Overreaction in Football Wagers

Gary Smith
Department of Economics
Pomona College
Corresponding author:
Gary Smith
Department of Economics
Pomona College
Claremont, California 91711
gsmith@pomona.edu

Andrew Capron
Department of Economics
Pomona College
Overreaction in Football Wagers

John Maynard Keynes\(^1\) observed that “day-to-day fluctuations in the profits of existing investments, which are obviously of an ephemeral and nonsignificant character, tend to have an altogether excessive, and even absurd, influence on the [stock] market.” This overreaction can be explained by investors’ inadequate appreciation of the importance of random events on corporate earnings, and the subsequent regression of earnings towards the mean.

We investigate whether this is also true of wagering on professional football games. Do bettors overreact to weekly score fluctuations and not fully anticipate the subsequent regression of performances towards the mean?

**Mean Reversion and Regression to the Mean**

A compelling body of evidence has accumulated supporting Keynes’ observation about overreaction in the stock market. For example, Keil, Smith, and Smith\(^2\) and Smith\(^3\) found that stock analyst predictions of earnings growth rates can be improved by shrinking predicted growth rates toward the average forecast. DeBondt and Thaler\(^4,5\) found that U. S. stocks that had done relatively well in the past (“winners”) typically do worse subsequently than do stocks that had done poorly (“losers”). Fama and French\(^6\) and Poterba and Summers\(^7\) also found mean reversion in stock returns in that stocks with above-average returns over various horizons tend to experience below-average returns subsequently.

Mean reversion in stock prices may be due to an insufficient appreciation of the statistical principle of regression to the mean. To the extent that corporate earnings are affected by temporary, unexpected events—and they surely are—the most profitable companies in any given
year are more likely to have had positive surprises than negative ones. They typically did not do as well the year before and or the year after.

If investors do not anticipate this subsequent regression toward the mean, they will bid up the stock prices of the most profitable companies. Then, when earnings regress, these stock prices will suffer. This pattern of price increases followed by price decreases is mean reversion.

This argument that regression to the mean in earnings causes a mean reversion in prices explains the conclusion of Lakonishok, Shliefer, and Vishny: “a likely reason that these value strategies have worked so well relative to the glamour strategies is the fact that the actual future growth rates of earnings, cash flow, [and sales] etc. of glamour stocks relative to value stocks turned out to be much lower than they were in the past.”8

Similar reasoning applies to athletic performances that depend on chance as well as skill. Extreme performances tend to reflect abilities that are closer to the mean than are the performances, and are consequently typically followed by performances closer to the mean.

Consider a very simple model in which observed performance $Y$ fluctuates randomly about ability $\mu$,

$$ Y = \mu + \varepsilon $$

(1)

where $\varepsilon$ is a random error term with an expected value of zero. A person’s ability is the expected value of his or her performance and is consequently an unbiased predictor of performance.

The variance of performances across individuals is larger than the variance of abilities:

$$ \sigma_Y^2 = \sigma_\mu^2 + \sigma_\varepsilon^2 $$

(2)

Therefore, extreme performances usually reflect less extreme abilities.
The regression-to-the-mean insight is that, if we compare the performance of one person relative to the performance of others, we should take into account the statistical argument that an outstanding performance probably exaggerates a person’s ability. It would be unusual if an athlete who outperformed others had performed below his or her ability. Those with outstanding performances generally have somewhat less exceptional ability. The larger is the variance of the error term in Equation 1, the lower is the correlation between performance and ability, and the more regression we should anticipate.

For example, the first pick in the National Football League (NFL) draft is considered the best available player by the team that drafts him, based on his performance in college and at the draft combine. However, the top draft pick seldom turns out to be the best NFL player in his draft cohort. During the 28-year period 1990–2017, only two number-one picks were voted rookie of the year. Almost half (13) never played in the Pro Bowl in their career (although some are still young enough that they may play someday).

Perhaps the disappointing performance of these number-one picks is due to the fact that a player’s football performance is affected by the other players on his team, and the number-one picks are normally made by the worst team the previous season. However, the same pattern is seen in baseball, where a player’s performance is largely independent of one’s teammates. There have been 53 number-one Major League Baseball picks from 1965 through 2017. Only three have been voted Rookie of the Year, even though there are Rookie of the Year awards in both the National League and American League and there were two winners in the National League in 1976 and two winners in the American League in 1979. Ken Griffey, Jr., is the only number-one pick who made it to the Hall of Fame.
Are these number-one picks the exception to the rule or are they the rule? Massey and Thaler\textsuperscript{9} analyzed NFL draft picks for the years 1983 through 2008. They estimated that the chances that a drafted player will turn out to be better than the next player in his position drafted (for example, the first quarterback drafted compared to the second quarterback drafted) is only 52 percent, barely better than a coin flip. Yet, teams pay much more for early draft picks than for later picks, evidently because they do not fully anticipate the regression that usually occurs.

This regression is also true of professional football teams because their performance depends on both ability and chance—unpredictable variations in the opponent’s play calling, the players’ health, the officiating, and even the proverbial bounces of the football. When such chance events make observed performance an imperfect measure of ability, observed differences in team performances tend to exaggerate underlying differences in team abilities. Performances consequently regress toward the mean, in that those teams that perform far from average typically do not perform as far from average subsequently.

Kahneman and Tversky\textsuperscript{10} argue that most people are blind to regression to the mean in that they are surprised when regression occurs and invent fanciful theories to explain it. If pilots who excel in a training session do not do as well in the next session, it is because the flight instructors praised them for doing well. If low-scoring students do better after special tutoring, it must have been because of the tutoring.

Many sports fans underestimate the role of chance in athletic contests. They consequently interpret fluctuations in performance as fluctuations in skill. For example, inevitable runs in performance are thought to be meaningful hot and cold streaks in ability.\textsuperscript{11} Similarly, when an outstanding performance regresses downward to the mean, fans attribute it to a Cy Young Award
jinx, rookie-of-the-year jinx, or *Sports Illustrated* cover jinx. More recently, the Madden Curse says that the football player whose picture appears on the cover of Madden NFL, a football video game, will not perform as well afterward.

If sports bettors do not appreciate fully that chance events make outcomes an imperfect measure of ability, they will anticipate that teams doing well and teams doing poorly will continue doing as well and as poorly, overlooking the statistical fact that the former have undoubtedly had more than their share of good luck and the latter more than their share of bad luck. Gamblers will be too optimistic about successful teams and too pessimistic about unsuccessful ones, creating potentially exploitable inefficiencies in betting odds.

**Football Wagers**

The *spread* set by bookmakers is a prediction of a team’s margin of victory. In the first week of the 2014 season, the New England Patriots played the Miami Dolphins and the bookmakers made New England a 4-point favorite. Gamblers who bet on New England “give” 4 points and win their bets if New England covers the spread by winning by more than 4 points; those who bet on Miami “take” 4 points and win their bets if Miami covers the spread either by winning or by losing by fewer than 4 points. If New England wins by exactly 4 points, this is called a *push* and the bettors get their money back.

Another popular wager is the *totals* (or *over/under*) bet that is based on the total number of points scored by both teams. In the New England-Miami game, the totals line was 47. Those who pick *over* win their bets if the total score is greater than 47; those who pick *under* win their bets if the total score is less than 47. In this game, Miami won 33-20, so those who picked Miami and those who chose over won their bets.
Typically, for every $11 a bettor wagers, a winning bet pays $10, so that a bettor needs a probability $P = 0.5238$ of winning for a bet to have a positive expected value:

$\$10P + (-\$11)(1 - P) > 0$ \text{if } P > 0.5238.

The large amount of money wagered on football games is evidence that bettors do not think that it is difficult to win 52.38% of the time. The longevity of bookmakers is evidence that picking winners is more difficult than bettors believe. Yet, gamblers continue betting because they interpret their winning bets as a confirmation of their skill, and explain away their losses as flukes and near-wins.\textsuperscript{12}

The bookmakers’ profit is risk-free if it does not depend on the outcome of the game. If the spread equalizes the dollars wagered on each team (and the totals equalizes the dollars wagered over and under), then the losers pay the winners $10 and pay the bookmaker $1, no matter how the game turns out. This $1 profit (a 4.55\% \textit{vigorish} or \textit{rake} relative to the $22 bet by both sides of the wager) compensates bookmakers for the expenses of making a market and for the risk that the spread (or totals line) may be set incorrectly—leading to a disproportionate number of winning bets.

Bookmakers attempt to take into account all the factors that bettors take into account: player strengths and weaknesses, injuries, home-field advantage, historic rivalries. If bookmakers are aware of systematic bettor irrationalities, they will take these into account, too.\textsuperscript{14} For example, if gamblers are inclined to bet on the Green Bay Packers for sentimental reasons, bookmakers will set the spread as if Green Bay is a better team than it really is. If, objectively, the Packers’ margin of victory against a certain opponent is equally likely to be above or below 3 points, bookmakers might make the Packers a 4-point favorite in order to equalize the dollars wagered for and
against the Packers. The “smart money” will bet against the Packers and win more than half the time. Of course, bookmakers must be wary of setting betting lines that reflect the biases of naive gamblers but create financially ruinous opportunities for smart gamblers.\textsuperscript{13} (Bruce and Johnson, 2000).

Several studies\textsuperscript{15,16,17} have concluded that the point spread is an unbiased predictor of the margin of victory. Sauer\textsuperscript{18} surveys the evidence for several different sports and concludes that wagering markets are generally efficient. Gray and Gray\textsuperscript{19} argue that NFL bettors put too much emphasis on recent results relative to long-run performance. However, Dare and MacDonald\textsuperscript{20,21} reanalyze their data and dispute their conclusion.

Brown and Sauer\textsuperscript{22} and Camerer\textsuperscript{23} detect evidence that bettors in the professional basketball market are prone to the hot-hand belief that teams on winning and losing streaks are likely to continue their streaks, but Durham, Hertzel, and Martin\textsuperscript{24} analyze college football wagers and report that bettors expect winning and losing streaks to be reversed (the gambler’s fallacy). Many of these studies do not explicitly analyze the profitability of strategies based on these discovered patterns.

**A Betting Strategy**

Our strategy is based on the simple presumption that a blindness to regression toward the mean causes gamblers and/or bookmakers to overrate successful teams and to underrate unsuccessful ones. Since betting lines are intended to equalize the dollars wagered on opposing bets, point spreads will be too high for successful teams and under/over totals will be too high for teams that have been scoring a lot of points.

We expect to find statistically persuasive evidence of regression to the mean in football
performances, but we do not expect large unexploited profits. There are just too many people looking too hard for betting strategies that will win more than 52.38% of the time. In fact, there is at least one betting blog\textsuperscript{25} that explicitly mentions regression to the mean.

**Spread Bets**

When a team beats the spread, one possible interpretation is that the team is better than previously believed, and estimates of the team’s ability should be revised upward. However, the regression-to-the-mean argument cautions that an exceptional performance overstates a team’s ability. If bettors do not fully appreciate the regression argument, they will overreact to an outstanding performance and push the spread too high the next week (this team will be too big a favorite). If so, it may be profitable to bet against teams that beat the spread and in favor of teams that do not.

Because gamblers have memories, we keep track of each team’s cumulative record against its spreads. We restart this cumulative tabulation each season because of the large number of personnel changes between seasons. Our strategy is to bet against whichever team has been more successful in covering its spreads. This strategy does not require an explicit comparison of the teams’ estimated abilities with the spread. We simply assume that the bettors push the spread too high for teams that have been covering the spread and too low for teams that have not been covering the spread.

In one version of our strategy, we measure the success differential $D$ when Team 1 plays Team 2 by the difference between the cumulative number of points by which each team has covered the spread over the preceding $m$ weeks.
If $D$ is positive, we bet $D$ dollars on Team 2; if $D$ is negative, we bet $D$ dollars on Team 1. In the second version, we measure success by the cumulative number of games a team has beaten its spread: $s$ is equal to 1 if a team covers the spread and –1 if it does not cover the spread.

Either way, we scale the wagers up or down, as needed, to make the total amount bet each week equal to $1000$.

As an example, consider the game between New England and Minnesota in the second week of the 2014 season. In the first week, New England was a 4-point favorite and lost by 13 points, putting it 17 points under the spread. In Minnesota’s first game of the 2014 season, Minnesota had been a 3-point underdog against the St. Louis Rams and won by 28 points, putting it 31 points over the spread.

Because New England was 17 points under the spread and Minnesota was 31 points over the spread, the success differential $D$ is 48 points. We bet on New England and we bet twice as much money on this game as on a game where the success differential is 24 points, subject to the constraint that the sum of all wagers is $1,000$. Notice that our bet is not based on the spread for the New England-Minnesota game, only on the teams’ past performance relative to their spreads.

Minnesota had done better than New England against the spread, so we figured that bettors would, on average, be too optimistic about Minnesota’s chances and bookmakers would adjust the spread accordingly in order to even out the amount of money wagered on each team. So, we bet on New England. New England was a 3-point favorite over Minnesota and, as it turned out, New England won 30-7 and we won this bet.
An alternative strategy is to count the number of times the teams covered or failed to cover the spread, without regard for the exact number of points. For the New England-Minnesota game, New England had failed to cover the spread the first week and Minnesota had covered its spread, so the New England-Minnesota success differential is 2 games. We bet on New England and we bet twice as much money on this game as on a game where the success differential is 1 game.

**Over/Under Bets**

Our reasoning is similar for over/under bets. The regression-to-the-mean argument suggests that in a game that goes over the totals line, the teams are likely to have experienced good luck in scoring points and/or bad luck on defense. If bettors do not fully appreciate the fact that chance factors most likely played a role in the score going over the totals line, they will make the totals lines for these teams too high in subsequent weeks. If so, it may be profitable to bet that teams that have been going over the line will go under the line. Again, this argument does not require an explicit comparison of offensive or defensive abilities, only knowledge about whether the two teams have been going over or under the totals line in previous weeks.

As with the spread, we considered two strategies, one based on the cumulative number of points teams had been over or under the line, and the other on the cumulative number of games teams had been over or under. Again, consider the game between New England and Minnesota in the second week of the 2014 season. In the first week, New England’s totals line was 47 and the 33-20 score was 6 points over. In Minnesota’s first game of the 2014 season, the totals line was 43 and the 34-6 score was 3 points under the line.

With New England 6 points over the line and Minnesota 3 points under the line, these two
teams had, on balance, been 3 points over their lines and we figured that bettors would be too optimistic about the total score, leading bookmakers to set the totals too high—so, we bet under. The New England-Minnesota totals was 49 and the score turned out to be 30-7, so we won our under bet.

For our second strategy, counting the number of times the teams were over or under the line, without regard for the exact number of points, New England had been over the first week and Minnesota under, so we did not make an over/under bet.

**Cutoffs**

Our strategy is based on the hypothesis that gamblers underestimate the role of chance in determining football scores. Logically, this myopia will be most exploitable when teams have performed exceptionally well or poorly. If one team has cumulatively beaten its spreads by 2 points and the opposing team has beaten its spreads by one point, we might argue that bettors will be slightly too optimistic about the first team and we should consequently bet against it. However, this excess optimism is so slight that it should not bias the spread enough to create an attractive betting opportunity, considering the 4.55% vigorish. The slight bias is likely to be overwhelmed by the randomness in game scores.

This argument suggests setting a cutoff so that we do not bet on a game unless the success differential is larger than the cutoff. We looked at cutoffs that yielded at least (approximately) 100 bets for either the spread or totals. In practice, this turned out to be 0 to 150 points for the cumulative-points strategy and 0 to 10 games for the cumulative games strategy.

**Data**

Betting lines and results for all National Football League games from 1993 through 2017 were
collected from Football Locks\textsuperscript{26} and The Gold Sheet.\textsuperscript{27} The data are virtually identical with an occasional half-point difference in the quoted spread or totals line. Competing bookmakers generally offer nearly identical betting lines; otherwise, gamblers would pick and choose where to place their wagers, undermining each bookmaker’s attempt to equalize the number of dollars wagered on each side of a bet.

We analyzed each season independently of other seasons because of the substantial player turnover between seasons. We did not analyze postseason games because these are played by selected teams at a higher level of intensity with somewhat different objectives. In regular season games, teams are encouraged to score as many points as possible, since net points are part of the tie-breaker system to determine qualification, seeding, and home field for post-season games. Post-season games may be played more conservatively since one loss eliminates a team and the margin of victory is unimportant.

Table 1 provides some summary statistics for spreads. The calculation of the number of times the favorite went over or under the spread excludes “pick-em” games with a zero spread. The average spreads and actual margins are from the standpoint of the team favored to win the game. The standard deviations and correlation coefficients are from the standpoint of the home team.

The total number of times that the favorite covered the spread is very close to the number of times it has not (49.7 percent covered; 50.3 percent not covered), and the average actual margin of victory is close to the average spread (both from the standpoint of the favored team). However, the overall correlation between the spread and victory margin is only 0.41.

The over/under data in Table 2 are similar. The number of games over and under are nearly equal (50.2 percent over, 49.8 percent under), as are the average total score and the average totals
line. The correlation between totals lines and total scores is 0.27, even smaller than the correlation between the spread and margin of victory.

Even though the spreads and totals are approximately correct, on average, the low correlations with the actual scores implies that the outcomes are heavily affected by unpredictable events. Coaches making bad decisions, coaches making good decisions. Defensive players making bad guesses, offensive players making good guesses. Defensive players slipping, offensive players not slipping. Fumbles lost and recovered. Passes caught and dropped. Holding penalties called and not called. The list is very long. Luck—good and bad—is why the best team doesn’t win every game and the final score is so hard to predict.

The importance of chance is also demonstrated by the fact that the standard deviation of the actual margin of victory is much larger than the standard deviation of the spread, and the standard deviation of the total score is much larger than standard deviation of the totals line.

This uncertainty is why performances regress toward the mean.

Results

To implement our betting strategy, we begin betting during the second week of each season. We start each season with a $1,000 stake and bet $1,000 each week. Weekly profits and losses are accumulated and do not earn interest or pay interest. At the end of the football season, we calculate the simple percentage rate of return on our $1,000 stake, without adjustment for it being only a four-month investment.

Table 3 shows a summary of the results of wagers based on the cumulative points a team has been over or under its spreads or totals. The win percentages are relative to the number of wins and losses, excluding pushes. The $p$ values use the binomial distribution to calculate the
probability of winning so many bets if the probability of making a winning wager were 0.50. (The strategies are sometimes profitable even though the win percentage is less than 52.38 percent because relatively more money was wagered on what turn out to be winning bets.)

All of the strategies win more than 50% of the time, although the spread bets with low cutoffs do not win consistently enough to beat the vigorish. The totals strategies are consistently profitable. The spread and totals strategies that use larger cutoffs, reflecting more extreme differences in past performance, tended to be the most profitable.

Table 4 shows a summary of the results of wagers based on the cumulative number of games that a team has been over or under its spreads or totals. As with wagers based on cumulative points, spread bets with low cutoffs do not win often enough to beat the vigorish and, again, spread and totals bets with relatively large cutoffs are the most profitable.

The most profitable strategies involve situations where there has been a large disparity in the performance of the teams involved relative to their spreads or a large difference in their scores relative to their totals. For example, a spread bet with a cutoff of 10 games might involve betting on a team that has a net performance relative to its spreads of –4 when it plays a team that has a net performance relative to its spreads of +6. An over/under wager with a cutoff of 10 games might be an over bet when two teams have each been under a net of 5 times.

Logically, the spread and line strategies are essentially independent wagers; empirically, the correlation between the actual margin of victory minus the point spread and the actual total points scored minus the totals line is 0.019. Assuming these to be independent tests, we can multiply the two $p$ values in order to calculate the probability that both tests would show such large deviations from the expected number of winning bets if there is only a 0.50 probability of
making a winning bet. These joint $p$ values range from 0.00003 to 0.02292.

Kahneman and Tversky\textsuperscript{28} argued that people tend to draw conclusions from very limited data, what they call “the law of small numbers.” It is a mistake, for example, to see a team win one football game convincingly and conclude that it is a great team. This error overlaps the regression argument that performances that are far from the mean exaggerate how far abilities are from the mean, but one way to distinguish between these two errors is to compare the success of a betting strategy that is based solely on the results of the most recent game with a strategy (like ours) that considers results from earlier weeks.

We consequently redid all of our calculations using horizons of 1 to 15 games (each team plays 16 games). For example, with a horizon of four games, we assume that gamblers look back no more than four games. The second week of the season, they look back one game; the third week, they look back two games; after the fourth week, they look back four games.

Figures 1 and 2 show the results. For all horizons, all the wagers were successful at least half the time. Strategies based just on the most recent game were not more successful than strategies based on all games, suggesting that the success of our strategy is due not to a myopic overweighting of recent information but rather to an insufficient appreciation of regression to the mean.

Borghesi\textsuperscript{29} finds a consistent pattern within seasons; specifically, that point-spread bets on home underdogs after week 14 of the season are consistently profitable. To investigate whether our results reflect late-season anomalies, we redid the calculations with betting restricted to weeks 15 through 18 (still using data from earlier weeks to determine the success differentials). Tables 5 and 6 show no consistent discrepancies with the full-season bets in Tables 3 and 4, other
than the $p$ values increasing because of the smaller sample sizes. If anything, late-season point-spread wagers, while still winning more than half the time, tend to be less profitable than full-season wagers.

**Summary**

There is compelling evidence that stock market investors do not fully appreciate the importance of random events on corporate performance and consequently do not fully anticipate regression to the mean in corporate earnings, which creates profitable investing opportunities. This paper investigates whether the same is true of football wagers. If gamblers do not fully appreciate the importance of random events in football games and consequently do not fully anticipate regression to the mean in team performance, there may be profitable betting opportunities.

Data from 1993 through 2017 on two very different kinds of football wagers—spread and over/under—indicate that gamblers do not take regression to the mean into account fully. A strategy of betting against teams that have been much more successful than their opponents in covering their spreads would have been profitable, even taking into account the bookie’s 4.55% vigorish. So, too would a strategy of betting under when the two teams have been going far over their totals lines, and betting over when the reverse is true. Not surprisingly, the most profitable opportunities occur less frequently because they involve unusually extreme past performances.

As is true in the stock market, gamblers evidently overreact to chance fluctuations in football scores, and this has an excessive effect on betting lines, which creates profitable betting opportunities.
Financial disclosure statement: No competing financial interests exist.
References


3. Smith G. Companies are Seldom as Good or as Bad as They Seem at the Time, in Guzman M. Ed. Toward a Just Society: Joseph Stiglitz and Twenty-First Century Economics, Columbia University Press. 2018. 95-110.


20. Dare W.H., and MacDonald S.S. A generalized model for testing home and favourite team


