

Performance Measurement & Matching: The Market for Football Coaches

Todd Brown*
Stephen F. Austin State University

Kathleen A. Farrell
University of Nebraska–Lincoln

Thomas Zorn
University of Nebraska–Lincoln

The matching hypothesis asserts that it is the matching of the employee and the firm rather than the qualifications of the employee alone that matter. Using a unique and large data set of college football coaches, we perform two different tests of the matching hypothesis. The relative accuracy with which performance in college football is observed allows for a direct test of matching. We find that matching is a significant factor in team performance. A good match accounts, on average, for approximately a 5 percent improvement in performance. We also find that the hazard rate is increasing for the first five years and then subsequently decreasing over a coach's tenure. Our evidence is consistent with a four- or five-year contracting period for college football coaches.

Introduction

College football coaches occupy a unique position in college athletics and athletics in general. College football is the most prominent college sport. Unlike professional football, a successful coach over time is identified more strongly with the team than the athletes who leave after a few years. Their influence in some cases extends beyond athletics. Tom Osborne's successful career as Nebraska's coach, for example, led to a seat in Congress. College football coaches act as the CEOs of sizeable organizations. As such, they provide an example of a labor market where a high level manager's performance is measured with relative ease.

* We appreciate the helpful comments and suggestions of Richard DeFusco, Mostafa Mashayekhi, Manfred Peterson, Colin Ramsay, Warren Luckner, and seminar participants at the University of Nebraska–Lincoln.

Finding the best person for a particular position easily can be confused with merely finding the best person. The matching hypothesis, first proposed by Jovanovic (1979), suggests it is not the most able person that should be placed in a particular position, but the best person-to-firm match that should be hired. An individual's productivity is influenced by the way in which he or she fits with the firm. This fit can be value-increasing for the firm by allowing an average employee to achieve above average results; conversely, a bad match for the firm will result in below average results.

For the firm, fit involves factors such as the individual's ability to be productive in a specific work environment. Higher level positions are multidimensional, and a simple one-dimensional metric does not capture the factors that determine a fit. As important as it is for the firm to find an employee that fits the firm, it is just as important for an employee to find employment that is optimal given his or her alternatives. The labor literature stresses the importance of employers attracting and retaining qualified employees. This implies designing competitive compensation packages that include pay, incentives, retirement, and health benefits. There are also non-pecuniary aspects of a position that the firm may or may not control completely such as the location and prestige of the firm. If a firm cannot at least meet the next best alternative available to the employee, the employee is likely to quit.

In the college coaches market, fit involves a complex set of interrelated aspects of the coach's personality, the school, alumni, student body, the athletes, the community, the local press, and the school's football tradition. While turnovers either initiated by the school or the coach are common in college football, some tenures are lengthy; Penn State coach Joe Paterno is only one example. Lengthy tenure is *ex post* evidence of a good match.

Information costs imply that the matching process in the case of higher level positions requires search and post-hire experience on both sides of the market. When an organization hires or fills a particular position, there is no guarantee, even after an extensive search process, that they have found an individual that fits the organization. Conversely, an employee will not be fully aware of the characteristics of a position prior to on-the-job experience.

Testing the matching hypothesis is difficult in traditional labor markets because the theory requires knowledge of unobservable characteristics of the worker, the firm, and the match (Garen, 1999). We use the approach of a number of studies that attempt to overcome at least some of these difficulties by analyzing an example from the sports market—namely, the market for college football coaches. In this market it is easier to observe the match effect. We also are able to make use of a large sample. While this is not the first study that analyzes sports teams to test the matching hypothesis, the football coaches market is particularly suited for examining the matching hypothesis as a determinant of labor market mobility. We find evidence of a coach-team match effect on team performance for a large sample of college foot-

ball coaches from 1968 to 2003. We also find that the hazard rate increases early in a coach's tenure and subsequently decreases. Our results suggest, on average, that a good match results in a five percent improvement in a team's winning percentage.

The Matching Hypothesis

A good match between an individual and a firm can take many forms. The firm obviously needs an employee with the proper skills to perform the job. Moreover, it needs someone to fit the corporate culture. Conversely, an employee's preferences include factors such as compensation (pay and benefits), as well as job satisfaction. When both sides of this employment equation are satisfied, it can be expected that the tenure and productivity of the employee will be elevated, leading to greater rewards for both worker and organization.

Jovanovic (1979) theorizes that the probability of employee turnover increases at first as information is revealed about the quality of the match between the employee and the firm and then decreases as the likelihood of a good match increases. Another implication of the matching hypothesis is that if an employee is a match for the firm, his or her productivity should increase in proportion to the strength of the match. Those who are well suited for the job will receive increased compensation and will be more likely to stay. Conversely, a poorly matched employee of a firm will fail to reach his or her potential level of productivity, will fail to receive wage increases, and likely will be terminated or quit.

Match theory has been used to explain empirical results relating to the positive relation between wages and tenure and the negative association of turnover and tenure. [See Garen (1999) for a review of the matching hypothesis empirical literature.] Hersch and Reagan (1990) find that a match does have a positive effect on pay. Their study, however, relies on data gathered from 18 firms in Oregon that have few observed firm changes for employees. That study also fails to control for uniform employment characteristics of the sampled employees. To overcome this obstacle, more recent literature has used data from sports teams as a proxy for firms. Sports provide the opportunity to test the matching hypothesis using a relatively homogeneous sample of firms and employees.

Chapman and Southwick (1991) and Ohtake and Ohkusa (1994) use professional baseball data (American and Japanese, respectively) to test the impact of performance on manager-team matches. They attempt to isolate the effect of the match by controlling for player quality with statistics for hitting, fielding, and pitching. Both studies test for a match effect by using an interaction term between the team and the coach. While Chapman and Southwick (1991) find support for a match effect, Ohtake and Ohkusa (1994) do not. Prisinzano (2000) extend the data set of Chapman and Southwick (1991) and analyze the team's expected wins minus their

actual wins to proxy for the managerial effect on the team.¹ Borland and Lye (1996) extend Chapman and Southwick (1991) by analyzing the match effect in Australian Rules football. They use player awards to proxy for player quality and coaching experience to control for learning on the job in an effort to improve the isolation of the match variable.

The Labor Market of College Football Coaches

The college football coach market has some distinct advantages in testing the matching hypothesis. In professional sports the division of responsibilities between the coach, the general manager, and the owner can vary greatly. The legendary Chicago Bears coach George Halas ran the entire operation for much of his career; Al Davis, the general manager of the Oakland Raiders, was often credited as the genius behind the success of the team; and Jerry Jones, the charismatic owner of the Dallas Cowboys, was reputed to have a great deal of influence concerning on-the-field decisions. The players that are drafted by a professional team and other important personnel decisions are not entirely in the hands of the head coach or manager. In contrast, the responsibilities of a college football coach tend to be similar at most universities.

The lack of any front office personnel in collegiate football allows the coach to control ongoing operations of the team. Although a head coach in college football serves under a university president and the athletic director, the responsibility for the conduct, success, or failure rests with the head football coach. An athletic director or other university administrator rarely is involved except when a coaching change is made. Thus, a coach has identical responsibilities at different universities. The sample is therefore more homogeneous in college athletics than in professional athletics, a clear methodological advantage.

To test for the match effect between a coach and a team, it is necessary to isolate his or her performance, subject to the abilities of the team's players as well as the inherent capabilities of the program. Previous researchers analyzing professional sports have used player statistics as a proxy for player ability (e.g., Chapman and Southwick, 1991). Coaching decisions can impact the statistics of a player, reducing the significance of the coach-team match. This study uses the college draft to proxy for a player's talent level, which is only marginally influenced by coaching decisions. National Football League (NFL) teams are likely to choose a talented player

¹ This variable attempts to capture coaching efficiency. Expected wins is computed using a complex formula with such variables as the amount of runs scored and runs allowed. These numbers can be skewed due to the nature of a game, such as whether or not the game is close at the end. Managers tend to make different strategic decisions depending on the score that affect the number of runs scored/allowed.

over a statistically successful player when they decide who to select in the draft.² In college athletics, team talent is more under the control of the coach than is that in professional sports. Coach Bill Parcells, expressing his frustration at a lack of control over which players he had to coach, famously said “If you’re going to prepare the meal, you should be able to buy the groceries.”³ Many college coaches are known as outstanding recruiters rather than as on field generals⁴. We therefore also examine the match hypothesis ignoring player talent, because talent in college athletics may be a primary responsibility of the coach⁵.

Both Chapman and Southwick (1991) and Ohtake and Ohkusa (1994) concede that their model suffers from a simultaneous equation bias from the mid-season trading of professional baseball players around the league. This can lead to a contamination of the variable capturing the match effect. College football prohibits mid-season student transfers, and only rarely do players change schools even in the off season.

Data

Data on college football coaches and their respective teams were obtained through the College Football Data Warehouse at <http://www.cfbdatawarehouse.com/>. The entire coaching lineage from all current Division I-A football schools was available. To maintain homogeneity among coaches, we only include coaches that were members of Division I-A schools. Other divisions have to comply with different rules such as the number of scholarships that may be awarded. We gather data on schools beginning in 1925 because many schools began football programs after World War I. Our data set continues through 2003. Some data, however, are not available over the entire sample period. Because we include a player draft variable in Tables 3 and 4, we use data starting with the merger of the NFL and the American Football League in 1967 in order to avoid problems associated with two drafts. Table 1 below reports descriptive statistics for the sample period 1968-2003 used in the primary regression analysis shown in Tables 3 and 4.

To test for the existence of a coach-team match, it is necessary to look at output differences of a coach with different teams. We measure a coach’s output with the winning percentage of his team. The average winning percentage in the sample is 0.51 (Table 1), which includes wins over non-Division I teams. The highest and lowest overall winning percentage based on the sample between 1925 and 2003 is

² The Heisman Trophy winner, presumably the best college football player, is often not the top draft choice. Performance of players in college is influenced, to a considerable extent, by coaching.

³ http://www.usatoday.com/sports/football/nfl/cowboys/2003-01-02-parcells-cover1_x.htm

⁴ For example, Florida State coach Bobby Bowden’s success often is attributed in large part to his recruiting skills.

⁵ We are indebted to an anonymous referee for this point.

Tennessee (0.729) and Kent State (0.319), respectively. A team's performance also can be affected by the quality of the opponents a team faces during a season, by the quality of players on the team, and the inherent traits of the particular program.

Table 1—Means and Standard Deviations for Division I-A Schools 1968-2003

Variables	Mean	Std Dev	Min	Max
Winning Percentage	0.51	0.23	0.00	1.00
Tradition	0.52	0.15	0.12	0.90
Coaching Tenure	5.29	5.01	1.00	38.00
Coaching Experience	7.60	6.60	1.00	38.00
Strength of Schedule	0.56	0.18	0.05	0.95
Number Drafted	2.29	2.36	0.00	17.00

Notes: Tradition is a moving average of a team's previous ten years winning percentage, regardless of their coach. Coaching Tenure is the tenure of a coach at a given college. Coaching Experience is a coach's years of experience in collegiate football. Strength of schedule is based on James Howell's calculation (<http://www.jhowell.net/cf/cfindex.htm>). Number Drafted is the number of players selected for the NFL Draft for a given team in a given year

Because teams play against a subset of the Division 1-A teams, the schedule can have a significant effect on team performance. We specify a strength-of-schedule variable to capture opponent difficulty.⁶ As shown in Table 1, the strength of schedule varies from 0.05 to 0.95, with the average strength of schedule equal to 0.56. Notre Dame has the strongest strength of schedule, averaging 0.739 during the sample period from 1925-2003.

Because high school athletes have a choice of which college to attend, there is often an imbalance of talent among college teams.⁷ We use data on the NFL draft, obtained in the Football Encyclopedia, as a measure of player talent. While previous research focused on statistics as a proxy of player skill, the NFL draft order permits us to measure talent independent of coaching decisions. Players selected at the top of the draft typically receive far greater compensation, indicating the relative scarcity of their exceptional skills relative to players selected toward the bottom of the draft. We assume that talent level of the drafted players is nonlinear and is subject to far greater estimation error toward the later rounds of the draft. We, therefore, model the players drafted as an exponential decay function, namely $e^{-0.01(\text{Draft Choice})}$.⁸ To avoid problems

⁶ Strength of schedule is a common statistic in college football computer models used to determine the national champion. For our study, we used James Howell's calculation of SOS at <http://www.jhowell.net/cf/cfindex.htm>.

⁷ In professional sports, however, drafts and salary caps maintain a balance of talent.

⁸ This assumption is consistent with NFL teams using draft value charts. A draft value chart determines the numeric value of a draft pick to aid teams in trading draft picks. For example, one team's draft value chart assigns the top pick a value of 3000, the second pick 2600, third pick 2200, and fourth pick 1800. Then pick values fall less rapidly with the sixteenth pick worth 1000 points and the thirtieth pick worth 620. To see the decay of the function, the one hundredth pick is worth 100 points. See, for example, <http://www.nfl.com/draft/story/9341444>

associated with two drafts, we collect data starting with the merger of the NFL and the American Football League in 1967.⁹

The average number of players drafted in the sample is 2.29, as shown in Table 1. Joe Paterno has the most players drafted (215) over his last 35 years of tenure. His overall tenure in the sample is 38 years; however, we restricted the sample to 1968 and beyond for the draft variable. Larry Coker had 9.34 players drafted per year, on average, over his first three years of tenure at Miami. USC has the most players drafted overall of 230. Additional stats regarding the draft variable for coaches and teams are presented in Table 2.

Table 2—NFL Draft Data by Coach and by Team for Division I-A Schools 1968-2003

Panel A:			
Coach	Number Drafted	Total Tenure	Average per Year
Larry Coker	28	3	9.33
Jim Tressel	27	3	9.00
Jim Hilles	8	1	8.00
Tom Osborne	168	25	6.72
Phillip Fulmer	72	11	6.55
Ron Zook	13	2	6.50
Barry Switzer	102	16	6.38
Mark Richt	19	3	6.33
Joe Paterno	215	35	6.14
Dick Vermeil	12	2	6.00
Panel B:			
Team	Total Number of Players Drafted		
USC	230		
Nebraska	228		
Notre Dame	216		
Penn State	215		
Ohio State	209		
Miami, Florida	198		
Tennessee	194		
Michigan	188		
Oklahoma	177		
Florida	174		

We assume that the major difference in talent exists between star players and journeyman players. Among the journeyman players, talent differences tend to be minor. Many journeyman players are, to a considerable extent, interchangeable. Because recruiting is a major component of a college coach's job and talent is not

for a discussion of one NFL team's value chart obtained by the NFL. Our choice of the decay function closely resembles this team's value chart.

⁹ The draft when the two leagues were competing not only reflected the player's talent, but also the likelihood that the team would be able to sign the player.

independent of coaching, we also conduct the analysis without controlling for player talent.

Previous studies (Chapman and Southwick, 1991; Borland and Lye, 1996) have tried to test for the existence of team fixed effects. This controls for the possibility that any coach-team match can be confused with inherent differences in team abilities. Neither study finds a team-influence effect. Fort (2003), however, did find a significant team effect. He shows that there is a winning percent imbalance in college football that has persisted for at least 30 years. We therefore use a tradition variable based on a moving average of a team's previous ten years winning percentage.¹⁰ The tradition variable controls for coaches that switch between schools with sharply different traditions and allows for changes in a team's fixed effect in the model.

Not surprisingly, the average tradition variable shown in Table 1 is close to fifty percent, but the maximum is 0.924 (Oklahoma in 1958) which suggests that some colleges in the sample have strong winning traditions. Despite the high strength of schedule, Notre Dame also has the highest average tradition of 0.747 calculated over the sample period 1925-2003. Kent State has the lowest average tradition of 0.36, although Texas-El Paso had the lowest tradition variable in a given year (1985) of 0.122.

Finally, we control for the fixed effect of each coach's ability. This captures the difference in a coach's output, positive or negative, across different teams. We create a dummy variable for each coach and an experience variable to control for coaches with different levels of experience.

Testing Match Theory

A Direct Test of the Coach-Team Match

To determine the value of a coach-team match, we attempt to control for other factors that may impact the winning percentage of a football team. We specify the following regression equation:

$$P_{it} = \beta_1 \text{Schedule}_{it} + \beta_2 \text{Tradition}_{it} + \beta_3 \sum_{d_t \in D_t} e^{-0.01(\text{DRAFT}_{it})} \quad (1) \\ + \beta_4 \text{EXP}_{it} + \beta_5 \text{EXP}_{it}^2 + \sum_{c \in C} \delta_c \text{COACH}_{ci} + e_{it} ,$$

where the dependent variable, P , equals the team winning percentage and the independent variables control for the talent level of the players (DRAFT), quality of opponents (Schedule), and the varying fixed effects of the team (Tradition). To control for on-the-job learning and increases in skill over a coach's career, we include

¹⁰ Results are robust to alternative proxies for the tradition variable, including calculating average winning percentage over five and fifteen years.

EXP and EXP², the coach's years of experience and experience squared.¹¹ We include COACH_i as a measure of the fixed value that each coach brings to the profession. We specify COACH_{it} as a dummy variable ($c = 1, \dots, 556$), less one coach to avoid colinearity, to allow each coach to have a fixed effect. The marginal value of each of these variables is computed by using panel data from all 117 teams from Division 1-A for the years 1968-2003.

Matching theory implies that if there is a coach-team match, then the effect should be reflected in performance. Following Chapman and Southwick's (1991) approach, we directly test for a coach-team match by modifying the above equation to include an interaction variable as follows:

$$P_{it} = \beta_1 \text{Schedule}_{it} + \beta_2 \text{Tradition}_{it} + \beta_3 \sum_{d_t \in D_t} e^{-0.01(\text{DRAFT}_{it})} \quad (2)$$

$$+ \beta_4 \text{EXP}_{it} + \beta_5 \text{EXP}_{it}^2 + \sum_{c \in C} \delta_c \text{COACH}_{ci} + \sum_{m \in M} \gamma_m \text{MATCH}_{mci} + e_{it},$$

where MATCH_{mci} ($m = 1, \dots, 188$) is a dummy variable for each coach-team match. As described in Borland and Lye (1996), the match dummy variables are restricted to the set of repeat coaching spells. Finding statistical significance of the coefficients on the match dummy variable suggests that a coach's output varies across teams and, therefore, across matches. Thus, a direct test of the matching hypothesis is performed by comparing the explanatory power of Equations 1 and 2. That is, the coach is more productive at a given college than at another; symmetrically, for any given college a particular coach will be more productive than will another coach. Not all coaches change teams—Tom Osborne spent his entire career at the University of Nebraska, Barry Switzer did the same at Oklahoma, and Joe Paterno continues to coach at Penn State. Others, such as Lou Holtz with five coach changes, have been the head coach at a number of universities over their careers.

The results from estimating Equations 1 and 2 are reported in Table 3. The coefficients of the dummy variables have been suppressed for brevity. All non-dummy variables are significant at the 1 percent level in both equations, except years of coaching experience in the first regression. The coefficients of the strength of schedule and players drafted are consistent with the expectation that a weaker set of opponents and higher quality players results in a higher winning percentage. The tradition variable, however, changes signs when the match dummy variables are included in Equation (2). Because each match dummy variable is associated with only one team, it influences the effect of the tradition variable.

The players drafted variable assumes that very talented players have a significantly greater impact on performance. The case also can be made that the players drafted variable should not be included because it reflects one of the important func-

¹¹ A quadratic is used because learning frequently is modeled as increasing at a decreasing rate.

tions of a successful college coach, namely recruiting player talent. We, therefore, show the results of excluding the draft variable in columns four and five of Table 3. The exclusion of the player draft variable does not have a large impact on the remaining variables.¹²

Table 3—OLS Regressions on Winning Percentage for Division I-A Schools 1968-2003

Variable	Model 1	Model 2	Model 1	Model 2
Coach Dummies	Yes	Yes	Yes	Yes
Match Dummies	No	Yes	No	Yes
Strength Of Schedule	-0.315 (-12.45)	-0.325 (-11.01)	-0.260 (-9.80)	-0.318 (-10.32)
Tradition	0.263 (7.36)	-0.300 (-5.05)	0.398 (10.75)	0.243 (3.94)
Players Drafted	0.078 (18.98)	0.068 (16.32)		
Coaching Experience	0.00001 (0.01)	0.00700 (3.66)	0.00250 (1.49)	0.00891 (4.48)
Coaching Experience Squared	-0.00020 (-3.43)	-0.00026 (-4.10)	-0.00025 (-4.06)	-0.00030 (-4.50)
Sum of Squared Errors:	89.4644	76.5290	99.4843	83.2554
Degrees of Freedom:	3218	3030	3219	3031
Adjusted R ² :	0.4766	0.5245	0.4182	0.4829
Wald Statistic		512.15		590.83

Notes: Columns two and four report the results of OLS regressions on Equation (1), and columns three and five are OLS regressions on Equation (2). The Wald statistic is a test of the significance of inclusion of the match dummy variables between Equations (1) and (2). The coefficients of the dummy variables have been suppressed due to difficulty of interpretation and to conserve space. $Coach_{it}$ is a dummy variable for each coach ($c = 1, \dots, 556$), less one coach. $Match_{mci}$ ($m = 1, \dots, 188$) is a dummy variable for each coach-team match (less one match for each coach). Strength of schedule is based on James Howell's calculation at <http://www.jhowell.net/cf/cfindex.htm>. Tradition is a moving average of a team's previous ten years winning percentage, regardless of their coach. Players Drafted is a decaying function of the draft order in which players are selected for the NFL Draft. Coaching Experience is a coach's years of experience and is allowed to enter as a quadratic. t-statistics given in parentheses

We calculate a Wald statistic as shown in the final row of Table 3 to test the significance of the inclusion of the match dummy variables. The Wald test rejects the hypothesis that the match dummy variables do not explain the variation in winning percentage at the one percent level in all model specifications. As expected, the model fit is also enhanced by including match dummy variables and increases the R squared from 0.477 to 0.525 in columns two and three with similar improvements in fit across columns four and five.¹³

¹² We also expand the sample to include the data from 1925 through 2003 because the draft variable is the primary reason for restricting the data to post 1967. The results are qualitatively the same as those reported in Tables 3 and 4 and are available from the authors upon request.

¹³ The introduction of match dummies in our model increases the R squared in a comparable manner as that shown in Table 2 of Chapman and Southwick (1991) whose R squared increases from 0.765 to 0.823.

Because the average winning percentage for all teams during any given year must equal roughly fifty percent, the cross-sectional errors for each team will be correlated.¹⁴ This will cause the OLS estimates to be inefficient and may lead to upward-biased estimates of their variance. To correct for this bias, we re-estimate both Equations (1) and (2) using generalized least squares (GLS) for all model specifications in Table 3 and report the results in Table 4. The coefficients are similar to the OLS regressions, but the Wald test statistic is somewhat lower. The test statistic for the impact of the inclusion of the match dummy variables remains significant at the 1 percent level. Our evidence presented in Tables 3 and 4 is consistent with the matching hypothesis and robust to alternative specifications.

Table 4—GLS Regressions on Winning Percentage for Division I-A Schools 1968-2003

Variable	Model 1	Model 2	Model 1	Model 2
Coach Dummies	Yes	Yes	Yes	Yes
Match Dummies	No	Yes	No	Yes
Strength Of Schedule	-0.316 (-11.14)	-0.324 (-10.97)	-0.307 (-10.30)	-0.317 (-10.28)
Tradition	-0.137 (-2.73)	-0.294 (-4.97)	-0.062 (-1.19)	-0.238 (-3.87)
Players Drafted	0.074 (17.94)	0.068 (16.31)		
Coaching Experience	0.0023 (1.28)	0.0070 (3.66)	0.004 (2.18)	0.009 (4.48)
Coaching Experience Squared	-0.0002 (-2.98)	-0.0003 (-4.12)	-0.0002 (-3.42)	-0.0003 (-4.52)
Sum of Squared Errors	82.6652	76.5911	90.8997	83.3131
Degrees of Freedom:	3218	3030	3219	3031
R ² :	0.4646	0.4953	0.4100	0.4498
Wald Statistic		240.30		276.00

Notes: Columns two and four report the results of GLS regressions on Equation (1) and columns three and five are GLS regressions on Equation (2). The Wald statistic is a test of the significance of inclusion of the match dummy variables between Equations (1) and (2). The coefficients of the dummy variables have been suppressed due to difficulty of interpretation and to conserve space. $Coach_{it}$ is a dummy variable for each coach ($c = 1, \dots, 556$), less one coach. $Match_{mci}$ ($m = 1, \dots, 188$) is a dummy variable for each coach-team match (less one match for each coach). Strength of schedule is based on James Howell's calculation at <http://www.jhowell.net/cf/cfindex.htm>. Tradition is a moving average of a team's previous ten years winning percentage, regardless of their coach. Players Drafted is a decaying function of the draft order in which players are selected for the NFL Draft. Coaching Experience is a coach's years of experience and is allowed to enter as a quadratic. t-statistics given in parentheses

An Alternative Test of the Coach-Team Match

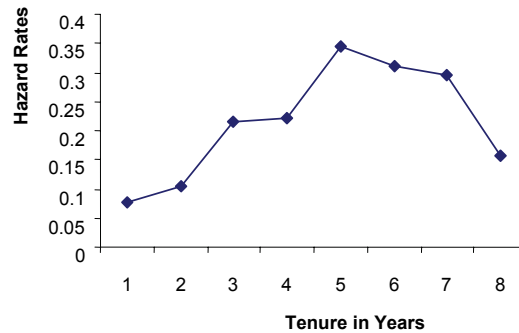
In his description of the matching model, Jovanovic (1979) proposes that employers and employees learn about match quality over time. Initially, the probability of job separation will increase early in a worker's tenure. As information

¹⁴ Due to schools playing teams from other divisions as well as some schools moving from Division I-A to lower divisions, the actual average winning percentage of the current Division I-A teams is 0.51, as shown in Table 1.

is revealed regarding match quality, the probability of job separation subsequently will decline.

We analyze the probability of separation by calculating the hazard rates for college football coaches over the sample period of 1968 through 2003 and document a pattern consistent with that predicted by Jovanovic (1979). Figure 1 displays college football coaches' hazard rates by tenure. An examination reveals that the hazard rate is increasing for the first five years and then subsequently decreasing over a coach's tenure. The hypothesis that the hazard rate is constant over tenure is rejected. Our evidence is consistent with a four or five year contracting period for college football coaches. Given the length of a college football player's eligibility, schools often allow a coach enough time to recruit his own players to decide on his coaching effectiveness. This result is also consistent with the findings of Allgood and Farrell (2003) in their study of the CEO labor market where the peak hazard rate for CEO tenure is also five years.

Figure 1—Hazard Rates by Tenure



Note: The hazard rates are based on the data for Division I-A schools from 1968-2003

While the peak does occur at the same time, the severity of the peak for college coaches is more than twice that of CEOs. Furthermore, both Chapman and Southwick (1991) and Borland and Lye (1996) find a peak of over 40 percent at the three year point in tenure with their data on baseball and Australian Rules football, respectively. This is consistent with the notion that performance is easier to measure in sports than it is in the corporate context, which accelerates the determination of a successful match.

To measure the effect of matching, we draw on the information provided by the hazard rate. The hazard indicates that if a coach's tenure endures beyond five years,

the coach-team combination is a good match. To further analyze this relation, we calculate the average winning percentage for all coaches in the sample from 1925 to 2003 who have completed tenures that last between one and five years. Eight hundred and twenty-six coaches have tenures between one and five years in the sample and have an average winning percentage of 0.44. We also calculate the average winning percentage for all coaches that have a tenure that lasts beyond five years (these include incomplete spells). We have 484 coaches in the sample with tenures greater than five years and an average winning percentage across all years equal to 0.56. The average winning percentage of 0.56 for coaches with tenures greater than five years is significantly greater than the average winning percentage of 0.44 for coaches with tenures of five years or less, at the 1 percent level.

Based on the preceding analysis, we define a dummy variable equal to one if a coach's tenure is greater than five years with a given college. The dummy variable takes the value of one for every year the coach-team match is in the sample.¹⁵ If a coach is with a team for five years or less and then departs, we define the dummy variable equal to zero. We have a right censoring issue because some coaches have been with a team less than five years, but have not left the team at the end of our sample period. To capture a sufficient number of good and bad matches in the sample, we use the entire sample period to include all Division I-A schools from 1925-2003. Given our data for the players drafted variable begin in 1968, we drop this variable from the analysis.

In Table 5, we show the regression of winning percentage on the good match dummy variable, strength of schedule, tradition, coaching experience, and coaching experience squared. The signs and significance levels of the control variables are consistent with expectations, except an insignificant relation between coaching experience and winning percentage. The good match dummy variable is positive and significant in all model specifications. The coefficients in Table 5 suggest that a good match between a coach and team increases the winning percentage of a team 5.5 percent or 4.5 percent, according to the OLS (column two) or the GLS (column three) specification, respectively.¹⁶

The good match variable is correlated with the tradition variable because all good matches are defined as those that last longer than five years, and tradition is a ten year moving average. Therefore, we exclude the tradition variable from the analysis in Table 5. The significance levels of the good match variables are slightly

¹⁵ Allgood and Farrell (2003) define a good match variable equal to one if a CEO's tenure is greater than or equal to five years. They only analyze performance of the firm over the first three years of the CEO's tenure, not performance over the CEO's entire career. Their unit of observation is a CEO-firm match.

¹⁶ As a robustness test, we restrict the data to the sample period 1968-2003. The good match dummy variable remains significant, but the coefficient is between 4 percent and 3.5 percent depending on the model specification.

higher in columns four and five relative to columns two and three of Table 5, although the inferences remain virtually unchanged. These results suggest that a good match results in a 5 percent improvement in a team's winning percentage.

Table 5—Regressions on Winning Percentage while Controlling for a Good Match for Division I-A Schools from 1925-2003

Variable	OLS Regression	GLS Regression	OLS Regression	GLS Regression
Coach Dummies	Included	Included	Included	Included
Good Match	0.055 (4.96)	0.045 (3.23)	0.061 (5.45)	0.045 (3.28)
Strength Of Schedule	-0.298 (-15.75)	-0.310 (-15.23)	-0.290 (-15.24)	-0.306 (-15.08)
Tradition	0.290 (9.67)	0.088 (2.49)		
Coaching Experience	0.00022 (0.18)	0.00161 (1.22)	0.00087 (0.69)	0.00187 (1.43)
Coaching Experience Squared	-0.00014 (-3.01)	-0.00019 (-4.01)	-0.00013 (-2.72)	-0.00019 (-3.93)
Sum of Squared Errors:	192.9596	183.5478	196.0038	183.7013
Degrees of Freedom:	5924	5924	5925	5925
R ² :	0.4591	0.3940	0.4505	0.3925

Notes: The second and fourth columns report the results of OLS regressions on Equation (3), and the third and fifth columns are GLS regressions on Equation (3). Good Match is a dummy variable where a coach's tenure is greater than five years. Strength of schedule is based on James Howell's calculation at <http://www.jhowell.net/cf/cfindex.htm>. Tradition is a moving average of a team's previous ten years winning percentage, regardless of their coach. Coaching Experience is a coach's years of experience and is allowed to enter as a quadratic. t-statistics given in parentheses.

We caution that it is difficult to quantify the coach-team match effect. For example, in our original model specifications (Tables 3 and 4), our coach dummy variables have information concerning the match with the deleted coach-team combination. Also, the match dummy variable represents the quality of the coach-team match relative to the omitted coach-team match for that individual. Also, it is worth noting that coach-team matches such as Tom Osborne from Nebraska and Joe Paterno from Penn State will not have match dummies because both coaches only have one match in the sample period. Thus, following Chapman and Southwick (1991), we avoid ranking coaches from best to worst because a ranking would combine information about both the coach and about his coach-team match.

Conclusion

Using a large data set for football coaches, a direct test of the matching hypothesis shows that the productivity of coaches varies across teams. Specifically, we find that a coach-team match significantly impacts the variation in collegiate football teams winning percentages. We also show that coaches tend to face an increasing probability of turnover early in their tenure and then subsequently face a decreasing probability of turnover once they have been with a team for approximately four or

five years, consistent with the pattern predicted by Jovanovic (1979). Defining a good match as one that endures beyond five years, we estimate the value of a good match as an improvement in a team's winning percentage of approximately 5 percent. We conclude that mobility in the market for college football coaches follows a process similar to that predicted by the matching model.

References

1. Allgood, S., and K.A. Farrell, "The Match between CEO and Firm," *Journal of Business*, 76, no. 2 (2003), pp. 317-342.
2. Borland, J., and J. Lye, "Matching and Mobility in the Market for Australian Rules Football Coaches," *Industrial and Labor Relations Review*, 50, no. 1 (1996), pp. 143-158.
3. Chapman, K.S., and L. Southwick, Jr., "Testing the Matching Hypothesis: The Case of Major-League Baseball," *American Economic Review*, 81 (December 1991), pp. 1352-1360.
4. Fort, R.D., *Sports Economics* (Upper Saddle River, New Jersey: Prentice Hall, 2003).
5. Garen, John E., "Empirical Studies of the Job Matching Hypothesis," *Research in Labor Economics*, 9 (1999), pp. 187-224.
6. Hersch, J., and P. Reagan, "Job Match, Tenure and Wages Paid by Firms," *Economic Inquiry*, 28 (1990), pp. 488-507.
7. Jovanovic, B., "Job Matching and the Theory of Turnover," *Journal of Political Economy*, 87 (1979), pp. 972-990.
8. Ohtake, F., and Y. Ohkusa, "Testing the Matching Hypothesis: The Case of Professional Baseball in Japan with Comparisons to the United States," *Journal of the Japanese and International Economies*, 8 (1994), pp. 204-219.
9. Prisinzano, R., "Investigation of the Matching Hypothesis: The Case of Major League Baseball," *Journal of Sports Economics*, 1 (August 2000), pp. 277-298.

Copyright of Quarterly Journal of Business & Economics is the property of College of Business Administration/Nebraska and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.