receiving Grade*, continuous *nfl-team-pass* (grade of his entire offense's performance on passing plays over one season), and discrete *nfl-career-year*, for the NFL seasons 2007-2014.

Hypothesis and Model

In accordance with previous research and available data, discussed above, we proposed a linear relationship between a prospective running back's pre-draft and post-draft performance. We theorized that a strong passing attack affects the defense's strategy, and consequently control with *nfl-team-pass*. We also controlled for differing years of NFL experience, using fixed-effects for the observed *nfl-career-years*. We hypothesized that the following model describes the annual PFF performance grades of an NFL running back in career year i .

$$\mathbf{H}_1: \quad PFF \text{ Rushing and Receiving Grade}_i = x_0 + x_1 \text{height} + x_2 \text{weight} + x_3 \text{forty-yard-dash} + x_4 \text{power-conference} + x_5 \text{collegiate-touches} + x_6 \text{collegiate-touchdowns} + x_7 \text{nfl-team-pass} + x_{8,i} \text{nfl-career-year}_i$$

H₀: There is no relationship between a running back’s pre-draft observables and his PFF grade.

RESULTS

Among the model’s seven independent variables, three—*power-conference*, *collegiate-touchdowns*, and *nfl-team-pass*—have a statistically significant, positive association with *PFF Rushing and Receiving Grade* for the 435 player-seasons in our sample. These effects hold amongst the other controlling variables and across time through a running back’s NFL career. Consequently we rejected the null hypothesis that these variables have no relationship to an NFL running back’s PFF rushing and receiving performance grade. Four of the seven fixed effects of *nfl-career-year* have statistically significant, negative associations with NFL performance. The model found the remaining variables to be insignificant. More detailed findings are below.

Table 1

PFF Rushing and Receiving Grade

Observations	435
R-Squared	0.1094
Independent Variable	Coefficient
<i>height</i>	0.225
<i>weight</i>	-0.025
<i>forty-yard-dash</i>	2.08
<i>power-conference</i>	2.047***
<i>collegiate-touches</i>	-0.003
<i>collegiate-touchdowns</i>	0.089***
<i>nfl-team-pass</i>	0.03***
<i>nfl-career-year-1</i>	-6.403**
<i>nfl-career-year-2</i>	-5.831**
<i>nfl-career-year-3</i>	-5.626**
<i>nfl-career-year-4</i>	-4.610*
<i>nfl-career-year-5</i>	-5.814**
<i>nfl-career-year-6</i>	-4.642
<i>nfl-career-year-7</i>	-1.602
*P-Value<0.10 **P-Value<0.05 ***P-Value<0.01	

power-conference

Playing in a *power-conference* within FBS football was associated with a two-unit

increase in a running back's PFF rushing and receiving grade over the course of a single NFL season. With a strong statistical ($p=0.004$) and substantive (coefficient = 2.047) effect, this dummy variable was the strongest relationship in the model. For more detail about what constitutes a *power-conference*, please examine the above methodology.

collegiate-touchdowns

An additional touchdown scored by a running back during his collegiate career was associated with a roughly one-tenth-unit increase in his PFF rushing and receiving grade. This effect was strongly statistically significant ($p=0.008$), with a coefficient of 0.089. All else equal, the model indicated that among NFL Combine attendees, an FBS, non-*power-conference* running back who scores 22 more touchdowns than an FBS *power-conference* running back over their respective collegiate careers will perform equally in the NFL.

nfl-team-pass

The model found an NFL running back's team's passing attack to have a very strong statistical, yet very weak substantive, effect on the running back's PFF grade. A one-unit increase in his team's season passing performance was associated with only a three-hundredths-unit increase in his own season rushing and receiving performance. Nonetheless, this relationship was the strongest statistical effect in the model ($p<0.001$).

nfl-career-year

Rookie seasons were the worst among all player-seasons, with the sampled running backs demonstrating statistically significant improvement through the early years of their careers. Independent of the other included variables, second-year running backs performed roughly one-half-unit better than first-year backs, while third-year backs performed roughly one-fifth-unit better than second-year backs. These findings were statistically sound ($p<0.05$). Beyond the early

years, the sample size dwindled and the effects muddled from significant to insignificant, with curious coefficients, from year to year.

Insignificant Variables

The model suggested that a running back's *height* ($p=0.195$), *weight* ($p=0.369$), and *forty-yard-dash* time ($p=0.400$), as recorded at the NFL combine, and his career *collegiate-touches* ($p=0.107$), were all unrelated to his NFL performance. Due to the statistical insignificance observed, we failed to reject the null hypothesis that these variables have no relationship to an NFL running back's *PFF Rushing and Receiving Grade*.

DISCUSSION

The model adds to the conversation about how to better predict NFL performance based on various pre-draft measures. Yet even with the inclusion of previously unstudied metrics and variables, it is still far from a perfect science. There are some significant correlations and relationships between variables examined in this study, but our model does not replace the need for non-statistical human evaluation of running back prospects.

Our results support the research of Hendricks et al. which suggests that it is much riskier to trust the collegiate performance of running back prospects from non-FBS conferences (2003). It also helps explain why the ten schools that produce the most running backs selected in the first round of the NFL draft all belong to *power-conferences* (Noel, 2012). Although this does not significantly narrow the pool of available running backs who declare for the NFL draft, as the majority come from FBS conferences, it speaks to the value of consistently playing against stronger collegiate competition prior to playing in the NFL—the pinnacle of football competition. There have occasionally been successful running backs from non-FBS conferences. Our data simply reinforces the notion that non-FBS collegiate performance is not on the same

plane as FBS competition, at least at the running back position. The distinction between FBS and non-FBS running backs is even more evident when considering touchdowns scored in a player's collegiate career. All else equal, a non-FBS running back must score some 22 more touchdowns than an FBS running back for this model to value their prospective careers equally. Most collegiate running backs score only 20 or 30 touchdowns in their entire career, demonstrating the profound importance of playing FBS football. An average FBS running back will likely perform better than a great non-FBS running back.

The NFL combine is generally perceived as having some sort of predictive power in determining NFL success, but to what degree has been up for debate. McGee & Burkett theorized that the NFL combine measures could accurately predict the draft status of running backs, wide receivers, and defensive backs, but not with other positions (2003). Kuzmitz & Adams noted that most of the NFL combine statistics were not good indicators of NFL performance, but that the *forty-yard-dash* for running backs did have some correlation (2008). Our results do not support either of those findings. This model posits that there is no significant correlation between any of the combine measures and NFL performance. It may take a lot more than pure athleticism to perform in the NFL.

According to our study, the nature of a player's collegiate conference was the strongest indicator of professional performance, followed closely by how many touchdowns he scored in college and the quality of his NFL offense's passing attack. We suggest that these findings, while placing less stock in combine performance, are useful to consider when evaluating potential running backs. In their current infancy, these ideas should *not* be the sole determinants of an NFL decision-maker's player evaluations.

LIMITATIONS

While these results have contributed to the field, their conclusions are inherently limited. This research attempted to isolate the indicators of running back performance, but football is in fact a team sport and the actions of others also play into the equation. This may have biased our sample. For example, the performance of an offensive line is imperative and will either help or hinder a running back's opportunity to perform at a high level. PFF grades attempt to control for such things, but they may not do so perfectly. Similarly, the ability of the offensive play caller may also distort a running back's true performance.

With only eight years of data available, it was also difficult to track successful players for the duration of a career. Injuries and player movement resulting in different situations (play calling, scheme, and offensive line) could greatly influence a running back's chances of finding success in the NFL. While many independent variables were included in the study, spurious third causes may misconstrue the results.

With that in mind, the limitations of this project have created new avenues in which future research could capitalize on. Isolating additional variables, such as injuries, may lead to better findings. Interaction effects and non-linear relationships may also be present. Such findings could lead to more succinct results in predicting successful running backs. Rather than focusing on performance, future research may use similar data to investigate whether a collegiate running back will play in the NFL at all. Character concerns also play a huge role in evaluating young professional athletes. How running backs perform before and after injuries is also a good topic for future research.

CONCLUSION

The new millennium has enriched the football world with numbers. Sports fans and analysts alike have numbers evaluating how a running back performed given the blocks made (or

missed) by his offensive line, the tackles made (or missed) by his defensive opposition, etc.

Along with the conventional combine and collegiate measures dating back decades, these numbers present new avenues that talent evaluators can utilize when making tough decisions.

From looking at nearly a decade of statistical information, this study discovers that playing in a *power-conference* is the most indicative predictor for professional running back success. While intended for NFL talent evaluators, this knowledge may also be useful for high school football players entering college: it really matters which conference a running back plays in. Other statistically significant variables are the number of *collegiate-touchdowns* scored and the *nfl-team-pass* grade of the player's professional squad. The model represents merely a piece of the puzzle when it comes to evaluating running back prospects. The study's limitations section detailed the fact that injury, offensive line performance, offense scheme, play calling ability, and other factors may affect professional success as well, complicating these findings. While our model identifies some significant variable associations, there are also many insignificant variables. This analysis is one step towards understanding signals of NFL running back performance, but this understanding is far from complete.

References

- Berri, D. J., & Simmons, R. (2011). Catching a draft: On the process of selecting quarterbacks in the national football league amateur draft. *Journal of Productivity Analysis*, 35(1), 37-49.
- Coates, D., & Oguntimein, B. (2010). The length and success of NBA careers: Does college production predict professional outcomes? *International Journal of Sport Finance*, 5(1), 4-26.
- Hendricks, W., DeBrock, L., & Koenker, R. (2003). Uncertainty, hiring, and subsequent performance: The NFL draft. *Journal of Labor Economics*, 21(4), 857-886.
- Kuzmits, F.E., & Adams, A.J. (2008). The NFL combine: Does it predict performance in the National Football League? *Journal of Strength and Conditioning*, 22(6), 1721-1727.
- Lyons, B. D., Hoffman, B. J., & Michel, J. W. (2009). Not much more than g? An examination of the impact of intelligence on NFL performance. *Human Performance*, 22, 225-245.
- Lyons, B. D., Hoffman, B. J., Michel, J. W., & Williams, K. J. (2011). On the predictive efficiency of past performance and physical ability: The case of the National Football League. *Human Performance*, 24, 158-172.
- Massey, C., and Thaler, R.H. (2005). The loser's curse: overconfidence vs. market efficiency in the national football league draft. *NBER Working Paper*, (W11270).
- McGee, K.J. & Burkett, L.N. (2003). The national football league combine: a reliable predictor of draft status? *Journal of Strength and Conditioning Research*, 17(1), 6-11.
- Noel, T. (2012). The chosen few: colleges in the NFL draft. *Coffin Corner*. 34(3), 22-23
- Quinn, K. G., Geier, M., & Berkovitz, A. (2007, June). Passing on success? Productivity outcomes for quarterbacks chosen in the 1999-2004 National Football League player entry drafts. In *IASE Conference Papers* (No. 0711). International Association of Sports

Economists.

Robbins, D.W. (2010). The National Football League (NFL) combine: Does normalized data better predict performance in the NFL draft? *Journal of Strength & Conditioning Research*, 24(11), 2888-2899.

Spieler, M., Czech, D. R., Joyner, A. B., Munkasy, B., Gentner, N., & Long, J. (2007). Predicting athletic success: Factors contributing to the success of NCAA division IAA collegiate football players. *Athletic Insight*, 9(2), 22-33.