Do Firms Maximize? Evidence from Professional Football

David Romer

University of California, Berkeley and National Bureau of Economic Research

This paper examines a single, narrow decision—the choice on fourth down in the National Football League between kicking and trying for a first down—as a case study of the standard view that competition in the goods, capital, and labor markets leads firms to make maximizing choices. Play-by-play data and dynamic programming are used to estimate the average payoffs to kicking and trying for a first down under different circumstances. Examination of actual decisions shows systematic, clear-cut, and overwhelmingly statistically significant departures from the decisions that would maximize teams' chances of winning. Possible reasons for the departures are considered.

I. Introduction

A central assumption of most economic models is that agents maximize simple objective functions: consumers maximize expected utility, and firms maximize expected profits. The argument for this assumption is not that it leads to perfect descriptions of behavior, but that it leads to reasonably good approximations in most cases.

The assumption that consumers successfully maximize simple objective functions frequently makes predictions about how a single individual will behave when confronted with a specific, easily describable decision.

I am indebted to Ben Allen, Laurel Beck, Sungmun Choi, Ryan Edwards, Mario Lopez, Peter Mandel, Travis Reynolds, Evan Rose, and Raymond Son for outstanding research assistance; to Christina Romer for invaluable discussions; to Steven Levitt and Richard Thaler for important encouragement; and to numerous colleagues and correspondents for helpful comments and suggestions. An earlier version of the paper was titled "It's Fourth Down and What Does the Bellman Equation Say? A Dynamic-Programming Analysis of Football Strategy."

[*Journal of Political Economy,* 2006, vol. 114, no. 2]

Thus it can often be tested in both the laboratory and the field. The assumption that firms maximize profits is much more difficult to test, however. Particularly for large firms, the decisions are usually complicated and the data difficult to obtain. But the a priori case for firm maximization is much stronger than that for consumer maximization. As Alchian (1950), Friedman (1953), Becker (1957), Fama (1980), and others explain, competition in the goods, capital, and labor markets creates strong forces driving firms toward profit maximization. A firm that fails to maximize profits is likely to be outcompeted by more efficient rivals or purchased by individuals who can obtain greater value from it by pursuing different strategies. And managers who fail to maximize profits for the owners of their firms are likely to be fired and replaced by ones who do. Thus the case for firm maximization rests much more on logical argument than empirical evidence. As Friedman puts it, "unless the behavior of businessmen in some way or other approximated behavior consistent with the maximization of returns, it seems unlikely that they would remain in business for long. . . . The process of 'natural selection' thus helps to validate the hypothesis [of return maximization]" (1953, 22).

This paper takes a first step toward testing the assumption that firms maximize profits by examining a specific strategic decision in professional sports: the choice in football between kicking and trying for a first down on fourth down. Examining strategic decisions in sports has two enormous advantages. First, in most cases, it is difficult to think of any significant channel through which strategic decisions are likely to affect a team's profits other than through their impact on the team's probability of winning. Thus the problem of maximizing profits plausibly reduces to the much simpler problem of maximizing the probability of winning. Second, there are copious, detailed data describing the circumstances teams face when they make these decisions.[1]

The predictions of simple models of optimization appear especially likely to hold in the case of fourth-down decisions in professional football. There are three reasons. First, the market for the coaches who make these decisions is intensively competitive. Salaries average roughly \$3 million per year, and annual turnover exceeds 20 percent.[2] Second, winning is valued enormously (as shown by the very high salaries commanded by high-quality players). And third, the decisions are unusually amenable to learning and imitation: the decisions arise repeatedly, and

[1] Thaler (2000) stresses the potential value of sports decision making in testing the hypothesis of firm optimization.

[2] The salary figure is based on the 23 coaches (out of 32) for whom 2004 salary information could be obtained from publicly available sources. The turnover data pertain to 1998–2004.

information about others' decisions is readily available. Thus a failure of maximization in this setting would be particularly striking.

This paper shows, however, that teams' choices on fourth downs depart in a way that is systematic and overwhelmingly statistically significant from the choices that would maximize their chances of winning. One case in which the departure is particularly striking and relatively easy to see arises when a team faces fourth down and goal on its opponent's 2-yard line early in the game.[3] In this situation, attempting a field goal is virtually certain to produce 3 points, while trying for a touchdown has about a three-sevenths chance of producing 7 points. The two choices thus have essentially the same expected immediate payoff. But if the team tries for a touchdown and fails, its opponent typically gains possession of the ball on the 2-yard line; if the team scores a touchdown or a field goal, on the other hand, the opponent returns a kickoff, which is considerably better for it. Thus trying for a touchdown on average leaves the opponent in considerably worse field position. I show later that rational risk aversion about points scored, concern about momentum, and other complications do not noticeably affect the case for trying for a touchdown. As a result, my estimates imply that the team should be indifferent between the two choices if the probability of scoring a touchdown is about 18 percent. They also imply that trying for a touchdown rather than a field goal would increase the team's chances of winning the game by about three percentage points, which is very large for a single play. In fact, however, teams attempted a field goal all nine times in my sample they were in this position.

Analyzing the choice between kicking and trying for a first down or touchdown in other cases is more complicated: the immediate expected payoffs may be different under the two choices, and the attractiveness of the distributions of ball possession and field position may be difficult to compare. Fortunately, however, the problem can be analyzed using dynamic programming. The choice between kicking and going for it leads to an immediate payoff in terms of points (which may be zero) and to one team having a first down somewhere on the field. That first down leads to additional scoring (which again may be zero) and to another possession and first down. And so on. Section II of the paper therefore uses data from over 700 National Football League (NFL) games to estimate the values of first downs at each point on the field (as well as the value of kicking off). To avoid the complications introduced when one team is well ahead or when the end of a half is approaching, I focus on the first quarter.

Section III uses the results of this analysis to examine fourth-down decisions over the entire field. To estimate the value of kicking, I use

[3] The Appendix summarizes the rules of football that are relevant to the paper.

the outcomes of actual field goal attempts and punts. Decisions to go for it on fourth down (i.e., not to kick) are sufficiently rare, however, that they cannot be used to estimate the value of trying for a first down or touchdown. I therefore use the outcomes of third-down plays instead. I then compare the values of kicking and going for it to determine which decision is better on average as a function of where the team is on the field and the number of yards it needs for a first down or touchdown. Finally, I compare the results of this analysis with teams' actual choices. I find that teams' choices are far more conservative than the ones that would maximize their chances of winning.

Section IV considers various possible complications and biases and finds that none change the basic conclusions. Section V considers the results' quantitative implications. Because the analysis concerns only a small fraction of plays, it implies that different choices on those plays could have only a modest impact on a team's chances of winning. But it also implies that there are circumstances in which teams essentially always kick even though the case for going for it is clear-cut and the benefits of going for it are substantial.

Finally, Section VI discusses the results' broader implications. The hypothesis that firms maximize simple objective functions could fail as a result of either the pursuit of a different, more complex objective function or a failure of maximization. I discuss how either of these possibilities might arise and how one might be able to distinguish between them.[4]

II. The Values of Different Situations

A. *Framework*

The dynamic-programming analysis focuses on 101 situations: a first down and 10 on each yard line from a team's 1 to its opponent's 10, a first and goal on each yard line from the opponent's 9 to its 1, a kickoff from the team's 30 (following a field goal or touchdown, or at the beginning of the game), and a kickoff from its 20 (following a safety). Let V_i denote the value of situation i. Specifically, V_i is the expected long-run value, beginning in situation i, of the difference between the points scored by the team with the ball and its opponent when the two teams are evenly matched, average NFL teams.

[4] Two recent papers that apply economic tools to sports strategy—and in doing so use sports data to test hypotheses about maximization—are the study of serves in tennis by Walker and Wooders (2001) and the study of penalty kicks in soccer by Chiappori, Levitt, and Groseclose (2002). In contrast to this paper, these papers find no evidence of large departures from optimal strategies. Carter and Machol (1971, 1978) and Carroll, Palmer, and Thorn (1998, chap. 10) are more closely related to this paper. I discuss how my analysis is related to these studies below.

By describing the values of situations in terms of expected point differences, I am implicitly assuming that a team that wants to maximize its chances of winning should be risk-neutral over points scored. Although this is clearly not a good assumption late in a game, I show in Section IV that it is an excellent approximation for the early part. For that reason, I focus on the first quarter.

Focusing on the first quarter has a second advantage: it makes it reasonable to neglect effects involving the end of a half. Because play in the second quarter begins at the point where the first quarter ended, the value of a given situation in the first quarter almost certainly does not vary greatly with the time remaining.

Let g index games and t index situations within a game. Let D^i_{gt} be a dummy that equals one if the tth situation in game g is a situation of type i. For example, suppose that $i = 100$ denotes a kickoff from one's 30; then, since all games begin with a kickoff, $D^{100}_{g1} = 1$ for all g and $D^i_{g1} = 0$ for all g and for all $i \neq 100$. Let P_{gt} denote the net points scored by the team with the ball in situation g, t before the next situation. That is, P_{gt} is the number of points scored by the team with the ball minus the number scored by its opponent. Finally, let B_{gt} be a dummy that equals one if the team with the ball in situation g, t also has the ball in situation $g, t+1$ and that equals minus one if the other team has the ball in situation $g, t+1$.

The realized value of situation g, t as of one situation later has two components. The first is the net points the team with the ball scores before the next situation, P_{gt}. The second is the value of the new situation. If the same team has the ball in that situation, this value is simply the V_i corresponding to the new situation. If the other team has the ball, this value is minus the V_i corresponding to the new situation (since the value of a situation to the team without the ball is equal and opposite to the value of the situation to its opponent). In terms of the notation just introduced, the value of situation $g, t+1$ to the team with the ball in situation g, t is $B_{gt}\sum_i D^i_{gt+1}V_i$.

The value of situation g, t as of that situation must equal the expectation of the situation's realized value one situation later. We can write the value of situation g, t as $\sum_i D^i_{gt}V_i$. Thus we have

$$\sum_i D^i_{gt}V_i = E\left[P_{gt} + B_{gt}\sum_i D^i_{gt+1}V_i\right], \tag{1}$$

where the expectation is conditional on situation g, t.

Now define e_{gt} as the difference between the realized value of situation

g, t one situation later and the expectation of the realized value conditional on being in situation g, t:

$$e_{gt} = \left[P_{gt} + B_{gt}\sum_i D^i_{gt}V_i\right] - E\left[P_{gt} + B_{gt}\sum_i D^i_{gt+1}V_i\right].$$

By construction, e_{gt} is uncorrelated with each of the D^i_{gt}'s. If e were correlated with a D^i, this would mean that when teams were in situation i, the realized value one situation later would differ systematically from V_i; but this would contradict the definition of V_i.

Using this definition of e_{gt}, we can rewrite (1) as

$$\sum_i D^i_{gt}V_i = P_{gt} + B_{gt}\sum_i D^i_{gt+1}V_i - e_{gt}, \tag{2}$$

or

$$P_{gt} = \sum_i V_i(D^i_{gt} - B_{gt}D^i_{gt+1}) + e_{gt}. \tag{3}$$

To think about estimating the V_i's, define $X^i_{gt} = D^i_{gt} - B_{gt}D^i_{gt+1}$. Then (3) becomes

$$P_{gt} = \sum_i V_i X^i_{gt} + e_{gt}. \tag{4}$$

This formulation suggests regressing P on the X's. But e may be correlated with the X's. Specifically, e_{gt} is likely to be correlated with the $-B_{gt}D^i_{gt+1}$ terms of the X^i_{gt}'s. Recall, however, that e_{gt} is uncorrelated with the D^i_{gt}'s. Thus the D^i_{gt}'s are legitimate instruments for the X^i_{gt}'s. Further, since they enter into the X^i_{gt}'s, they are almost surely correlated with them. We can therefore estimate (4) by instrumental variables, using the D^i_{gt}'s as the instruments.[5]

There is one final issue. There are 101 V_i's to estimate. Even with a large amount of data, the estimates of the V_i's will be noisy. But the value of a first down is almost certainly a smooth function of a team's position on the field. Thus forcing the estimates of the V_i's to be smooth will improve the precision of the estimates while introducing minimal bias. I therefore require the estimated V_i's to be a quadratic spline as a function of the team's position on the field, with knot points at both 9-, 17-, and 33-yard lines and at the 50. I do not impose any restrictions

[5] There is another way of describing the estimation of the V_i's. Begin with an initial set of V_i's (such as $V_i = 0$ for all i). Now for each i, compute the mean of the realized values of all situations of type i one situation later using the assumed V_i's and the actual P_{gt}'s. Repeat the process using the new V_i's as an input, and iterate until the process converges. One can show that this procedure produces results that are numerically identical to those of the instrumental variables approach.

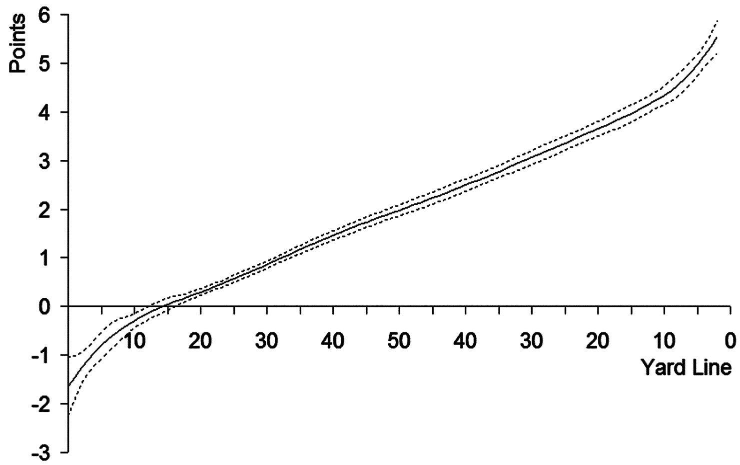


FIG. 1.—The estimated value of situations (solid line) and two-standard-error bands (dotted lines). The estimated value of a kickoff is −0.62 (standard error 0.04); the estimated value of a free kick is –1.21 (standard error 0.51).

on the two estimated V_i's for kickoffs. This reduces the effective number of parameters to be estimated from 101 to 12.[6]

B. *Data and Results*

Play-by-play accounts of virtually all regular-season NFL games for the 1998, 1999, and 2000 seasons were downloaded from the NFL Web site, nfl.com.[7] Since I focus on strategy in the first quarter, I use data only from first quarters to estimate the V_i's. These data yield 11,112 first-quarter situations. By far the most common are a kickoff from one's 30-yard line (1,851 cases) and a first and 10 on one's 20 (557 cases). Because 98.4 percent of extra-point attempts were successful in this period, all touchdowns are counted as 6.984 points.

Figure 1 reports the results of the instrumental variables estimation. It plots the estimated V for a first and 10 (or first and goal) as a function of the team's position on the field, together with the two-standard-error bands. The estimated value of a first and 10 on one's 1-yard line is −1.6 points. The V's rise fairly steeply from the 1, reaching zero at about the 15. That is, the estimates imply that a team should be indifferent between

[6] Carter and Machol (1971) also use a recursive approach to estimate point values of first downs at different positions on the field, using a considerably smaller sample from 1969. There are two main differences from my approach. First, they arbitrarily assign a value of zero to kickoffs and free kicks. Second, they divide the field into 10-yard intervals and estimate the average value for each interval.

[7] Data for two games in 1999 and two games in 2000 were missing from the Web site.

a first and 10 on its 15 and having its opponent in the same situation. The *V*'s increase approximately linearly after the 15, rising a point roughly every 18 yards. The value of a first and 10 equals the value of receiving a kickoff from the 30—0.6 points—around the 27-yard line. That is, receiving a kickoff is on average as valuable as a first and 10 on one's 27. Finally, the *V*'s begin to increase more rapidly around the opponent's 10. The estimated value of a first and goal on the 1 is 5.55 points; this is about the same as the value of an 80 percent chance of a touchdown and a 20 percent chance of a field goal. The *V*'s are estimated relatively precisely: except in the vicinity of the goal lines, their standard errors are less than 0.1.

III. Kicking versus Going for It

This section uses the results of Section II to analyze the choice between kicking and going for it on fourth down. The analysis proceeds in four steps. The first two estimate the values of kicking and going for it in different circumstances. The third compares the two choices to determine which is on average better as a function of the team's position on the field and its distance from a first down. The final step examines teams' actual decisions.

A. Kicking

If one neglects the issue of smoothing the estimates, analyzing the value of kicks is straightforward. To estimate the value of a kick from a particular yard line, one simply averages the realized values of the kicks from that yard line as of the subsequent situation (where "situation" is defined as before). This realized value has two components, the net points scored before the next situation and the next situation's value. In contrast to the previous section, there is no need for instrumental variables estimation.

I constrain the estimated values of kicks to be smooth in the same way as before, with one modification. Teams' choices between punting and attempting a field goal change rapidly around their opponents' 35-yard line. Since one would expect the level but not the slope of the value of kicking as a function of the yard line to be continuous where teams switch from punts to field goal attempts, I do not impose the slope restriction at the opponent's 33. And indeed, the estimates reveal a substantial kink at this knot point.

The data consist of all kicks in the first quarters of games. Since what we need to know is the value of deciding to kick, I include not just actual punts and field goal attempts, but blocked and muffed kicks and kicks nullified by penalties. There are 2,560 observations.[8]

The results are reported in figure 2. Figure 2*a* shows the estimated value of kicking as a function of the team's position on the field. Figure 2*b* plots the *difference* between the estimated value of a kick and of the other team having a first down on the spot. From the team's 10-yard line to midfield, this difference is fairly steady at around 2.1 points, which corresponds to a punt of about 38 yards. It dips down in the "dead zone" around the opponent's 35-yard line, where a field goal is unlikely to succeed and a punt is likely to produce little yardage. It reaches a low of 1.5 (a punt of only 25 yards) at the 33 and then rises to 2.2 at the 21. As the team gets closer to the goal line, the probability of a successful field goal rises little, but the value of leaving the opponent with the ball rises considerably. The difference between the values of kicking and of the opponent receiving the ball therefore falls, reaching 0.7 at the 1. The estimates are relatively precise: the standard error of the difference in values is typically about 0.1.[9]

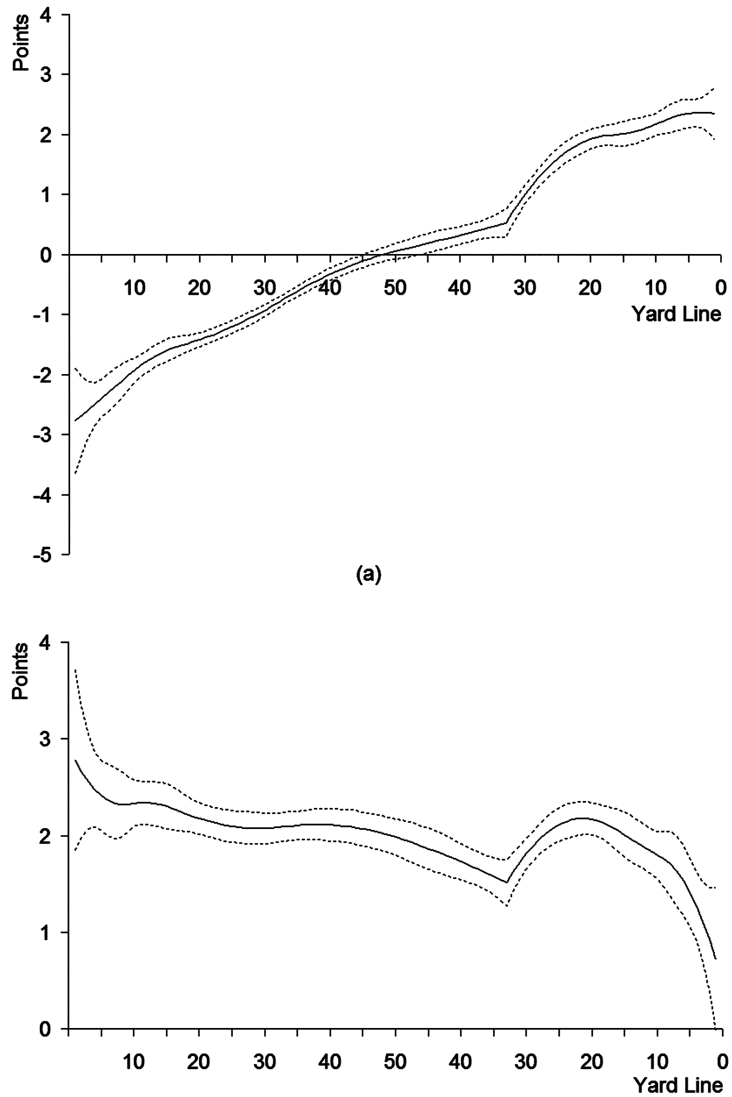
B. Going for It

The analysis of the value of trying for a first down or touchdown parallels the analysis of kicking. There are two differences. First, because teams rarely go for it on fourth down, I use third-down plays instead. That is, I find what third-down plays' realized values as of the next situation would have been if the plays had taken place on fourth down.

Second, the value of going for it depends not only on the team's position on the field, but also on the number of yards to go for a first down or touchdown. If there were no need to smooth the estimates,

[8] There are several minor issues involving the data. First, fourth-down plays that are blown dead before the snap and for which the play-by-play account does not say whether the kicking squad was sent in are excluded. Since such plays are also excluded from the analysis of the decision to go for it, this exclusion should generate little bias. Second, it is not clear whether fake kicks should be included; it depends on whether one wants to estimate the value of deciding to kick or the value of lining up to kick. There are only five fake kicks in the sample, however, and the results are virtually unaffected by whether they are included. The results in the text include fakes. Finally, since teams occasionally obtain first downs on kicking plays (primarily through penalties), the value of a kick is affected by the number of yards the team has to go for a first down. But there are only six kicking plays in the sample on which the team had 5 yards to go or less and moved the ball 5 yards or less and obtained a first down. Thus to improve the precision of the estimates, I do not let the estimated value of kicks vary with the number of yards needed for a first down.

[9] The standard errors account for the fact that the V_i's used to estimate the values of kicks are themselves estimated. This calculation is performed under the assumption that the differences between the realized and expected values of kicks are uncorrelated with the errors in estimating the V_i's. Although this assumption will not be strictly correct, it is almost certainly an excellent approximation.



(b)

Fig. 2.—*a*, The estimated value of kicks. *b*, The estimated value of the difference between the values of kicks and of turning the ball over. The dotted lines show the two-standard-error bands.

one could use averages to estimate the value of going for it for a specific position and number of yards to go. That is, one could consider all cases in which the corresponding circumstance occurred on third down, find what the plays' realized values would have been if they had been fourth-down plays, and average the values. In fact, however, there are over a thousand different cases in the sample. Smoothing the estimates is therefore essential.

To smooth the estimates, I focus on the difference between the values of going for it and of turning the ball over on the spot rather than estimating the value of going for it directly. In general, this difference depends on three factors. The first is the difference between the values of having a first down on the spot and of the other team having a first down there. Since the V's are essentially symmetric around the 50-yard line, this factor is essentially independent of the team's position on the field. The second factor is the probability that the team succeeds when it goes for it. As long as the team is not close to its opponent's goal line, there is no reason for this probability to vary greatly with the team's position. The third (and least important) factor is the average additional benefit from the yards the team gains when it goes for it. Again, as long as the team is not close to the opponent's goal line, there is no reason for this factor to vary substantially with its position.

Close to the opponent's goal line, however, the team has less room to work with, and so its chances of success and average number of yards gained are likely to be lower. On the other hand, because the value of a touchdown is much larger than the value of a first down on the 1, the additional benefit from gaining yards may be higher. Thus near the goal line, we cannot be confident that the difference between the values of going for it and of turning the ball over does not vary substantially with the team's position.

The difference between the values of going for it and of turning the ball over on the spot is $G_{iy} - (-V_{i'})$, or $G_{iy} + V_{i'}$, where G_{iy} denotes the value of going for it on yard line i with y yards to go and i' denotes the yard line "opposite" yard line i. From the team's goal line to the opponent's 17, I assume that this difference is independent of i and quadratic in y:

$$G_{iy} + V_{i'} = a_0 + a_1 y + a_2 y^2. \tag{5}$$

From the opponent's 17 to its goal line, I let the difference depend quadratically on both i and y:

$$G_{iy} + V_{i'} = b_0 + b_1 y + b_2 i + b_3 y^2 + b_4 yi + b_5 i^2 + b_6 y^2 i + b_7 yi^2 + b_8 y^2 i^2. \tag{6}$$

At the 17, where the two functions meet, I constrain both their level

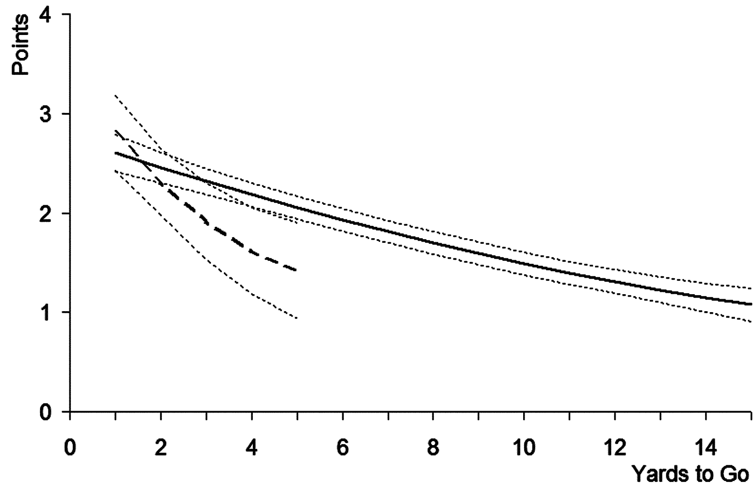


FIG. 3.—The estimated difference between the values of going for it and of the other team having the ball on the spot at a generic yard line outside the opponent's 17 (solid line) and at the opponent's 5 (dashed line). The dotted lines show the two-standard-error bands.

and their derivative with respect to i to be equal for all y. This creates six restrictions.

The data consist of all third-down plays in the first quarter; there are 4,733 observations.[10] Figure 3 summarizes the results. The solid line shows the estimates of $G_{iy} + V_{i'}$ as a function of y for a generic position on the field not inside the opponent's 17, and the dashed line shows the estimates at the opponent's 5. Outside the opponent's 17, the estimate of $G_{iy} + V_{i'}$ for a team facing fourth and 1 is 2.64. On third-and-1 plays from the goal line to the opponent's 17, teams are successful 64 percent of the time, and they gain an average of 3.8 yards; this corresponds to an expected value of 2.66 points.[11] Thus the estimate of 2.64 is reasonable. The estimated difference falls roughly linearly with the number of yards to go. It is 2.05 with 5 yards to go (equivalent to a 45 percent chance of success and an average gain of 6.3 yards), 1.49 with 10 yards to go (a 30 percent chance of success and an average gain of 6.6 yards), and 1.08 with 15 yards to go (an 18 percent chance of

[10] To parallel the analysis of kicking, plays that are blown dead before the snap for which it would not have been possible to determine whether the kicking team had been sent in are excluded (see n. 8). And to prevent outliers that are not relevant to decisions about going for it from affecting the results, plays on which the team had more than 20 yards to go are excluded.

[11] The translations of average outcomes into point values in this paragraph are done for a team at midfield. Since the V's are not exactly symmetric around the 50 or exactly linear, choosing a different position would change the calculations slightly.

success and an average gain of 7.7 yards). These estimates are similar to what one would obtain simply by looking at the average results of the corresponding types of plays.

At the opponent's 5, the estimate of $G_{iy} + V_{i'}$ with 1 yard to go is 2.94 (equivalent to a 38 percent chance of a first down with an average gain of 2 yards plus a 25 percent chance of a touchdown), which is slightly higher than the estimate elsewhere on the field. The estimate falls more rapidly with the number of yards to go than elsewhere on the field, however. With 5 yards to go, it is 1.42 (equivalent to a 26 percent chance of a touchdown). The estimate for 5 yards to go is quite similar to what one would obtain by looking at averages; the estimate for 1 yard to go is somewhat higher, however.

The dotted lines show the two-standard-error bands. For the range in which $G_{iy} + V_{i'}$ is constrained to be independent of i, the standard errors are small: for 15 yards to go or less, they are less than 0.1. Inside the 17, where fewer observations are being used, they are larger, but still typically less than 0.2.

C. *Recommended Choices*

Figure 4 combines the analyses of kicking and going for it by showing the number of yards to go where the estimated average payoffs to the two choices are equal as a function of the team's position. On the team's own half of the field, going for it is better on average if there is less than about 4 yards to go. After midfield, the gain from kicking falls, and so the critical value rises. It is 6.5 yards at the opponent's 45 and peaks at 9.8 on the opponent's 33. As the team gets into field goal range, the critical value falls rapidly; its lowest point is 4.0 yards on the 21. Thereafter, the value of kicking changes little while the value of going for it rises. As a result, the critical value rises again. The analysis implies that once a team reaches its opponent's 5, it is always better off on average going for it. The two dotted lines in the figure show the two-standard-error bands for the critical values.[12] The critical values are estimated fairly precisely.

Although these findings contradict the conventional wisdom, they are quite intuitive. As described in Section I, one case for which the intuition is clear is fourth and goal on the 2. The expected payoffs in terms of immediate points to the two choices are very similar, but trying for a touchdown on average leaves the other team in considerably worse field position. Another fairly intuitive case is fourth and 3 or 4 on the 50. If the team goes for a first down, it has about a 50-50 chance of success;

[12] For example, the lower dotted line shows the point where the difference between the estimated values of going for it and kicking is twice its standard error.

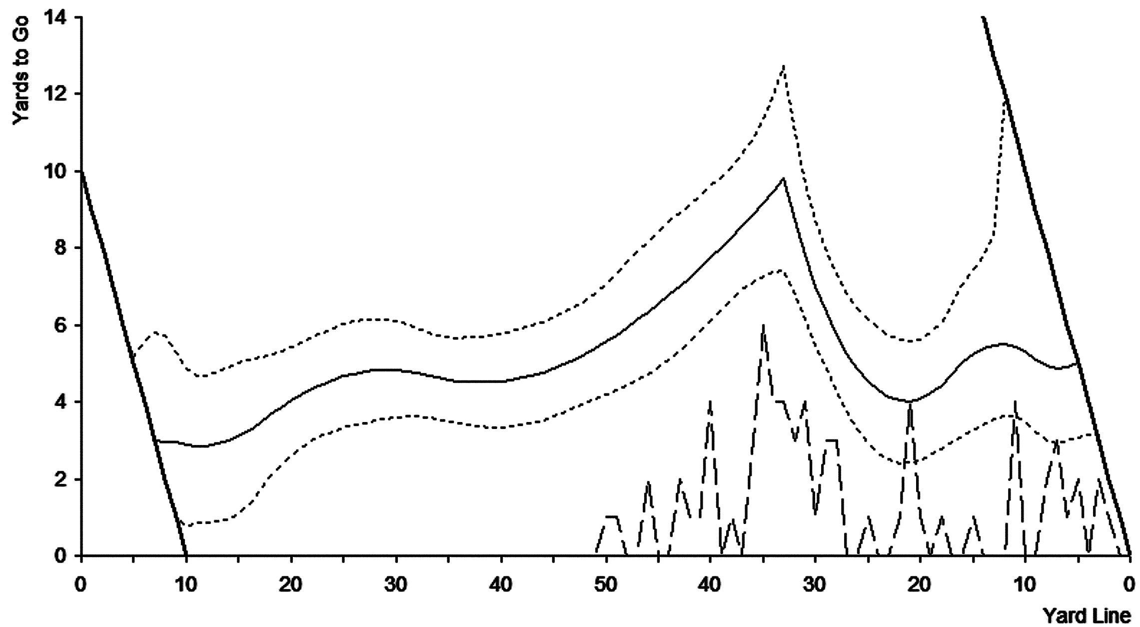


Fig. 4.—The number of yards to go where the estimated values of kicking and going for it are equal (solid line) and two-standard-error bands (dotted lines), and the greatest number of yards to go such that when teams have that many yards to go or less, they go for it at least as often as they kick (dashed line).

thus both the team and its opponent have about a 50 percent chance of a first and 10. But the team will gain an average of about 6 yards on the fourth-down play; thus on average it is better off than its opponent if it goes for it. If the team punts, its opponent on average will end up with a first and 10 around its 14. Both standard views about football and the analysis in Section II suggest that the team and its opponent are about equally well off in this situation. Thus, on average the team is better off than its opponent if it goes for a first down, but not if it punts. Going for the first down is therefore preferable on average.

The very high critical values in the dead zone also have an intuitive explanation. The chances of making a first down decline only moderately as the number of yards to go increases. For example, away from the opponent's end zone, the chance of making a first down or touchdown on third down is 64 percent with 1 yard to go, 44 percent with 5 yards to go, and 34 percent with 10 yards to go. As a result, the large decrease in the gain from kicking in the dead zone causes a large increase in the critical value.

D. *Actual Choices*

Teams' actual choices are dramatically more conservative than those recommended by the dynamic-programming analysis. On the 1,604 fourth downs in the sample for which the analysis implies that teams are on average better off kicking, they went for it only nine times. But on the 1,068 fourth downs for which the analysis implies that teams are on average better off going for it, they kicked 959 times.[13]

The dashed line in figure 4 summarizes teams' choices. It shows, for each point on the field, the largest number of yards to go with the property that when teams have that many yards to go or less, they go for it at least as often as they kick. Over most of the field, teams usually kick even with only 1 yard to go. Teams are slightly more aggressive in the dead zone, but are still far less aggressive than the dynamic-programming analysis suggests. On the line summarizing teams' choices, the null hypothesis that the average values of kicking and going for it are equal is typically rejected with a *t*-statistic between three and seven.[14]

[13] These figures exclude the 28 cases for which we cannot observe the team's intent because of a penalty before the snap.

[14] Carter and Machol (1978) and Carroll et al. (1998, chap. 10) also examine fourth-down decisions. Carter and Machol consider only decisions inside the opponent's 35-yard line. They use estimates from their earlier work (described in n. 6 above) to assign values to different situations. To estimate the payoff from going for it, they pool third-down and fourth-down plays. They assume that all successful plays produce exactly the yards needed for a first down, that all unsuccessful plays produce no yards, and that the probability of success does not depend on the team's position on the field. They then compare the estimated payoffs to going for it with the payoffs to field goal attempts and punts. They

IV. Complications

A. *Rational Risk Aversion*

I have assumed that a win-maximizing team should be risk-neutral concerning points scored. This is clearly not exactly correct. The analysis may therefore overstate the value of a touchdown relative to a field goal, and thus overstate the benefits of going for it on fourth down.

Three considerations suggest that this effect is not important. First, as I show in Section V, teams are conservative even in situations in which win-maximizing behavior would be risk-loving over points scored. Second, it is essentially irrelevant to decisions in the middle of the field. Near midfield, a team should maximize the probability that it is the first to get close to the opponent's goal line, since that is necessary for either a field goal or a touchdown. But teams are conservative over the entire field.

Third, direct evidence about the impact of points on the probability of winning suggests that risk neutrality is an excellent approximation for the early part of the game. Because teams adjust their play late in the game on the basis of the score, one cannot just look at the distribution of actual winning margins. Instead, I try to approximate what the distribution of winning margins would be in the absence of late-game adjustments and use this to estimate the value of a field goal or touchdown early in the game. I begin by dividing the games into deciles according to the point spread. I then find the score for the favorite and the underdog at the end of the first half; the idea here is that these scores are relatively unaffected by adjustments in response to the score. I then construct synthetic final scores by combining the first-half scores of each pair of games within a decile. This yields a total of 74(73)/2 or 73(72)/2 synthetic games for each decile, for a total of 26,718 observations. I use the results to estimate the impact of an additional field goal or touchdown in the first quarter. For example, the estimated effect of a field goal on the probability of winning is the sum of the probability that a team would trail by 1 or 2 points at the end of the game plus half the probability that the score would be tied or the team would trail by 3 points.

---

conclude that teams should be considerably more aggressive than they are. Carroll et al. consider decisions over the entire field. They do not spell out their method for estimating the values of different situations (though it appears related to Carter and Machol's), and it yields implausible results. Similarly to Carter and Machol, they pool third-down and fourth-down plays and assume that successful plays produce one more yard than needed for a first down, that unsuccessful plays yield no gain, and that the chances of success do not vary with field position. They again conclude that teams should be considerably more aggressive. Their specific findings about when going for it is preferable on average are quite different from mine, however. Finally, neither Carter and Machol nor Carroll et al. investigate the statistical significance of their results.

This exercise suggests that 7 points are in fact slightly *more* than seven-thirds as valuable as 3. An additional 3 points are estimated to raise the probability of winning by 6.8 percentage points; an additional 7 points are estimated to raise the probability by 16.2 percentage points, or 2.40 times as much. The source of this result is that the distribution of synthetic margins is considerably higher at 4 and 7 points than at 1 or 2. To put it differently, to some extent what is important about a touchdown is not that its usual value is 7 points, but that its usual value is between two and three times the value of a field goal.

B. Third Down versus Fourth Down

There are two ways to investigate the appropriateness of using third-down plays to gauge what would happen if teams went for it on fourth down. The first is to consider how teams' incentives are likely to affect outcomes on fourth downs relative to third downs. Relative payoffs to different outcomes are different on the two downs. In particular, the benefit from a long gain relative to just making a first down is smaller on fourth down. As a result, both the offense and defense will behave differently: the offense will be willing to lower its chances of making a long gain in order to increase its chances of just making a first down, and the defense will be willing to do the reverse.

This suggests that the direction of the bias from using third-down plays should depend on which team has more influence on the distribution of outcomes. Since it seems unlikely that the defense has substantially more influence than the offense on the distribution of outcomes, it follows that the use of third downs is unlikely to lead to substantial overestimates of the value of going for it.

More important, the relative payoffs to different outcomes do not differ greatly between third and fourth downs. For example, consider a team that is on its 30 and needs 2 yards for a first down. On third down (under the realistic assumption that the team will punt if it fails to make a first down), the benefit of gaining 15 yards rather than none is 1.4 times as large as the benefit of gaining 2 yards rather than none. On fourth down, the benefit of gaining 15 yards rather than none is 1.2 times as large as the benefit of gaining 2 yards rather than none. Thus one would not expect either side to behave very differently on the two downs. And when a team has goal to go, the payoff on either third down or fourth down depends almost entirely on whether the team scores a touchdown. Thus one would expect both sides' behavior to be essentially the same on the two downs. These considerations suggest that any bias from the use of third-down plays is likely to be small.

The second approach is to directly compare the realized values of plays where teams went for it on fourth downs (i.e., the immediate points scored plus the value of the resulting field position) with what one would expect on the basis of the analysis of third downs. This comparison is potentially problematic, however, for two reasons. First, as described above, teams went for it only 118 times in the sample. Second, times when teams choose to go for it may be unusual: the teams may know that they are particularly likely to succeed, or they may be desperate.

To increase the sample without bringing in fourth-down attempts that are likely to be especially unusual, I include the entire game except for the last two minutes of each half (and overtimes). This increases the sample to 1,338 plays. And as a partial remedy for the second problem, I experiment with controlling for the amount the team with the ball is trailing by and the amount it is favored by.

The results suggest that fourth downs are virtually indistinguishable from third downs. The mean of the difference between the realized value of the fourth-down attempts and what is predicted by the analysis of third downs is 0.006 (with a standard error of 0.7), which is essentially zero. When controls for the prior point spread and the current point differential are included, the coefficient falls to -0.042 and remains highly insignificant. The point estimate corresponds to the probability of success being one percentage point lower on fourth downs than on third downs, which would have almost no impact on the analysis.

C. *Additional Information*

In making fourth-down decisions, a team has more information than the averages used in the dynamic-programming analysis. Thus it would not be optimal for it to follow the recommendations of the dynamic-programming analysis mechanically.

Additional information cannot, however, account for the large systematic departures from the recommendations of the dynamic-programming analysis. Over wide ranges, teams almost always kick in circumstances in which the analysis implies that they would be better off on average going for it. For example, on the 512 fourth downs in the sample in the offense's half of the field for which the dynamic-programming analysis suggests going for it, teams went for it only seven times. Similarly, on the 175 fourth downs with 5 or more yards to go for which the analysis suggests going for it, teams went for it only 13 times.

Additional information can account for this behavior only if teams know on a large majority of fourth downs that the expected payoff to going for it relative to kicking is considerably less than average, and know on the remainder that the expected payoff is dramatically larger than average. This possibility is not at all plausible. Further, it predicts that when teams choose to go for it, the results will be far better than one would expect on the basis of averages. As described above, this prediction is contradicted by the data.

D. Momentum

Failing on fourth down could be costly to a team's chances of winning not just through its effect on possession and field position, but also through its effect on energy and emotions. Thus it might be more costly for the other team to have the ball as a result of a failed fourth-down attempt than for it to have the ball at the same place in the course of a normal drive or because of a punt. The analysis might therefore overstate the average payoff to going for it.

There are two reasons to be skeptical of this possibility. First, the same reasoning suggests that there could be a motivational benefit to succeeding on fourth down, and thus that the analysis could understate the benefits of a successful fourth-down attempt. Second, studies of momentum in other sports have found at most small momentum effects (e.g., Gilovich, Vallone, and Tversky 1985; Albright 1993; Klaassen and Magnus 2001).

More important, it is possible to obtain direct evidence about whether outcomes differ systematically from normal after plays whose outcomes are either very bad or very good. To obtain a reasonable sample size, for very bad plays I consider all cases in which from one situation to the next (where a situation is defined as before), possession changed and the ball advanced less than 10 yards. For very good plays, I consider all cases in which the offense scored a touchdown. These criteria yield 636 very bad plays and 628 very good plays. I then examine what happens from the situation immediately following the extreme play to the next situation, from that situation to the next, and from that situation to the subsequent one. In each case, I ask whether the realized values of these situations one situation later differ systematically from the V's for those situations. That is, I look at the means of the relevant e_{gt}'s (always computed from the perspective of the team that had the ball before the very bad or very good play).

The results provide no evidence of momentum effects. All the point estimates are small and highly insignificant; the largest t-statistic (in absolute value) is less than 1.3. Moreover, the largest point estimate (again in absolute value) goes the wrong direction from the point of view of the momentum hypothesis: from the situation immediately following a very bad play to the next, the team that lost possession does somewhat better than average.[15]

[15] The working paper version of the paper (Romer 2005) considers two additional complications. The first is the possibility of sample selection bias in the estimation of the V's

V. Quantitative Implications

An obvious question is whether the potential gains from different choices are important. There are in fact two distinct questions. The first is whether there are cases of clear-cut departures from win maximization. If there were not, then small changes in the analysis might reverse the conclusions.

The answer is that there are clear-cut departures. One example is the case of fourth and goal on the 2 discussed above. The estimates imply that trying for a touchdown and failing is only slightly worse than kicking a field goal. As a result, they imply that going for a touchdown is preferable on average as long as the probability of success is at least 18 percent. The actual probability of success, in contrast, is about 45 percent. Thus there are no plausible changes in the analysis that could reverse the conclusion that trying for a touchdown is preferable on average. Moreover, the average benefit of trying for a touchdown is substantial. The estimated value of going for it is about 3.7 points, whereas the estimated value of kicking is about 2.4 points. Since each additional point raises the probability of winning by about 2.3 percentage points, trying for a touchdown on average increases the chances of winning by about three percentage points. Yet teams attempted a field goal every time in the sample they were in this position.

Two other examples are fourth and goal on the 1 and fourth and 1 between the opponent's 35 and 40. For the first, the estimates imply that the critical and actual probabilities of success are 16 percent and 62 percent, and that trying for a touchdown on average increases the chances of winning by about five percentage points. For the second, the critical and actual probabilities are 39 percent and 64 percent, and going for a first down raises the probability of winning by about 2.5 percentage points. In these cases, teams do not always kick, but they do about half the time. These decisions are consistent with win maximization only if teams have substantial additional information that allows them to identify times when their fourth-down attempts are especially likely to succeed. As described in the previous section, there is no evidence of such large additional information.

The second question is whether the analysis implies that teams could increase their overall chances of winning substantially. Since the analysis considers only a small fraction of plays and only a single decision on those plays, one would not expect it to show large potential increases

stemming from the fact that teams are not assigned to situations randomly. The second is general equilibrium effects: different decisions on fourth downs could affect other choices. I conclude that the effects of sample selection bias are small and of ambiguous sign, and that general equilibrium effects are small and most likely strengthen the case for being more aggressive on fourth downs.

in the chances of winning. And indeed, the potential gains are small. In the 732 first quarters in the sample, there are 959 cases in which a team kicked when the difference between the estimated values of going for it and kicking was positive. The average estimated value of the expected gain from going for it in these cases is 0.35 points. Thus the expected payoff to a typical team of being more aggressive on fourth downs in the first quarter is approximately 0.23 points per game, which corresponds to an increase in the probability of winning of about one-half of a percentage point.

Teams could also benefit by being more aggressive on fourth downs in the remaining quarters. A full-fledged analysis of fourth-down decisions over the entire game would require accounting for the score and the time remaining, which would complicate the analysis enormously. Nonetheless, the evidence is clear. Consider first all fourth-down plays in the second, third, and fourth quarters. On the 9,233 such plays, the analysis of decisions in the first quarter suggests that going for it is preferable on average 3,555 times, yet teams went for it only 1,426 times.[16] That is, teams are almost as conservative over the last three quarters as they are in the first. But there is no reason to think that the average benefits to various outcomes are much different in the later quarters.

Stronger evidence comes focusing on cases in which win maximization implies that teams should be risk-loving over points scored (and in which they are not so far behind that they might reasonably view the game as unwinnable). Specifically, I consider fourth downs in the second quarter when the team with the ball is trailing by at least 4 points, in the third quarter when it is trailing by between 4 and 28 points, and in the fourth quarter when it is trailing by between 4 and 16 points. In the 3,065 such cases, the first-quarter analysis suggests going for it 1,147 times, but teams went for it only 596 times. That is, in cases in which win-maximizing behavior is risk-loving over points scored, teams are considerably more conservative than they would be if they were risk-neutral over points.

This evidence suggests that a rough estimate of the potential gains over the whole game is four times the gains from the first quarter, or an increase of about 2.1 percentage points in the probability of winning. Since an NFL season is 16 games long, this corresponds to slightly more than one additional win every three seasons. This is a modest (though not trivial) effect. Thus one cannot rule out the possibility that I have merely identified a clear-cut but modest and isolated departure from

[16] As before, these figures (and those in the next paragraph) exclude cases for which it is not possible to determine the team's intent because of a penalty before the snap.

maximization. Because I have examined one particular type of decision in detail, there is simply no evidence either for or against this hypothesis.

VI. Conclusion

This paper shows that the behavior of National Football League teams on fourth downs departs from the behavior that would maximize their chances of winning in a way that is highly systematic, clear-cut, and statistically significant. This is true even though the decisions are comparatively simple, the possibilities for learning and imitation are unusually large, the compensation for the coaches who make the decisions is extremely high, and the market for their services is intensively competitive. Despite these forces, the standard assumption that agents maximize simple objective functions fails to lead to reasonably accurate descriptions of behavior.

The departures from win maximization are toward "conservative" behavior: the immediate payoff to a punt or field goal attempt has a lower variance than the immediate payoff to going for it. Nonetheless, conventional risk aversion cannot explain the results. At the end of the game, one team will have won and the other will have lost. Thus even a decision maker who faces a large cost of losing and little benefit of winning should maximize the probability of winning.

At a broad level, two forces could lead to departures from the maximization of simple objective functions. First, the relevant actors could have more complicated objective functions. In the context of the decisions considered in this paper, the natural possibility is that the actors care not just about winning and losing, but about the probability of winning during the game, and that they are risk-averse over this probability. That is, they may value decreases in the chances of winning from failed gambles and increases from successful gambles asymmetrically. If such risk aversion comes from fans (and if it affects their demand), teams' choices would be departures from win maximization but not from profit maximization. If it comes from owners, they would be forgoing some profits to obtain something else they value. And if it comes from coaches and players, teams' choices could again be profit-maximizing (if coaches and players are willing to accept lower compensation to follow more conservative strategies), or they could be the result of agency problems. But even if the departures from win maximization reflect the pursuit of a more complex objective function—and, indeed, even if they reflect profit maximization through subtle channels—the results would still support the view that the determinants of firm behavior cannot be derived a priori but must be determined empirically.

The second broad possibility would involve an even more significant departure from standard models: perhaps the decision makers are systematically imperfect maximizers. Many skills are more important to running a football team than a command of mathematical and statistical tools. And it would hardly be obvious to someone without knowledge of those tools that they could have any significant value in football. Thus the decision makers may want to maximize their teams' chances of winning, but rely on experience and intuition rather than formal analysis. And because they are risk-averse in other contexts, experience and intuition may lead them to behave more conservatively than is appropriate for maximizing their chances of winning.[17]

The experimental and behavioral literatures have documented many aspects of behavior that are systematically more conservative than standard models predict. The classic "Ellsberg paradox" (Ellsberg 1961) and later research (e.g., Hogarth and Kunreuther 1989) show that individuals typically act as though they are risk-averse over probabilities when probabilities are ambiguous. Selten, Sadrieh, and Abbink (1999) show that when subjects face choices among lotteries whose payoffs are themselves lottery tickets, they often exhibit strong risk aversion over probabilities. Rabin (2000) shows that individuals tend to make conservative decisions concerning monetary gambles in ways that cannot be rationalized by any plausible degree of risk aversion. And both individuals and firms tend to view risky decisions in isolation and to exhibit risk aversion regarding them even when their implications for the risk of the individual's wealth or the firm's profits are minimal (e.g., Kahneman and Lovallo 1993; Read, Loewenstein, and Rabin 1999).

Much of the previous evidence of systematically conservative behavior involves highly stylized laboratory settings with small stakes and inexperienced decision makers devoting relatively little effort to their choices. Thus previous work provides little evidence about the strength of the forces pushing decision makers toward conservatism. The results of this paper suggest that the forces may be shockingly strong.

Unfortunately, there is little evidence about whether conservative behaviors arise because individuals have nonstandard objective functions or because they are imperfect maximizers. For example, individuals may exhibit risk aversion over probabilities either because they genuinely dislike uncertainty about probabilities or because they misapply their usual rules of thumb to settings where risk involves probabilities rather than payoffs. Similarly, as Read et al. observe, individuals may choose to forgo a sequence of gambles that is virtually certain to have a positive total payoff either because the expected utility from the eventual payoff

[17] In addition, herding (e.g., Scharfstein and Stein 1990) could magnify departures from win maximization: if coaches who deviate from standard practice are punished more for failures than they are rewarded for successes, departures from win maximization will be self-reinforcing. But herding cannot explain why the departures are in one particular direction.

is not enough to compensate them for the disutility they would suffer from the many small setbacks along the way, or because they do not understand how favorable the distribution of final outcomes would be. And as described above, the departures from win maximization in football could also arise from either source.

The hypotheses of nonstandard objective functions and imperfect optimization do, however, make different predictions about the future evolution of football strategy. If conservative choices stem from preferences concerning the probability of winning during the game, behavior will not change. But if they stem from imperfect optimization, then trial and error, increased availability of data, greater computing power, and the development of formal analyses of strategy will cause behavior to move toward victory-maximizing choices. Thus the future evolution of football strategy will provide evidence about the merits of these two competing explanations of systematic departures from the predictions of models of complete optimization of simple objective functions.

Appendix

This appendix describes the main rules of football that are relevant to the paper.

A football field is 100 yards long. Each team defends its own goal line and attempts to move the ball toward its opponent's. The yard lines are numbered starting at each goal line and are referred to according to which team's goal line they are closer to. Thus, for example, the yard line 20 yards from one team's goal line is referred to as that team's 20-yard line.

The game begins with a kickoff: one team puts the ball in play by kicking the ball from its own 30-yard line to the other team. After the kickoff, the team with the ball has four plays, or downs, to move the ball 10 yards. If at any point it gains the 10 yards, it begins a new set of four downs. Plays are referred to by the down, number of yards to go for a first down, and location. For example, suppose that the receiving team returns the opening kickoff to its 25-yard line. Then it has first and 10 on its own 25. If it advances the ball 5 yards on the first play, it has second and 5 on its own 30. If it advances 8 yards on the next play (for a total of 13), it now has first and 10 on its own 38. The team with the ball is referred to as the offense, the other team as the defense.

If a team advances the ball across its opponent's goal line, it scores a touchdown. A touchdown gives the team 6 points and an opportunity to try for an extra point, which almost always produces 1 point. If a team has a first and 10 within 10 yards of its opponent's goal line, it cannot advance 10 yards without scoring a touchdown. In this case, the team is said to have first and goal rather than first and 10.

On fourth down, the offense has three choices. First, it can attempt a conventional play. If the play fails to produce a first down or touchdown, the defense gets a first down where the play ends. Second, it can kick (or "punt") the ball to the defense; this usually gives the defense a first down, but at a less advantageous point on the field. Third, it can attempt to kick the ball through the uprights located 10 yards behind the opponent's goal line (a "field goal"). If it succeeds, it scores 3 points. If it fails, the defense gets a first down at the point where the kick was made, which is normally 8 yards farther from its goal line than the play started. (If the field goal was attempted from less than 20 yards from the goal line, however, the defense gets a first down on its 20-yard line rather than at the point of the attempt.) After either a touchdown or a field goal, the scoring team kicks off from its 30-yard line, as at the beginning of the game.

The final (and by far the least common) way to score is a safety: if the offense is pushed back across its own goal line, the defense scores 2 points, and the offense puts the ball in play by kicking to the other team from its 20-yard line (a "free kick").

The game is divided into four 15-minute periods. At the beginnings of the second and fourth quarters, play continues at the point where it left off. At the beginning of the third quarter, however, play begins afresh with a kickoff by the team that did not kick off at the beginning of the game.

References

Albright, S. Christian. 1993. "A Statistical Analysis of Hitting Streaks in Baseball." *J. American Statis. Assoc.* 88 (December): 1175–83.

Alchian, Armen A. 1950. "Uncertainty, Evolution, and Economic Theory." *J.P.E.* 58 (June): 211–21.

Becker, Gary S. 1957. *The Economics of Discrimination.* Chicago: Univ. Chicago Press.

Carroll, Bob, Pete Palmer, and John Thorn, with David Pietrusza. 1998. *The Hidden Game of Football: The Next Edition.* New York: Total Sports.

Carter, Virgil, and Robert E. Machol. 1971. "Operations Research on Football." *Operations Res.* 19 (March–April): 541–44.

———. 1978. "Optimal Strategies on Fourth Down." *Management Sci.* 24 (December): 1758–62.

Chiappori, Pierre-André, Steven D. Levitt, and Timothy Groseclose. 2002. "Testing Mixed-Strategy Equilibria When Players Are Heterogeneous: The Case of Penalty Kicks in Soccer." *A.E.R.* 92 (September): 1138–51.

Ellsberg, Daniel. 1961. "Risk, Ambiguity, and the Savage Axioms." *Q.J.E.* 75 (November): 643–69.

Fama, Eugene F. 1980. "Agency Problems and the Theory of the Firm." *J.P.E.* 88 (April): 288–307.

Friedman, Milton. 1953. "The Methodology of Positive Economics." In *Essays in Positive Economics.* Chicago: Univ. Chicago Press.

Gilovich, Thomas, Robert Vallone, and Amos Tversky. 1985. "The Hot Hand in Basketball: On the Misperception of Random Sequences." *Cognitive Psychology* 17 (July): 295–314.

Hogarth, Robin M., and Howard Kunreuther. 1989. "Risk, Ambiguity, and Insurance." *J. Risk and Uncertainty* 2 (April): 5–35.

Kahneman, Daniel, and Dan Lovallo. 1993. "Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking." *Management Sci.* 39 (January): 17–31.

Klaassen, Franc J. G. M., and Jan R. Magnus. 2001. "Are Points in Tennis Independent and Identically Distributed? Evidence from a Dynamic Binary Panel Data Model." *J. American Statis. Assoc.* 96 (June): 500–509.

Rabin, Matthew. 2000. "Risk Aversion and Expected-Utility Theory: A Calibration Theorem." *Econometrica* 68 (September): 1281–92.

Read, Daniel, George Loewenstein, and Matthew Rabin. 1999. "Choice Bracketing." *J. Risk and Uncertainty* 19 (December): 171–97.
Romer, David. 2005. "Do Firms Maximize? Evidence from Professional Football." Working paper (July), Univ. California, Berkeley.
Scharfstein, David S., and Jeremy C. Stein. 1990. "Herd Behavior and Investment." *A.E.R.* 80 (June): 465–79.
Selten, Reinhard, Abdolkarim Sadrieh, and Klaus Abbink. 1999. "Money Does Not Induce Risk Neutral Behavior, but Binary Lotteries Do Even Worse." *Theory and Decision* 46 (June): 211–49.
Thaler, Richard H. 2000. "Sudden Death Aversion." Working paper (November), Grad. School Bus., Univ. Chicago.
Walker, Mark, and John Wooders. 2001. "Minimax Play at Wimbledon." *A.E.R.* 91 (December): 1521–38.