Journal of Business Administration Online
ARKANSAS TECH UNIVERSITY College of Business HTTPS://JBAO.ATU.EDU
Spring 2024, Vol. 18 No. 1

Mixed Strategy and The Money Down: A Field Test from American College Football

Jacob Bundrick, University of Central Arkansas

Abstract

Minimax serves a foundational role in our understanding of strategic behavior, much like that found in business and political settings. Investigations of minimax have traditionally utilized laboratory experiments with novice players but have increasingly moved toward field studies of professional athletes. Implicit in this move has been the shift away from choice sets precisely defined by researchers to assumed choice sets that may not represent the real-world decisions of players. This paper harmonizes real-world choices with those analyzed by examining the offensive third-down play of the University of Central Arkansas's American college football team. Interviews with a coach revealed the use of grouping techniques to reduce the number of in-game decisions, conserving cognitive resources and facilitating strategic behavior. I model third-down play as the matching pennies game and test the equilibrium prediction of equalized success probabilities across strategies. Results indicate that grouping techniques can simplify the implementation of optimal play in complex settings, but that creating appropriate groupings can be difficult.

Key words: Matching pennies, mixed strategy, strategic behavior

Introduction

Strategic play is required in many business and political settings. At the heart of strategic behavior is von Neumann's minimax theorem. Whether the predictions of minimax hold in practice has been the focus of much scholarship. Early work largely studied novice players in laboratory settings where researchers precisely defined choice sets. Researchers later examined decision making in the real world by studying professional athletes in competition. This field research had two advantages: participants had strong incentives to play an optimal strategy and they were experienced.

The evolution of mixed-strategy analyses from laboratory to field has improved our real-world understanding of strategic behavior. However, an important feature was lost as research shifted domains. Researchers no longer defined the choices of participants and instead assumed the choice set players faced. Without intimate knowledge of the game, an assumed choice set may differ from the choices made by players, particularly in complex settings. Morabito (2024) found evidence of this disconnect when examining the actions of female professional soccer players and the previous theoretical assumptions of the discipline. Aligning theoretical assumptions with real-world actions is important because improper assumptions may bias researchers' conclusions and distort our understanding of strategic behavior.

This paper harmonizes the choices made by players with those analyzed using a unique dataset from a representative American college football team. Specifically, I identify the third-down choices made by the offense of the University of Central Arkansas's (UCA) football team. Interviews with a UCA coach revealed the use of grouping techniques to reduce the number of in-game decisions. Because playing minimax in real-world settings is difficult (Kovash & Levitt, 2009), such grouping techniques can conserve cognitive resources and simplify implementation of strategic play. Statistical analysis is used to test whether in-game play reflects the intended strategy of UCA's offense. I then model third downs as the matching pennies game and empirically test the equilibrium prediction of equalized success probabilities across strategies. Determining whether grouping techniques facilitate optimal play is important because deviations from equilibrium reduce output (Kovash & Levitt, 2009).

The results of this paper indicate that in most – but not all – distance groupings, UCA adheres to its pre-defined choice sets. Regression analysis yields that UCA exhibits optimal play both across and within distance bins, a finding that contradicts previous studies of play selection. These results suggest that reducing the number of decisions can simplify the implementation of mixed strategies in complex settings. Appropriately consolidating the many unique decisions one faces into a smaller number of choices, though, can be difficult.

Related Literature

Early investigations of minimax relied on data generated by laboratory experiments where researchers could precisely define choice sets. Most experiments found that participants deviated from equilibrium play because they lacked the appropriate experience or incentives to play strategically (Brown & Rosenthal, 1990; Ochs, 1995; Erev & Roth, 1998). Accordingly, scholars turned to data generated by sports where expert players are experienced and highly motivated to play strategically.

The evidence provided by sports offers stronger support for equilibrium play. Investigations of penalty kicks in soccer regularly indicate that players both equalize their probabilities of success across strategies and randomize their play (Chiappori, Levitt, & Groseclose, 2002; Palacios-Huerta, 2003; Coloma, 2007; Dohmen & Sonnabend, 2018; Malkov, 2018; Morabito, 2024). Most studies utilizing professional tennis reveal that players' win rates are equalized when serving left or right, though the evidence as to whether players randomize serve directions is more mixed (Walker & Wooders, 2001; Hsu, Huang, & Tang, 2007; Spiliopoulos, 2018; Gauriot, Page, & Wooders, 2023; Anderson et al., 2024).

More complex sports environments offer less-than-conclusive findings. The evidence from professional baseball is dependent on the strategic aspect of the game. For instance, hitters deciding whether to swing at the first pitch do not equalize payoffs across their respective strategies, but base runners deciding whether to steal second base do (Downey & McGarrity, 2015; Choe & Kim, 2019). Most – but not all – evidence indicates that pitchers selecting pitch types fail to equalize payoffs across their respective strategies, a result that Bhattacharya and Howard (2022) argue stems from rational inattention (Kovash & Levitt, 2009; Hsiao et al., 2024; White & Smith, 2024).

Moreover, Gmeiner (2019) finds evidence that pitchers and hitters differ in their ability to make history-based adjustments towards optimal play.

Evidence from professional American football is also mixed. Studying play selection in the National Football League (NFL), Kovash and Levitt (2009) found evidence to reject minimax play as passing plays outperform running plays. Moreover, teams switch too frequently between play types to be random (Kovash & Levitt, 2009; Emara et al., 2017). McGarrity and Linnen (2010), however, found that NFL offenses demonstrate strategic behavior in other ways. A team's mix of runs and passes does not change when a substitute quarterback replaces an injured starting quarterback and teams randomize play types when only first-and-ten plays are considered.

Why do some professionals play minimax, but others do not when all are highly experienced and motivated? One reason may be that the analyses of sports settings have often relied on assumed choice sets which may not reflect the real-world choices made by players. This problem becomes more acute as the complexity of the game increases, such as in the cases of baseball and American football. This paper addresses this issue by identifying the real-world choices made by a representative American college football team and subsequently examining whether players make optimal decisions.

Third Downs in College Football

The Role of Third Downs

Football games consist of drives, or sequences of plays in which one team's offense competes against the opposing team's defense. The offense attempts to advance the ball to score points while the defense attempts to prevent them from doing so. The offense is given four downs (i.e., plays) at the beginning of each drive to advance a total of ten yards. "First down" corresponds to the first play in the set of downs, "second down" corresponds to the second play, and so forth. Gaining the ten yards converts a first down, in which the offense is awarded a new set of four downs to continue its drive. Failing to convert transfers possession of the ball to the opposing team.

Scoring drives typically require the offense to execute a series of plays that intermittently earn first downs. Consequently, third downs play a more critical role in sustaining drives than any other down. An offense that fails to convert on first or second down has additional opportunities to gain the necessary yardage. Fourth down may be used to run an offensive play, but game strategy often dictates that teams use it in other ways. In this sample, UCA ran an offensive play on just 44 of the fourth downs following one of their 304 failed third-down attempts (14.5%). Accordingly, third-down plays frequently determine whether an offense will continue its drive. Hence, third down is known as "the money down."

Third downs are the preferred play to examine strategic behavior in American football because it is the down where each team's goals are the clearest. It can be assumed that, on a typical third down, the offense has the immediate goal of converting while the defense has the immediate goal of preventing the conversion. This assumption does not hold on first and second down given the additional opportunities to convert.

Previous literature has considered this and used other measures of success (i.e., yards gained, change in net expected points) to examine minimax. However, such measures may not reflect the offense's goals. For instance, offenses often use first or second down to manipulate the defense into playing a particular strategy or to learn about a defense's strategy. These plays are important to the offense's overall success but may not be captured by traditional measures of success. For example, an offense may throw a long pass on first down to manipulate the defense into defending the pass more frequently. Whether the pass is completed is irrelevant to the offense's strategy.

Game Planning to Conserve Cognitive Resources

In-game strategy is dictated by the situations that arise in the game. While the number of possible situations is numerous, coaches possess scarce cognitive resources to address them in real time. A key part of decision making, then, is game planning, or preparing for potential in-game scenarios before the game arrives.

Interviews with a UCA coach revealed that one component of third-down game planning is discretizing distance to a first down to simplify in-game decision making. Coaches create a set of contiguous bins that stretch across the possible third down distances to organize play calls. An offensive game plan then includes pools of plays that coaches favor in each of the various third-down groupings. The intent is to apply similar strategy to each distance within a bin, but the strategies applied across bins may be distinct.

Discretizing distance to a first down reduces the number of real-time decisions to be made. Coaches can conserve their cognitive resources by pre-determining the in-game situations that are similar enough to approach with the same strategy. This grouping practice can be empirically exploited to determine whether such simplifying strategies facilitate optimal decision making in complex settings.

Theoretical Framework

Third downs can be modeled as a 2x2 normal-form game between two players, where the offense competes against the defense. The offense attempts to maximize the probability of converting on third down while the defense attempts to minimize the offense's probability of converting. Though the strategy spaces for both players are nuanced, each team's actions can be simplified to two options. The offense can run (R) or pass (P) the ball and the defense can defend the run (R) or pass (P). Simplification is possible because the expectation in a mixed-strategy Nash equilibrium is that all plays called with positive probability have equal payoffs (Gauriot, Page, & Wooders, 2023). The outcome of the play is clear and determined quickly.

Figure A presents the offense's payoff matrix. The offense's probabilities of converting on third down are represented by the four quantities π_OD, where O is the offense's play call (O= R,P) and D is the defense's play call (D=R,P). Because the third-down play is a zero-sum game, where either the offense converts or the defense prevents conversion, the defense's probabilities of preventing a third-down conversion can be expressed as $1 - \pi_{OD}$.

Figure A

Third-Down Conversion Payoff Matrix

	Defensive Play Type	
Offensive Play Type	R	P
R	π_{RR}	π_{RP}
P	π_{PR}	π_{PP}

Note: Each cell represents the offense's probability of converting on third-down, π_{OD}.

The probability that the offense converts on third down depends, in part, on the choices made by the defense. Likewise, the probability of the defense preventing a conversion depends on the strategy employed by the offense. Thus, two assumptions can be made about the offense's probability of converting.

ASSUMPTION 1:

(1) $\pi_{PR} > \pi_{RR}$ and $\pi_{RP} > \pi_{PP}$

ASSUMPTION 2:

(2) $\pi_{RP} > \pi_{RR}$ and $\pi_{PR} > \pi_{PP}$

Assumption 1 states that the offense is more likely to convert when they choose a play type the defense is not prepared to defend. If the offense knew what play type the defense expected, they would call a play of the opposite type. Assumption 2 states the offense is less likely to convert when the defense calls a play designed to defend the offense's play type. If the defense knew what play type the offense would call, the defense would choose to match it. This is the matching pennies game (McGarrity & Linnen, 2010).

atching pennies requires simultaneous play. If not, one player has a second-mover advantage. Though challenges in observing the defense's strategy prevent a formal test, the simultaneous-play assumption is reasonable because game rules restrict time between plays. Moreover, strategic differences make this assumption stronger in college football than in the NFL as previously modeled by Kovash and Levitt (2009), McGarrity and Linnen (2010), and Emara et al. (2017). NFL offenses actively seek a second-mover advantage by authorizing quarterbacks to change the play selection after surveying the defense. College quarterbacks are not typically given this same autonomy.

With enough iterations, the matching pennies game converges to a mixed-strategy Nash equilibrium. Of interest to this analysis is the equilibrium prediction that a player's probabilities of success are equalized across each strategy. Specifically, the probability the offense converts on third down (defense prevents) should be the same when passing or running the ball (defending the pass or run). If not, the offense (defense) is failing to maximize the overall probability with which it converts (prevents) on third down. This equilibrium prediction is testable for the offense but not the defense as only the offense's strategy is readily observed.

Data

The data is comprised of UCA's offensive third downs played in regulation during the 2018, 2019, and 2020 football seasons. UCA is a member of the National Collegiate Athletic Association's Division 1 Football Championship Subdivision (FCS), the second most prominent division of college athletics. Though membership may vary from year to year, the FCS is typically comprised of nearly 130 member institutions. These schools span 37 states and the District of Columbia. Each institution must meet membership requirements, such as minimum financial aid and sport sponsorship requirements, that create a group of similar football-playing institutions. Accordingly, I use UCA as a representative university.

The data were obtained by viewing each of UCA's 33 games played during these three seasons. Among the observable features of each play are the defensive team, the offense's play type, whether the play resulted in a first down, and several in-game characteristics that may affect play calling and third-down conversion rates. These include the distance to a first down, point margin (UCA's score minus the opponent's), field position, and game clock features, such as the quarter of play and whether the game is within the two-minute mark of each half.

Each play is categorized as a run or pass. Pass plays encompass all designed pass plays, including those that did not result in the quarterback passing the ball. This distinction is important for two reasons. First, quarterbacks may scramble, or run the ball, instead of throwing the ball on designed pass plays. Quarterbacks with scrambling ability create additional optionality within passing plays, making called pass plays relatively more attractive. In this sample, UCA's quarterback scrambled on 5.7% of the team's pass plays, half of which converted a first down. Second, defenses can defend a pass by sacking the quarterback, or tackling him before he throws the ball. An offense's inability to avoid sacks make designed pass plays relatively less attractive. UCA's quarterback was sacked on 9.1% of its third-down pass plays.

Of note is that I include third downs occurring in the second half of games because they carry more pressure than those occurring earlier in the game. Teams have time to recover from failed third-down attempts in the first half, but the pressure to execute increases as the time remaining in the game dwindles. In these situations, the strategic decisions made by the two teams are critically important to the outcome of the game. Moreover, second-half plays reflect halftime adjustments. That is, they incorporate in-game learning.

Third-down plays that are not well modeled by the matching pennies game are excluded. Specifically, third downs converted because of defensive penalties rather than an offensive play are omitted. Likewise, third-down plays used to manage the game clock rather than attempt to convert were dismissed. Clock-management plays include "spikes" to stop a running clock or "kneels" to keep the clock running. In total, there are 476 third-down plays in the sample. Table 1 provides summary statistics.

Table 1

Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max
Conversion	0.361	-	0	1
Pass	0.739	-	0	1
Yards to First	7.286	4.911	1	35
Distance to First (Category)				
Short (1-3 yards)	0.233	-	0	1
Medium (4-6 yards)	0.263	-	0	1
Long (7-12 yards)	0.376	-	0	1
Extra Long (13+ yards)	0.128	-	0	1
Point Margin	-1.458	12.527	-34	42
Quarter				
1^{st}	0.252	-	0	1
2^{nd}	0.286	-	0	1
3^{rd}	0.231	-	0	1
4^{th}	0.231	-	0	1
Two Minute	0.076	-	0	1
Yards to TD	52.513	24.145	1	99
Field Position (Category)				
Shadow	0.040	-	0	1
Open Field	0.824	-	0	1
Red Zone	0.137	-	0	1
Home	0.408	-	0	1

Note: The sample includes 476 observations comprised of UCA's offensive 3^{rd} downs played in regulation during the 2018-2020 seasons.

Third Down by Distance to First

Interviews with a UCA coach informed the structure of game situation variables to reflect the decision-making processes involved in play selection. Most importantly is the way in which distance to a first down is viewed. Though traditionally measured in yards as a continuous variable (Kovash & Levitt, 2009; McGarrity & Linnen, 2010; Emara et al., 2017), play callers discretize distance to a first down to simplify in-game decision making. I construct the “distance to a first down” variable to reflect the way in which UCA groups third-down distances. Specifically, distance to a first down is a categorical variable that includes short-yardage (1-3 yards), medium-yardage (4-6 yards), long-yardage (7-12 yards), and extra-long-yardage situations (13+ yards). Grouping these decisions into bins effectively reduces the many unique decisions into four decisions.

This discretization technique leads to two empirical questions. First, does the offense treat each yardage within a bin as though it is the same distance? Second, does the offense play minimax across each distance grouping and within each distance grouping? Table 2 provides a first attempt to answer these questions using a series of Chi-squared tests similar to Walker and Wooders (2001) and Hsu, Huang, and Tang (2007).

Consider first the way in which the coach treats each distance within a bin. For each bin, the coach may choose from a pool of plays. Because a given pool of plays is available when facing each distance within a bin, the strategy applied to each respective distance should converge as the number of attempts increases so that all distances within the bin are treated as though they are the same distance. Given a large enough sample, discretization should yield homogeneity of strategy within each distance bin.

A Pearson’s Chi-squared test is applied to examine whether there are equal distributions of pass plays across each yardage included in a respective distance bin. Let O_P^d represent the probability with which UCA’s offense O elects to pass P when facing distance in yards d. The null hypothesis when examining the short yardage distance bin is then $H_o\colon O_P^1 = O_P^2 = O_P^3$. Similar null hypotheses are constructed for each respective distance bin.

Traditionally, p-values are used to determine how likely it is that the null hypothesis is *not* true. In this case, the p-value is examined to determine how likely it is that the data *is* compatible with the null hypothesis. To be clear, it cannot be proved that the null hypothesis is true, but a larger p-value signals an increased likelihood that the null is true.

The results in Table 2 reveal that homogeneity of strategy exists within the medium-, long-, and extra-long-yardage bins with reasonable likelihood (p-values of 0.354, 0.355, and 0.244, respectively). However, the null hypothesis is rejected at the 1% level in the short-yardage bin, suggesting that each distance within the short-yardage bin is not approached with similar strategy. This result is driven by the strategy on third-and-one plays, where runs are twice as likely to be called than on third-and-two and third-and-three. In fact, third-and-one is the only distance in the data set where runs are more likely to be called than passes, suggesting that third-and-one is a special case.

The second consideration is whether UCA's offense played mixed strategies across and within each distance bin, specifically with respect to the prediction of equalized success probabilities across runs and passes. This, too, can be examined using a Pearson's Chi-squared test. Let π_O^d represent the offense's probability of converting when choosing to run R or pass P when facing distance in yards d. The null hypothesis for each distance is then $H_o{:}\,\pi_R^d = \pi_P^d$.

Table 2 indicates that the null hypothesis of equalized success probabilities across runs and passes is rejected at the 10% level in short-yardage situations as run plays outperformed passing plays. UCA converted 68.3% of its runs in short-yardage distances compared to just 50% of its passes, driven primarily by the disparity in conversion rates in third-and-two situations. This statistically significant difference suggests that UCA did not run the ball enough on third-and-two to be consistent with optimal play.

The null hypothesis cannot be rejected at conventional levels in medium-, long-, and extra-long-yardage situations. The respective p-values are 0.358 in medium-yardage situations, 0.201 in long-yardage situations, and 0.889 in extra-long-yardage situations, which are comparable to the p-values found by Walker and Wooders (2001) and Hsu, Huang, and Tang (2007). These results provide support for equilibrium play and are strongest in the extra-long-yardage bin. However, it is worth noting that not all distances within each bin were played optimally as UCA did not pass the ball enough on third-and-five and third-and-nine to be optimal.

The Pearson's Chi-squared tests provide an important first approximation of the strategic behavior exhibited by UCA's offense. However, the decision to run or pass and the success of that decision is likely to be affected by game situations that are not controlled for in these tests. A more rigorous examination is required.

Table 2

Third Downs by Distance to First Down

	Play Type (#)			Run Pass Mix (%)				Conversions (#)		Conversion Rate		$H_0: \pi_R^d = \pi_P^d$	
Distance	R	P	Total	R	P	χ^2	*p*-value	R	P	R	P	χ^2	*p*-value
Short	**63**	**48**	**111**	**0.568**	**0.432**			**43**	**24**	**0.683**	**0.500**	**3.7940**	**0.051***
1	39	10	49	0.796	0.204			28	7	0.718	0.700	0.0126	0.911
2	12	19	31	0.387	0.613			10	8	0.833	0.421	5.1341	0.023**
3	12	19	31	0.387	0.613			5	9	0.417	0.474	0.0965	0.756
				$H_0: O_P^1 = O_P^2 = O_P^3$		18.6379	0.000***						
Medium	**20**	**105**	**125**	**0.160**	**0.840**			**6**	**43**	**0.300**	**0.410**	**0.8455**	**0.358**
4	8	37	45	0.178	0.822			3	16	0.375	0.432	0.0889	0.766
5	9	36	45	0.200	0.800			1	17	0.111	0.472	3.9120	0.048**
6	3	32	35	0.086	0.914			2	10	0.667	0.3125	1.5270	0.217
				$H_0: O_P^4 = O_P^5 = O_P^6$		2.0786	0.354						

Long	**32**	**147**	**179**	**0.179**	**0.821**			**6**	**44**	**0.188**	**0.299**	**1.6323**	**0.201**
7	6	32	38	0.158	0.842			3	10	0.500	0.313	0.7892	0.374
8	4	21	25	0.160	0.840			0	5	0.000	0.238	1.1905	0.275
9	5	21	26	0.192	0.808			0	9	0.000	0.429	3.2773	0.070*
10	11	50	61	0.180	0.820			3	16	0.273	0.320	0.0940	0.759
11	6	12	18	0.333	0.667			0	2	0.000	0.167	1.1250	0.289
12	0	11	11	0.000	1.000			N/A	2	N/A	0.181	N/A	N/A
$H_0: O_P^7 = O_P^8 = O_P^9 = O_P^{10} = O_P^{11} = O_P^{12}$						5.5298	0.355						
Extra Long	**9**	**52**	**61**	**0.148**	**0.852**			**1**	**5**	**0.111**	**0.096**	**0.0194**	**0.889**
$H_0: O_P^{13} = \cdots = O_P^{35}$						16.0871	0.244						
Totals	**124**	**352**	**476**	**0.261**	**0.739**			**56**	**116**	**0.452**	**0.330**	**5.9206**	**0.015***
$H_0: O_P^1 = \cdots = O_P^{35}$						111.0439	0.000***						

Notes: χ^2 represents the Pearson statistic. O_P^d represents the offense's probability of passing when facing distance d. π_O^d represents the offense's probability of converting when choosing to run R or pass P when facing distance d. Extra-long yardages are not individually displayed due to the limited number of plays when facing each respective distance. *P<0.10, **P<0.05, ***P<0.01

Econometric Approach

Formally investigating the predictions of mixed strategy involves the use of a binary response model. Probit and logit models are common in the mixed-strategy literature using sports data in which there is a need to control for the game situation (McGarrity & Linnen, 2010; Downey & McGarrity, 2015; Emara et al., 2017). However, the relatively small sample size in this paper requires more care in selecting the appropriate estimation technique.

To see this, consider the conditional probability of the logit and probit models given by equation (1).

$$p_i = Pr(y_i = 1|x) = F(x_i\beta) \tag{1}$$

In this equation, p is the probability with which the outcome variable is equal to 1, i indicates the third-down play, x is a vector of independent variables, and β is a vector of estimated coefficients for the independent variables. The logit model specifies that $F(\cdot)$ is the cumulative distribution function of the logistic distribution, while the probit model specifies that $F(\cdot)$ is the standard normal cumulative distribution function. In both models, β is estimated by way of maximum likelihood.

A maximum likelihood estimate exists if the solution to the log likelihood function is finite (Haberman, 1974). In small samples, though, finite solutions may not exist. One such case is when there is complete separation of the data. Complete separation occurs if the likelihood converges, but the outcome variable can be perfectly separated into a group (1 or 0 in a binary response model) by a single explanatory variable. Albert and Anderson (1984) show that if there is complete separation, the parameter estimate $\hat{\beta}$ is infinite. Thus, the maximum likelihood estimate of $\hat{\beta}$ does not exist.

The data utilized in this paper suffer from complete separation. Consider that UCA failed to convert on third down in ten attempts when playing the University of Tulsa. The indicator variable for Tulsa, then, completely separates the outcome variable into the non-response group. For each play in which $Tulsa = 1$, $Conversion\ =\ 0$ and the maximum likelihood estimator leads to an infinite estimate of $\hat{\beta}_{Tulsa}$. Consequently, traditional logit and probit models are inappropriate estimation techniques.

One way to address complete separation is to use the Firth logit. The Firth logit estimates $\hat{\beta}$ by way of a penalized maximum likelihood estimator in which a correction is applied to the score function that generates the maximum likelihood estimate (Firth, 1993). As Firth (p. 28) explains, “the bias in [$\hat{\beta}$] can be reduced by introducing a small bias into the score function.” Because the bias correction is applied to the score function rather than the estimate of $\hat{\beta}$ itself, the penalized maximum likelihood estimator can be employed in cases where there is separation of the data and $\hat{\beta}$ would otherwise be infinite. Heinze and Schemper (2002) demonstrate that the Firth logit guarantees finite estimates and is “an ideal solution” to the problem of separation (p. 2418). Accordingly, I employ the Firth logit.

Testing for Homogeneity Within Third-Down Bins

To test whether strategies are homogenous within each distance bin, I split the full sample into four subsamples based on the respective distance bins. I run the following regression on each of the four subsamples:

$$Pass_i = \alpha + \delta YardsToFirst_i + \beta X_i + \varphi_{SD} + \varepsilon_i \qquad (2)$$

where $Pass$ is a dummy variable equal to one if the offense passed on third-down play i, $YardsToFirst$ is a categorical variable indicating the yards to a first down in the respective distance bin, X is a vector of control variables including point margin, quarter of play, whether the play occurred with two minutes or less to play in a half, field position, and whether UCA was the home team. φ_{SD} represents season and defense dummies and ϵ is an error term. I then perform a Chi-squared test for joint significance on $YardsToFirst$ where the null hypothesis H_0 is that each distance to a first down within a given distance bin is treated the same.

The results of the short-, medium-, and long-yardage subsamples are presented in Table 3. The extra-long subsample is excluded due to the large quantity of unique distances (14), however, the Chi-2 statistic for joint significance of $YardsToFirst$ in this subsample is 2.66 with a p-value of 0.999. The null hypothesis cannot be rejected at conventional levels when the offense faces medium-, long-, and extra-long-yardage distances. This provides support for homogeneity of strategy for each distance within these respective bins. However, the null hypothesis is rejected at the 5% level when the offense is in short-yardage situations, suggesting that heterogeneous strategies are applied to the distances within the short-yardage bin.

Table 3

Testing for Homogeneity Within Third Down Bins

	Dependent Variable: Pass		
	Distance to Fist		
	Short	Medium	Long
Yards to First			
2	1.626**		
	(0.016)		
3	1.642**		
	(0.012)		
5		0.147	
		(0.839)	

6		0.875	
		(0.250)	
8			-0.820
			(0.362)
9			0.268
			(0.754)
10			-0.673
			(0.387)
11			-1.311
			(0.175)
12			1.604
			(0.377)
Chi-2 Statistic: Joint Significance of Yards to First	8.94**	1.52	5.02
	(0.011)	(0.468)	(0.413)
Full Set of Controls?	Yes	Yes	Yes
Season FE?	Yes	Yes	Yes
Defense FE?	Yes	Yes	Yes
Season X Defense?	Yes	Yes	Yes
Penalized Log Likelihood	-50.870	-43.574	-51.577
N	111	125	179

Notes: Dependent variable is a binary response equal to 1 if the offense passed. Base case for Short is 1 yard; for Medium is 4 yards; for Long is 7 yards. Controls include point margin, quarter, two-minute, field position, and home. P-values in parentheses. *P<0.10 **P<0.05 ***P<0.01.

The "Short" column in Table 3 reflects results for the third-and-short subsample. The base case for *Yards to First* is one yard. Both third-and-two and third-and-three are statistically significant at conventional levels, suggesting they are not treated the same as third-and-one. The coefficients suggest that UCA is more likely to pass on third-and-two and third-and-three than they are on third-and-one, consistent with the results in Table 2. The results do suggest, though, that the play calling strategies applied to third-and-two and third-and-three are not different from each other. This implies that third-and-one is a special case.

That third-and-one is a special case may be due to risk aversion (Romer, 2006; Goff & Locke, 2019; Yam & Lopez, 2019). Consider that across all UCA's offensive plays in the 2018-20 seasons, the median run play gained three yards while the median pass play gained just two yards. Moreover, 78.3% of UCA's runs gained at least one yard compared to just 52.9% of passes. Running the ball on third-and-one appears to provide more certainty than passing.

Running the ball on third-and-one may also be safer in terms of fan support. A team that regularly passes on third-and-one but fails to convert can expect to receive the ire of the fan base. Yet, the football community has a mantra to support failed rushing attempts on third-and-one: "we don't deserve to win if we can't gain a yard" (Pompei, 2012; Alper, 2019).

Testing for Equal Expected Payoffs

There are two considerations when examining the equilibrium prediction of equalized success probabilities when running or passing the ball. First, does UCA play each distance grouping optimally? Second, does UCA play each distance within a distance bin optimally?

To test whether UCA plays each distance grouping optimally, I split the full sample into three subsamples restricted to short-, medium-, and long-yardage plays, respectively. Extra-long-yardage situations are excluded because the small number of observations (61) and the volume of perfect predictions prevent valid estimations. The estimated equation is:

$$Convert_i = \alpha + \delta Pass_i + \beta X_i + \varphi_{SD} + \varepsilon_i \qquad (3)$$

where $Convert$ is a dummy variable equal to one if the offense converted on third-down play i, $Pass_i$ is an indicator variable equal to one if the offense passed, and X_i is the same vector of game controls described above. Season and defense indicator variables, φ_{SD}, are included to allow the payoffs to be season and matchup specific. A season-defense interaction is also included.

The parameter of interest, δ, is the coefficient on whether the offense passed. The null hypothesis H_O is that the likelihood of conversion is equal when passing or running. If the offense played mixed strategies, δ should be equal to zero. The null hypothesis cannot be proved, but a larger p-value signals an increased likelihood that the null is true.

To test whether UCA plays each distance within a distance bin optimally, I introduce interaction terms into equation (3). The estimated equation is:

$$Convert_i = \alpha + \delta Pass_i + \gamma YardsToFirst_i + \omega Pass_i YardsToFirst_i + \beta X_i + \varphi_{SD} + \varepsilon_i \qquad (4)$$

where $Convert$ is a dummy variable equal to one if the offense converted on third-down play i, $Pass_i$ is an indicator variable equal to one if the offense passed, $YardsToFirst_i$ is an indicator variable equal to one if the distance faced on play i is a specified distance, and $Pass_i YardsToFirst_i$ is an interaction term between $Pass_i$ and $YardsToFirst_i$. The parameters of interest in equation (4) are δ, γ, and ω. The null hypothesis H_O is that the likelihood of conversion is equal when passing or running. If the offense played optimally, $\delta, \gamma,$ and ω should be equal to zero.

Table 4 presents the results for the short-yardage bin. Column 1 reflects equation (3). The p-value on δ is 0.820, which fails to reject the null hypothesis. This provides reasonably strong support that UCA's success probabilities are equalized when running and passing the ball in short-yardage situations. Columns 2 and 3 reflect equation (4), with emphasis placed on third-and-two given the results of Table 2. In both estimations, the parameters of interest are statistically insignificant. This supports the notion that UCA played mixed strategies when facing each distance within the short-yardage bin.

Table 4

Testing Equal Payoffs – Short Yardage

	Dependent Variable: Conversion		
	(1)	(2)	(3)
Pass	0.113	0.230	1.409
	(0.820)	(0.690)	(0.194)
Yards to First			
2		0.679	0.903
		(0.460)	(0.359)
3			-0.721
			(0.398)
Pass × 2		-0.687	-2.119
		(0.540)	(0.173)
Pass × 3			-1.287
			(0.361)
Point Margin	-0.069	-0.064	-0.065
	(0.118)	(0.156)	(0.139)
Quarter			
2^{nd}	-0.200	-0.232	-0.175
	(0.768)	(0.732)	(0.803)
3^{rd}	-0.080	-0.098	-0.186
	(0.899)	(0.875)	(0.773)
4^{th}	1.748**	1.684**	1.729**
	(0.040)	(0.048)	(0.050)

Two Minute	-0.663	-0.481	0.221
	(0.527)	(0.656)	(0.849)
Field Position			
Red Zone	-0.269	-0.343	-0.455
	(0.637)	(0.551)	(0.451)
Shadow	2.154	2.185	2.632
	(0.277)	(0.270)	(0.225)
Home	0.448	0.508	0.653
	(0.678)	(0.648)	(0.581)
Constant	-0.013	-0.051	0.633
	(0.990)	(0.960)	(0.576)
Season FE?	Yes	Yes	Yes
Defense FE?	Yes	Yes	Yes
Season x Defense?	Yes	Yes	Yes
Penalized Log Likelihood	-54.721	-54.122	-52.244
N	111	111	111

Notes: Dependent variable is a binary response equal to 1 if the offense converted. Base case for *Yards to First* is 1 yard; for *Quarter* is 1st; for *Field Position* is Open Field. P-values in parentheses. **P<0.05 and ***P<0.01.

Table 5 presents results for medium-yardage plays. Column 1 reflects equation (3). The p-value on the estimated parameter for pass is 0.468. The null hypothesis cannot be rejected at conventional levels, supporting the notion that UCA played the group of medium-yardage distances optimally. Columns 2 and 3 reflect equation (4), with emphasis on third-and-five given the results of Table 2. The parameters of interest are statistically insignificant in both estimations. This suggests that UCA played minimax when facing each distance within the medium-yardage bin.

Table 5

Testing Equal Payoffs – Medium Yardage

	Dependent Variable: Conversion		
	(1)	(2)	(3)
Pass	0.531	0.508	0.747
	(0.468)	(0.549)	(0.468)
Yards to First			
5		0.478	0.857
		(0.713)	(0.546)
6			1.225
			(0.454)
Pass × 5		-0.075	-0.364
		(0.957)	(0.809)
Pass × 6			-1.075
			(0.547)
Point Margin	-0.086**	-0.083**	-0.079*
	(0.030)	(0.039)	(0.053)
Quarter			
2nd	0.685	0.755	0.785
	(0.277)	(0.240)	(0.230)
3rd	-0.168	-0.189	-0.157
	(0.808)	(0.786)	(0.827)
4th	0.414	0.367	0.362
	(0.557)	(0.614)	(0.626)
Two Minute	-1.198	-1.211	-1.163
	(0.167)	(0.187)	(0.219)
Field Position			
Red Zone	-0.252	-0.254	0.316
	(0.691)	(0.687)	(0.623)
Shadow	0.407	0.551	0.786
	(0.762)	(0.680)	(0.570)
Home	-0.341	-0.427	-0.366

	(0.811)	(0.766)	(0.798)
Constant	-0.593	-0.741	-1.020
	(0.664)	(0.608)	(0.508)
Season FE?	Yes	Yes	Yes
Defense FE?	Yes	Yes	Yes
Season x Defense?	Yes	Yes	Yes
Penalized Log Likelihood	-59.590	-58.966	-58.905
N	125	125	125

Notes: Dependent variable is a binary response equal to 1 if the offense converted. Base case for *Yards to First* is 4 yards; for *Quarter* is 1st; for *Field Position* is Open Field. P-values in parentheses. *P<0.10 and **P<0.05.

Table 6 presents results for the long-yardage bin. Column 1 reflects equation (3). The p-value on the parameter of interest is 0.163, which fails to reject the null hypothesis at conventional levels. Thus, the evidence does not rule out that UCA played mixed strategies when facing long-yardage distances. Columns 2 and 3 reflect equation (4), emphasizing third-and-nine given the results of Table 2. Note that no estimation exists for Pass x 12 because UCA did not run on any third-and-twelve plays. The parameters of interest are statistically insignificant in each estimation, providing support for optimal play when facing each distance within the long-yardage bin.

Overall, the evidence from regression analysis suggests that UCA plays optimally across the various distance bins the team has pre-defined. The evidence is strongest in the short-yardage bin and weakest in the long-yardage bin. The evidence also indicates that UCA plays each distance within the respective bins optimally. Recall from Table 2 that the fraction of third downs in which UCA passed ranged from a low of 43.2% in short-yardage situations to a high of 84.0% in medium-yardage situations (excluding extra-long yardage situations). The range of run-pass mixes coupled with these regression results suggest that UCA's play selection is sophisticated enough to maintain equilibrium play while adjusting to the distance they face.

Table 6

Testing Equal Payoffs – Long Yardage

	Dependent Variable: Conversion		
	(1)	(2)	(3)
Pass	0.880	0.472	-0.692
	(0.163)	(0.483)	(0.593)
Yards to First			

8			-0.799
			(0.703)
9		-1.011	-2.931
		(0.574)	(0.211)
10			-1.379
			(0.281)
11			-3.008
			(0.166)
12			0.376
			(0.714)
Pass × 8			-0.015
			(0.995)
Pass × 9		1.685	3.302
		(0.375)	(0.182)
Pass × 10			1.356
			(0.321)
Pass × 11			2.044
			(0.374)
Pass × 12			
Point Margin	-0.056**	-0.059**	-0.057**
	(0.037)	(0.028)	(0.029)
Quarter			
2nd	0.637	0.641	0.831
	(0.227)	(0.226)	(0.151)
3rd	0.234	0.343	0.188
	(0.696)	(0.571)	(0.756)
4th	0.083	0.242	0.447
	(0.895)	(0.708)	(0.499)
Two Minute	0.126	0.042	0.420
	(0.879)	(0.962)	(0.669)
Field Position			

Red Zone	-0.108	-0.087	-0.209
	(0.874)	(0.899)	(0.762)
Shadow	-0.278	-0.394	-0.507
	(0.772)	(0.689)	(0.620)
Home	-0.538	-0.453	-0.419
	(0.621)	(0.677)	(0.710)
Constant	-3.773**	-3.512**	-2.081
	(0.023)	(0.034)	(0.285)
Season FE?	Yes	Yes	Yes
Defense FE?	Yes	Yes	Yes
Season x Defense?	Yes	Yes	Yes
Penalized Log Likelihood	-71.787	-70.914	-68.950
N	179	179	179

Notes: Dependent variable is a binary response equal to 1 if the offense converted. Base case for *Yards to First* is 7 yards; for *Quarter* is 1st; for *Field Position* is Open Field. P-values in parentheses. *P<0.10 and **P<0.05.

Conclusion

Field studies of minimax have greatly improved our understanding of real-world strategic behavior. However, they have often relied on assumed choice sets that may deviate from the actual decisions made by players. This paper bridges the gap by identifying the real-world choice sets defined by UCA's American college football team on third down, determining whether the team adheres to those pre-defined choice sets in practice, and testing whether the team plays according to the equilibrium prediction of mixed strategy.

Playing minimax in complex, real-world situations – such as those provided in business and in politics – is difficult. To simplify this challenging task, UCA's coaches reduce the number of decisions they must make by grouping similar scenarios together and subsequently applying parallel strategies. Evidence provided by statistical tests suggests that UCA applies homogenous strategies when facing medium-, long-, and extra-long-yardage situations, but not in short-yardage situations. Third-and-one, third-and-two, and third-and-three are grouped into the short-yardage bin, but third-and-one is treated differently as run plays occur with higher frequency than on third-and-two and third-and-three. However, this run-pass mix equalized success probabilities, suggesting that third-and-one may warrant its own bin.

A casual analysis of the equilibrium prediction of equalized success probabilities across strategies yielded that UCA played mixed strategies in most distance scenarios.

A more rigorous examination provided stronger evidence that UCA played optimally both across distance bins and within distance bins. This implies that reducing the number of decisions to be made by creating groupings can save cognitive resources and simplify the implementation of mixed strategies in complex, real-world settings. However, creating appropriate groupings can be difficult.

I conclude with three caveats. First, this analysis relies on a relatively small sample of third-down plays, which restricts the ability to generate precise estimates. One may find different results should more plays become available. Second, this analysis is specific to a single team. Many American football teams implement discretization, but the exact nature of play calling described above is unique to UCA. Incorporating other teams into this analysis requires intimate knowledge of their play-calling schemes. While UCA's football program is similar to those of other FCS institutions, the generalizability of these results is dependent on the degree to which UCA is representative of other teams. Future work should consider the increasing role data analytics plays in college football and the opportunities that presents for understanding real-world decision making (Connelly, 2022). Academicians are not only well positioned to support their institution's sports programs, but to use that access to answer some of the most pressing questions regarding strategic behavior. Finally, this analysis says nothing of the second equilibrium prediction of minimax. I leave investigations of serial independence to further work.

References

Albert, A., Anderson, J. A., (1984), *On the existence of maximum likelihood estimates in logistic regression models*, Biometrika, Vol. 71, no. 1, pp. 1-10.

Alper, J. (2019, October 20). Anthony Lynn: You don't deserve to win if you can't get one yard. *Pro Football Talk, NBC Sports*. https://profootballtalk.nbcsports.com/2019/10/20/anthony-lynn-you-dont-deserve-to-win-if-you-cant-get-one-yard/

Anderson, A., Rosen, J., Rust, J., Wong, K. P., (2024), *Disequilibrium play in tennis*, Available at SSRN: https://ssrn.com/abstract=4383716.

Bhattacharya, V., & Howard, G., (2022), *Rational inattention in the infield*, American Economic Journal: Microeconomics, Vol. 14, no. 4, pp. 348-393.

Brown, J. N., Rosenthal, R. W., (1990), *Testing the minimax hypothesis: A re-examination of O'Neill's game experiment*, Econometrica: Journal of the Econometric Society, Vol. 58, no. 5, pp. 1065-1081.

Chiappori, P. A., Levitt, S., Groseclose, T., (2002), *Testing mixed-strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer,* American Economic Review, Vol. 92, no. 4, pp. 1138-1151.

Choe, J., Kim, J. S., (2019), *Minimax after money-max: Why major league baseball players do not follow optimal strategies,* Applied Economics, Vol. 51, no. 24, pp. 2591-2605.

Coloma, G., (2007), *Penalty kicks in soccer: An alternative methodology for testing mixed-strategy equilibria,* Journal of Sports Economics, Vol. 8, no. 5, pp. 530-545.

Connelly, B., (2022, March 30). State of analytics in college football: Fourth downs, 2-point conversions, Lane Kiffin and what's next, *ESPN+*. https://www.espn.com/college-football/insider/story/_/id/33616322/state-analytics-college-football-fourth-downs-2-point-conversions-lane-kiffin-next

Dohmen, T., Sonnabend, H., (2018), *Further field evidence for minimax play,* Journal of Sports Economics, Vol. 19, no. 3, pp. 371-388.

Downey, J., McGarrity, J. P., (2015) *Pick off throws, stolen bases, and southpaws: A comparative static analysis of a mixed strategy game*, Atlantic Economic Journal, Vol. 43, no. 3, pp. 319-335.

Emara, N., Owens, D., Smith, J., Wilmer, L., (2017), *Serial correlation in National Football League play calling and its effects on outcomes*, Journal of Behavioral and Experimental Economics, Vol. 69, pp. 125-132.

Erev, I., Roth, A. E., (1998), *Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria,* American Economic Review, Vol. 88, no. 4, pp. 848-881.

Firth, D., (1993), *Bias reduction of maximum likelihood estimates*, Biometrika, Vol. 80, no. 1, pp. 27-38.

Gauriot, R., Page, L., Wooders, J., (2023), *Expertise, gender, and equilibrium play*, Quantitative Economics, Vol. 14, no. 3, pp. 981-1020.

Gmeiner, M. W., (2019), *History-dependent mixed strategies: Evidence from Major League Baseball,* Journal of Sports Economics, Vol. 20, no. 3, pp. 371-398.

Goff, B., Locke, S. L., (2019), *Revisiting Romer: Digging deeper into influences on NFL managerial decisions*, Journal of Sports Economics, Vol. 20, no. 5, pp. 671-689.

Haberman, S. J., (1974): *The analysis of frequency data*, University of Chicago Press, Chicago.

Heinze, G., Schemper, M., (2002)., *A solution to the problem of separation in logistic regression*, Statistics in Medicine, Vol. 21, no. 16, pp. 2409-2419.

Hsiao, S. Y., Hu, S. H., Lin, M. J., Weng, W. C., (2024), *Do professional baseball players play mixed strategies? Evidence from MLB,* Available at SSRN: https://ssrn.com/abstract=4691177

Hsu, S. H., Huang, C. Y., Tang, C. T., (2007), *Minimax play at Wimbledon: Comment*, American Economic Review, Vol. 97, no. 1, pp. 517-523.

Kovash, K., Levitt, S. D., (2009), *Professionals do not play minimax: Evidence from Major League Baseball and the National Football League,* National Bureau of Economic Research, Working Paper, No. 15347, National Bureau of Economic Research, September, available online at https://www.nber.org/papers/w15347.

Malkov, E., (2018), *Minimax outside the top-5*, Available at SSRN: https://ssrn.com/abstract=3200795

McGarrity, J. P., Linnen, B, (2010), *Pass or run: An empirical test of the matching pennies game using data from the National Football League*, Southern Economic Journal, Vol. 76, no. 3, pp. 791-810.

Morabito, L., (2024), *Mixed-strategy equilibria and gender differences: The soccer penalty kick game*, Available at SSRN: https://ssrn.com/abstract=4819739

Ochs, J., (1995), *Games with unique, mixed strategy equilibria: An experimental study*, Games and Economic Behavior, Vol. 10, no. 1, pp. 202-217.

Palacios-Huerta, I., (2003), *Professionals play minimax*, The Review of Economic Studies, Vol. 70, no. 2, pp. 395-415.

Pompei, D. (2012, December 2). If Bears can't gain half-yard, they deserve to lose. *Chicago Tribune*. https://www.chicagotribune.com/sports/bears/ct-xpm-2012-12-02-ct-spt-1203-bears-pompei-chicago-20121203-story.html

Romer, D., (2006), *Do firms maximize? Evidence from professional football*, Journal of Political Economy, Vol. 114, no. 2, pp. 340-365.

Spiliopoulos, L., (2018), *Randomization and serial dependence in professional tennis matches: Do strategic considerations, player rankings and match characteristics matter?*, Judgment and Decision Making, Vol. 13, no. 5, pp. 413-427.

Walker, M., Wooders, J., (2001)., *Minimax play at Wimbledon*, American Economic Review, Vol. 91, no. 5, pp. 1521-1538.

White, D. R., Smith, B. O., (2024), *Changing it up: Determining the Nash equilibria for Major League Baseball pitchers*, American Behavioral Scientist, Available at https://doi.org/10.1177/00027642241235829.

Yam, D. R., Lopez, M. J., (2019), *What was lost? A causal estimate of fourth down behavior in the National Football League*, Journal of Sports Analytics, Vol. 5, no. 3, pp. 153-167.