The Louis-Schmelling Paradox and the League Standing Effect Reconsidered

Brad R. Humphreys* — West Virginia University
Li Zhou† — University of Alberta

July 7, 2014

Abstract

Fifty years on we examine two key propositions in Neale's (1964) "Peculiar Economics": the need for competitors in sport to have opponents of similar ability in order to earn large revenues and the effect of frequent changes sports leagues' standings on consumer demand. We develop a consumer choice model under uncertainty, and a structural econometric model, to motivate and test these ideas. Unfortunately, neither receives much empirical or theoretical support relative to alternative factors affecting consumer choice like loss aversion and home win preference.

JEL Codes: L83, D12

Key Words: Reference dependent preferences; outcome uncertainty; league standing effect, Louis-Schmelling Paradox

1 Introduction

2014 marks the fiftieth anniversary of the publication of "The Peculiar Economics of Professional Sports," Neale (1964), a seminal work in the sports economics literature. Neale's paper contains a wide-ranging discussion of topics related to observed outcomes in professional sports leagues. Among these are claims about the nature of outcome uncertainty at both the match or game level and league-wide. Neale (1964) refers to game or match level outcome uncertainty and the "Louis-Schmelling Paradox" and league-wide outcome uncertainty as the "League Standing Effect." In this paper, we address how models of consumer choice under uncertainty provide insight into Neale's claims about the impact of game level and league-wide outcome uncertainty on attendance.

*Paper prepared for the IX Gijon Conference on Sports and Economics: "Neale Golden Anniversary" 9-10 May 2014., Department of Economics, College of Business and Economics, PO Box 6025, Morgantown WV 26506-6025; email: brhumphreys@mail.wvu.edu; phone: 304-293-7871.

†Department of Economics, 8-14 Tory, Edmonton, AB T6G 2H4 Canada; email: Li.Zhou@ualberta.ca; phone: 780-492-4133.

Before the development of axiomatic models of game-level outcome uncertainty (Coates et al., 2014), the tools of economic theory had not been applied to Neale's outcome uncertainty discussion. The classical Uncertainty of Outcome Hypothesis (UOH), attributed to Rottenberg (1956), asserts that fan's interest is largest when the game has an uncertain outcome, other things equal, so attendance will be higher, and gate revenues maximized, when a team plays games with uncertain outcomes compared to games where the home team is expected to win most or all of the contests. While easy to empirically implement, given a variable reflecting game-level outcome uncertainty, the theoretical underpinnings of the classical UOH were unexamined for decades.

Neale's (1964) "League Standing Effect" follows immediately from the classical UOH. If each team in a league plays games against opponents with relatively equal talent, games have uncertain outcomes and there will be regular changes in the league standings, generating additional fan interest in games, and, according to Neale (1964), increasing gate revenues.

Coates et al. (2014) develop a model of consumer choice under uncertainty to explain the UOH. In this model, consumers have reference dependent preferences about game or match outcomes, and their utility includes both standard consumption utility and "gain-loss" utility that reflects differences between their expected outcome of games and actual game outcomes. Fan's can also exhibit loss aversion in this model, in that the marginal utility of a loss by the home team when the fan expected the home team to win may exceed the marginal utility of an home win when the fan expected a loss. The UOH emerges only as a special case in this model, and the version of the model with loss averse fans appears to have significant empirical support compared to the UOH version. This model also simplifies to a case where fans only care about seeing the home team win; in this case, fans have preferences for home team wins above all other possible outcomes, given that game outcomes are uncertain.

The loss aversion and home win preference cases generate tension in terms of both game level and league-wide outcome uncertainty. If fans exhibit either, then teams will not want to play large numbers of games with uncertain outcomes. In the loss aversion case, either the possibility of an upset of a stronger team or the home team overpowering weaker opponents both generate more expected utility for fans than games with uncertain outcomes, suggesting that attendance will be lower at games with an uncertain outcome. In the home win preference case, fans only want to see the home team win, and games with uncertain outcomes, or games where the home team is an underdog, generate less expected utility for fans than games the home team is expected to win. In either case, teams would not be interested in playing games with uncertain outcomes. In this case, the "League Standing Effect," where teams are close in the standings and frequent changes in the rank ordering of teams takes place, can be thought of as a public good that all teams must produce in order to generate a "league standing effect" that increases demand for all games played in the league, as Neale (1964) posited.

We develop a model of game outcomes that includes fans with both reference dependent preferences and a preference for a "League Standing Effect" that increases expected utility when there are frequent changes in the league standings. The model is general enough to capture the “Louis-Schmelling Paradox”, the “League Standing Effect,” reference dependent preferences, and loss aversion. Based on this model of consumer choice under uncertainty, we derive a structural econometric model that generates empirically testable predictions. We extend the model in Coates et al. (2014) by including the “League Standing Effect” and home win preference in the consumer choice model.

We test these predictions using data from Major League Baseball (MLB) games over the period 2006-2010. We estimate a structural econometric model of game-level attendance containing variables that reflect the “league standing effect” posited by (Neale, 1964): two variables reflecting daily changes in rank order standings in each of the six MLB divisions on each day of the regular season, and a variable reflecting the standard deviation of winning percentages of teams in each MLB division on every day of the season. The empirical results do not support the presence of a “League Standing Effect” in this sample. Day-to-day changes in the rank order standings are not associated with increases in attendance at MLB games, and variation in the daily standard deviation in winning percentages in divisions is also not associated with changes in attendance. The results support the existence of both reference dependent preferences with loss aversion, and home win preference, even when controlling for a “League Standing Effect.” The results are also not consistent with Neale’s (1964) “Louis-Schmelling” paradox. After 50 years, these two concepts put forth by Neale do not appear to have stood the test of time.

2 Modeling Attendance and Outcome Uncertainty

Neale’s (1964) represents a seminal work in sports economics, and has accumulated more than 500 citations, according to Google Scholar. The paper contains numerous insights into the working of sports leagues, including many specific predictions. Curiously, little subsequent research focused on converting Neale’s insights into axiomatic economic models of consumer choice. In this paper, we incorporate two key concepts from the paper, the “Louis-Schmelling Paradox” and the “League Standing Effect” into a consumer choice model. Similar to Coates et al. (2014), this model builds on the reference-dependent preference framework developed by Koszegi and Rabin (2006), where fans have reference-dependent preferences and preferences for home team success; the model also explicitly contains game outcome uncertainty.

Consider a game or match held with a predicted probability p that the home team wins. The outcome of this game is more uncertain when p is close to 0.5. We use $p(1-p)$ to measure the outcome uncertainty for a game: the larger $p(1-p)$, the more uncertain the outcome of the game.[1] Let $\alpha_{uo} \geqslant 0$ capture the marginal impact of uncertainty of game outcome $p(1-p)$ on fans’ utility. $\alpha_{uo} > 0$ implies that fans prefer games with uncertain outcomes.

Let y be an indicator variable that reflects game outcomes, where $(y = 1)$ represents a home team win and $(y = 0)$ a home team loss. In North America, people who attend games primarily reside in the city

[1] Tainsky and Winfree (2010) use a similar measure in an empirical analysis of the UOH.

where the game is played. It is reasonable to assume that fans' utility depends on whether the home team wins. Let $\alpha_{hw} \geqslant 0$ capture fans' marginal utility from attending a home win. The utility generated from a home win is $\alpha_{hw} y$ and $\alpha_{hw} > 0$ implies that fans have home win preference.

Neale (1964) posited that overall league-wide uncertainty of outcome may also affects fans' utility, through the "League Standing Effect." Fans derive more enjoyment from discussing the game when there are more or more frequent changes in the standings, or league table. Let $\alpha_l \geqslant 0$ reflect the marginal impact of league standing changes, indicated by LS, on fans' utility; $\alpha_l > 0$ is consistent with Neale's "League Standing Effect."

Following Coates et al. (2014), we assume that fans have reference-dependent preferences, like those developed by Koszegi and Rabin (2006). Fans form an expectation about any game or match outcome $E(y = 1) = p^r$. In addition to the intrinsic "consumption utility" from attending a game, fans also have "gain-loss utility" that depends on the deviation of actual game outcomes from the reference point p^r. Assume the marginal impact of positive deviations from the reference point when the home team wins $(1 - p^r)$ is $\beta_1 (\geqslant 0)$, and the marginal impact of a negative deviation from the reference point $(0 - p^r)$ when the home team loses is $\beta_2 (\geqslant 0)$. Fans have reference-dependent preferences if either $\beta_1 > 0$ or $\beta_2 > 0$. Fans have loss aversion if $\beta_2 > \beta_1 \geqslant 0$. We assume that the reference point, fans' subjective expectation of a home team win, is equal to the objective probability that the home team wins, i.e., $p^r = p$.

Fans may also care about the overall quality of play in a game or match, independent of their assessment that the home team will win a game. Let α_q reflect the marginal impact of the quality of a game independent of the probability of a home team win, captured by the variable q, on fans' utility. $\alpha_q > 0$ implies that fans get higher utility from watching high quality games regardless of expected game outcome.

Finally, we assume that fans derive some baseline utility from attending a live game, independent of specific game characteristics. This baseline utility contains two parts: a common deterministic component that represents generic "fandom" and an idiosyncratic component that varies across fans and reflects the relative intensity of this baseline utility. Let α_0 capture the common deterministic component of baseline utility from attending a live game and v_i capture the idiosyncratic part of baseline utility. v_i is a random variable that is uniformly distributed across the fan population over the interval $[0, 1]$.[2] As v_i gets larger, the baseline utility derived from attending a game gets larger.

For fan i, the utility from attending a home win $(y = 1)$ with *ex ante* home win probability p is

$$U_i(y = 1) = \alpha_0 + \alpha_{hw} + \alpha_{uo} p(1 - p) + \alpha_l LS + \beta_1(1 - p) + \alpha_q q + v_i$$

For a game that the home team loses $(y = 0)$ with *ex ante* win probability p, the total utility for fan i is

$$U_i(y = 0) = \alpha_0 + \alpha_{uo} p(1 - p) + \alpha_l LS + \beta_2(0 - p) + \alpha_q q + v_i$$

[2] The range of the distribution is normalized to $[0, 1]$ for simplicity. Generalizing the range to $[0, a]$ rescales the parameters by $\frac{1}{a}$ but will not change the signs of the parameters.

The expected utility from attending a game with *ex ante* home team win probability p varies with the outcome of the game. Expected utility is

$$\begin{aligned} E[U_i|p] &= pU_i(y=1) + (1-p)U_i(y=0) \\ &= \alpha_0 + \alpha_{hw}p + \alpha_{uo}p(1-p) - (\beta_2 - \beta_1)p(1-p) + \alpha_l LS + \alpha_q q + v_i \end{aligned}$$

Define $\beta \equiv \beta_2 - \beta_1$. $\beta > 0$ indicates that fans have loss aversion, in that the marginal utility of a home team loss when the fan expected the team to win exceeds the marginal utility of a home team win when the fan expected the team to lose. Expected utility becomes

$$E[U_i|p] = \alpha_0 + (\alpha_{hw} + \alpha_{uo} - \beta)p + (\beta - \alpha_{uo})p^2 + \alpha_l LS + \alpha_q q + v_i \quad (1)$$

Equation (1) motivates both the "Louis-Schmelling Paradox" and the "League Standing Effect" in the context of a consumer choice model with game outcome uncertainty. The "Louis-Schmelling Paradox", or the "classical UOH" (Coates et al., 2014), posits that individual game attendance first increases with the probability of home win p, then decreases with p, and reaches a maximum level of attendance at some p where $p < 1$.

From Equation (1), the "Louis-Schmelling Paradox" is a composite hypothesis. Neale (1964) did not consider the presence of loss aversion, so $\beta = 0$ is implicit in the "Louis-Schmelling Paradox". The other part of the hypothesis is that $\alpha_{hw} > 0$, $\alpha_{uo} > 0$, and $0 < \alpha_{hw} < \alpha_{uo}$.[3] In other words, for the classic UOH to hold, the marginal utility generated by a home win and the marginal utility of outcome uncertainty must be positive, and the marginal utility of outcome uncertainty must exceed the marginal utility generated by a home win.

The second element of this part of the classical UOH, emphasizing the importance of home win preference, has not been recognized in most UOH research. Fans surely derive utility from watching the home team win a game or match, regardless of how much uncertainty exists about the result. For the classical UOH to hold, the marginal utility from game outcome uncertainty must outweigh the marginal utility from a home win.

The League Standing Effect is simply $\alpha_l > 0$. Neale (1964) clearly had in mind the idea that the "League Standing Effect" belonged in consumers' utility function. In his description of the effect, he refers to "*race utility*" (italics in the original) when describing the effect.

2.1 The Consumer Choice Problem

Equation (1) can be used to generate predictions about fan's decisions to attend games. Assume fan i chooses to attend or not attend a live game to maximize total utility, which is the sum of utility from consumption

[3] With $\alpha_{hw} > 0$ and $\alpha_{uo} > 0$, the maximum attendance is achieved at $p = \frac{\alpha_{hw}+\alpha_{uo}}{2\alpha_{uo}}$. The classic UOH requires that the home win probability associated with maximum attendance be strictly smaller than 1, i.e., $\frac{\alpha_{hw}+\alpha_{uo}}{2\alpha_{uo}} < 1$ or $0 < \alpha_{hw} < \alpha_{uo}$.

of a standard composite consumption good c and expected utility from attending a game conditional on the probability of a home team win

$$U_i = c + E[U_i|p]$$

subject to a budget constraint $c + P_T = m$, where P_T is the price of the ticket, and m is income. Note that the price of the composite consumption good has been normalized to one for simplicity. Normalize $E[U_i|p]$ to equal 0 if fan i chooses to not attend the game. The decision rule for utility maximization is to attend the game if

$$m - P_T + E[U_i|p] \geqslant m.$$

Using the budget constraint, this can be rewritten as

$$E[U_i|p] - P_T \geqslant 0$$

and substituting Equation (1) into this expression and rearranging terms gives

$$v_i \geqslant P_T - [\alpha_0 + (\alpha_{hw} + \alpha_{uo} - \beta)p + (\beta - \alpha_{uo})p^2 + \alpha_l LS + \alpha_q q]$$

In other words, fan i will attend the game if the idiosyncratic utility from game attendance v_i is bigger than the difference between the ticket price and the part of expected utility from attending the game that is common to all fans.

Given this expression, total attendance at a game will be the probability mass of potential attendees for which this expression holds $E[U_i|p] - P_T \geqslant 0$

$$\begin{aligned}
Attendance &= PotentialAttend * P(E[U_i|p] - P_T \geqslant 0) \\
&= PotentialAttend * P(v_i \geqslant P_T - [\alpha_0 + (\alpha_{hw} + \alpha_{uo} - \beta)p + (\beta - \alpha_{uo})p^2 + \alpha_l LS + \alpha_q q]) \\
&= PotentialAttend * [1 + \alpha_0 + (\alpha_{hw} + \alpha_{uo} - \beta)p + (\beta - \alpha_{uo})p^2 + \alpha_l LS + \alpha_q q - P_T]
\end{aligned}$$

This equation is simply a demand function for game attendance that includes outcome uncertainty, fans' home win preference, Neale's (1964) "League Standing Effect," and fans' preference for game quality. Total game attendance rises with these factors associated with expected attendance, rises with the potential number of attendees (*PotentialAttend*), and falls with the ticket price, which is assumed to be set in advance of the beginning of the season by the team. Note that parameters reflecting Neale's (1964) "Louis-Schmelling Paradox" and "League Standing Effect" appear in this demand function.

3 A Structural Econometric Model

Given the demand function derived in the previous section, we next develop a structural econometric model based on this model. First, log linearize the demand function

$$\ln Attendance_{hgt} = \ln PotentialAttend_{hgt} + \ln[1 + \alpha_0(\alpha_{hw} + \alpha_{uo} - \beta)p_{hgt} + (\beta - \alpha_{uo})p_{hgt}^2 + \alpha_l LS_t + \alpha_q q_{hg} - P_T]. \tag{2}$$

Next, apply the log approximation $\ln(x+1) \approx x$ to the second term on the right hand side of Equation (2) to get

$$\ln Attendance_{hgt} = \ln PotentialAttend_{hgt} + \alpha_0 + (\alpha_{hw} + \alpha_{uo} - \beta)p_{hgt} + (\beta - \alpha_{uo})p_{hgt}^2 + \alpha_l LS_t + \alpha_q q_{hg} - P_T$$

$PotentialAttend_{hgt}$ is the number of individuals who would consider attending a game independent of the characteristics of the game; this variable should be a function of characteristics of the local market, such as population, income, travel costs to the facility, time-related factors, such the day of the week and the start time of the game, and random factors such as weather. Assume that the functional form of the expression for the number of individuals in the area who would consider attending a game is

$$PotentialAttend_{hgt} = e^{X_{ht}\mu + \eta_h D_h + \eta_t D_t + \varepsilon_{hgt}},$$

where X_{ht} is a vector stadium and local market characteristics, D_h is a local market fixed effect capturing any unobservable market-specific heterogeneity, μ is a vector of unobservable parameters to be estimated, D_t is a vector of time-related factors that affect decisions made by this group of potential consumers, and ε_{hgt} is a random error term clustered on h that captures all other factors that affect the size of the population of residents who would consider attending a game between home team h and visiting team g.

The quality of the game q_{hg} between home team h and visiting team g that is independent of the home win probability is assumed to be a function of the ability of the players on the rosters of the two teams. Assuming that team rosters are relatively stable within a season, the game quality is an additively separable function each team's roster $D_r, r = h, g$ or

$$q_{hg} = \theta_h D_h + \theta_g D_g$$

By substitution, the log attendance model becomes

$$\begin{aligned} \ln Attendance_{hgt} &= (\alpha_0 - 1) + (\alpha_{hw} + \alpha_{uo} - \beta)p_{hgt} + (\beta - \alpha_{uo})p_{hgt}^2 + \alpha_l LS_t - P_T + X_{ht}\mu \\ &+ (\eta_h + \alpha_q\theta_h)D_h + \alpha_q\theta_g D_g + \eta_t D_t + \varepsilon_{hgt}. \end{aligned}$$

Rearranging and gathering terms gives a structural econometric model

$$\ln Attendance_{hgt} = \gamma_0 + \gamma_1 p_{hgt} + \gamma_2 p_{hgt}^2 + \gamma_3 P_T + \alpha_l LS_t + X_{ht}\mu + \gamma_h D_h + \gamma_g D_g + \eta_t D_t + + \varepsilon_{hgt} \tag{3}$$

where $\gamma_0 = (\alpha_0 - 1)$, $\gamma_1 = (\alpha_{hw} + \alpha_{uo} - \beta)$, $\gamma_2 = \beta - \alpha_{uo}$, $\gamma_g = \alpha_q \theta_g$, and $\gamma_h = \eta_h + \alpha_q \theta_h$. γ_3 reflects the effect of variation in ticket prices on demand.

Again, the joint hypothesis that $\beta = 0$ (implicit), and $\alpha_{uo} > \alpha_{hw} > 0$, which, in Equation (3) implies that $\gamma_1 > 0$, $\gamma_2 < 0$, and $0 < \gamma_1 + \gamma_2 < -\gamma_2$ embodies Neale's (1964) "Louis-Schmelling Paradox." We first estimate the parameters γ_1, γ_2, and α_l to determine if their estimated signs are consistent with Neale's (1964) ideas. The "League Standing Effect" can be empirically tested using the estimates of α_l, since $\alpha_l > 0$ implies that attendance is higher when more turnover in the league standings take place.

Our second goal is to estimate the parameters from a full model of fans' attendance decisions when fans have home win preference ($\alpha_{hw} > 0$), preferences for uncertain game outcomes ($\alpha_{uo} > 0$), and loss aversion ($\beta > 0$). Notice that we have three structural parameters (α_{hw}, α_{uo},and β) but only two reduced form parameters (γ_1 and γ_2) that we can estimate from the attendance model, Equation (3). We cannot identify all three structural parameters separately.

This identification issue has not been noted in previous research. Coates et al. (2014) discuss estimates of γ_1 and γ_2 in the existing literature and document the lack of consensus in terms of signs of he estimates of these parameters in the literature. One reason for the variation in these estimates is identification of thee structural parameters (home win preference, loss aversion, and preference for outcome uncertainty) from only two reduced form parameters. This may also be related to the bias in estimates of γ_1 and γ_2 discussed by Štrumbelj (ress).

The parameter capturing home win preference can be separately identified because $\alpha_{hw} = \gamma_1 + \gamma_2$. Note that $\gamma_1 + \gamma_2 > 0$ implies fans get higher utility from home team wins. Although we can not separately identify α_{uo} and β, we can infer the relative size of the two parameters from γ_2 since $\gamma_2 = \beta - \alpha_{uo} > 0$ means that loss aversion dominates fans' preferences for uncertain outcomes ($\beta > \alpha_{uo}$) and $\gamma_2 < 0$ means that fans' loss aversion is dominated by their preference for uncertain outcomes ($\beta < \alpha_{uo}$).

Notice that $\gamma_2 > 0$ provides evidence for the existence of loss aversion ($\beta > 0$) but cannot reject the hypothesis that $\alpha_{uo} > 0$, the existence of fan preferences for uncertain game outcomes. When fans have loss aversion $\beta > 0$, $\gamma_2 > 0$ only indicates that this preference for uncertain game outcomes is dominated by another more powerful behavioral response to outcome uncertainty, loss aversion.

Finally, to estimate γ_1 and γ_2, we have to deal with the correlation between quality of the game and the home win probability. Recall that we parameterized the quality of the game as $q_{hg} = \theta_h D_h + \theta_g D_g$ when we specified the structural econometric model, Equation (3). However, the home win probability p_{hgt} may be a function of the players on the rosters of the home and visiting team plus some random factors such as injuries, player availability due to the need for rest, and momentum of both teams. Formally

$$p_{hgt} = \lambda_h D_h + \lambda_g D_g + \lambda_{hgt},$$

where λ_{hgt} is the home win probability that is not explained by indicators for the home team and visiting team.[4] We use $\widehat{\lambda_{hgt}}$, the residuals from p_{hgt} regressed on D_h and D_g, as a measure of home win probability independent unobservable team quality vectors D_h and D_g and estimate the following modified attendance model

$$\ln Attendance_{hgt} = \gamma_0 + \gamma_1 \widehat{\lambda_{hgt}} + \gamma_2 \widehat{\lambda_{hgt}}^2 + \gamma_3 P_T + \alpha_l LS_t + X_{ht}\mu + \gamma_h D_h + \gamma_g D_g + \eta_t D_t + \varepsilon_{hgt}. \quad (4)$$

This model can be used to test for the presence of home win preference in consumers' decisions about game attendance.

4 Empirical Analysis

4.1 Data

We collected data on attendance and other characteristics for all Major League Baseball games in the 2006 through 2010 regular seasons. Our data set contains data from all home games of every MLB team, over 12,000 games. The data come from a variety of sources. Game attendance data, and data on scoring in the games and the teams involved in each game were collected from the MLB web site (www.mlb.com). Average ticket price data come from the Fan Cost Index collected and published by Team Marketing Report (www.teammarketing.com).

The probability that the home team wins each game is an important factor in the model developed above. We estimate the probability that the home team will win each game in our sample using betting odds data that come from Sports Insights (www.sportsinsights.com), a sports gambling information web site. We collected data on money line betting on MLB games; in North America, fixed odds betting is referred to as "money line" betting to distinguish it from point spread betting. The MLB money line data collected and distributed by Sports Insights reflects the average money line offered by three off shore, on-line sports books: BetUS.com, FiveDimes.com, and Caribsports.com. We converted the money line to standard odds, and then to the probability that the home team wins each game using the formula in Kuypers (2000). This variable is a market-based measure of the probability that the home team will each game.

We also collected data on the average ticket price charged by each MLB team. The ticket price data are the ticket price component of the Fan Cost Index (FCI) which is published annually by Team Marketing Report (`www.tmr.com`). We also collected data on the total population in the Metropolitan Statistical Area (MSA) that is home to each MLB team in the sample, from the US Census Bureau and Statistics Canada.

Descriptive statistics for key game-, team-, and market-specific variables in the data are reported in Table 1.

Average attendance was just over 31,000 per game. Sold out games are relatively uncommon in MLB. In this sample, about 18% of the games played had attendance greater than or equal to the listed capacity

[4] Štrumbelj (ress) makes a similar point using data from European football leagues.

Table 1: Summary Statistics - Game and Team Level Variables

	Mean	std dev
Attendance	31387	11030
Prob Home Team Win	0.542	0.081
Average Ticket Price (FCI)	24.75	9.11
Home team MSA Population (mil)	6.51	5.32
Home Team Avg Runs Scored	4.719	0.350
Visiting Team Avg Runs Scored	4.719	0.350
Home Team Avg Runs Allowed	4.718	0.367
Visiting Team Avg Runs Allowed	4.719	0.367

of the home stadium. The home team winning percentage is greater than 5%, which reflects the well-known "home advantage" in sports leagues (Levernier and Barilla, 2007). Average ticket prices were about $25 in real 2010 dollars. The FCI ticket price data represent the average of the available ticket prices posted by each team at the beginning of the season. This does not reflect any weighting by actual attendance at each listed ticket price.

4.2 Measuring the League Standing Effect

Neale (1964) clearly spells out the "League Standing Effect" on page 3: "The closer the standings, and within any range of standings, the more frequently the standings change, the larger will be the gate receipts." While we lack data on gate receipts, attendance is observable for MLB games, as are the standings at any point in time. Based on game outcomes, we calculated each team's winning percentage and the standings in each of the six divisions in MLB on every day of the season for the five seasons in our sample. During this period, MLB contained six divisions (East, Central and West in the National and American Leagues) and the winning team in each division automatically qualified for postseason play. Thus changes in the rank standings in each division should reflect Neale's "League Standing Effect," since teams in each division are competing for postseason berths.

To measure the closeness of the standings in each division, we calculated the standard deviation of winning percentage on each day of the season in each division. The smaller is the standard deviation of the winning percentages across teams in a division, the closer are the teams in the division standings, other things equal. We also calculated the rank order standings in each MLB division on each day of the season, and used this rank order in standings to calculate two variables that reflect how frequently the standings changed from day to day in each season in the sample. The first is the total number of changes in the order of the divisional standings compared to the standings on the previous day. If Team A was in first place on

Day 1, Team B in second place, Team C was in third place, and Team D was in fourth place, and on day two Team B passed Team A to take over first place, and no other changes in the order of standings took place, then the total changes in order would be two in this division on this day. If both Team B and Team C passed Team A, then the total number of changes in order would be three.[5]

The second variable reflects how far in the rank ordering each team in an MLB division moved from one day to the next. This variable reflects both how many teams changed order, and how far in the rank order the teams moved. In the example above, if Team C moved from third place to first place, and Team A fell from first place to third place, the number total number of changes in order would be 2, but the total number of changes in rank order would be 4; Team A dropped two spots in the rank order in the division, and Team C rose two spots in the rank order.

Note that we treat teams that are tied at the beginning of any day as having the average rank of each team. So two teams tied for first place would each have a rank of 1.5. This makes it possible for only one change in order to occur on a day. For example, suppose Team A and Team B were tied for first place on day 1. The rank of each team is 1.5. Team A beats Team B, putting Team A into first place and Team B into second place. Team A increases in ranking by 0.5 and Team B decreases in ranking by 0.5. Since we calculate standings based on winning percentage, most of these instances occur in the early days of each season.

Neale (1964) posits that the "League Standing Effect" works like advertising: "we may treat this effect as a kind of advertising. (page 3)" In many empirical tests of advertising, the effects are only apparent after repeated exposure, suggesting that cumulative variables best reflect the impact of advertising on consumer choice (Bagwell, 2007). To capture this, we also create variables reflecting the cumulative number of changes in the rank standings in these MLB divisions.

Since Neale (1964) also mentions that the closer the standings, the larger are gate receipts and attendance, we also calculated the standard deviation of winning percentages in each division on each day of the season. This variable will reflect how close the standings are on each day in each division, as a division where all teams have the same winning percentage on a given day will have a standard deviation of winning percentage equal to zero.

Table 2 shows the summary statistics for variables that capture the League Standing Effect in MLB divisions over the sample. The standard deviation of winning percentages are relatively small, reflecting relatively little within-division variation in winning percentage. There is, on average, about 1 change in the order of the standings in the average MBL division on the average day. However, this masks considerable differences in the actual distribution of the variables reflecting changes in the standings and in the rank standings.

[5]Occasionally MLB teams play "double headers" which consist of two teams playing two games on the same day. Doubles headers complicate the calculation of variables reflecting the "League Standing Effect" because they raise the possibility of multiple changes in the league standings at different points each day. We ignore double headers in these calculations. There were 330 double headers in the sample period, or slightly less than 3% of the games played.

Table 2: Summary Statistics - League Standing Variables

League Standing Variable	Mean	St. Dev.	Minimum	Maximum
Standard Deviation of Winning %, Each Division-Day	0.09	0.06	0	0.58
Total Changes in Order, Each Division-Day	0.93	1.36	0	6
Total Change in Rank Standing, Each Division-Day	0.79	1.34	0	10

Table 3 shows the distribution of the variable that reflects the total number of daily changes in the order of standings in each MLB division over the sample period. On 36% of the division-days there is a change in the standings and on 64% of the division-days, there are no changes in the order of the standings. The most likely change in the standings is two teams flip-flopping positions, which happens on 1 division-day in 4. About 11% of the division-days feature more than two teams changing place in the standings in a division. Note that all of the instances when 6 changes took place in a day are in the National League Central division, which contained 6 teams during the sample period.

Table 3: Distribution - Total Changes in Rank Order

# of Changes	Division-Days	Percent
0	7,736	63.63
2	3,034	24.96
3	590	4.85
4	537	4.42
5	213	1.75
6	47	0.39
Total	12,152	100.00

A small body of research has analyzed the relationship between season-level outcome uncertainty and attendance using other approaches. Lahvička (ress) contains a thorough discussion of the issues associated with measuring season-level outcome uncertainty. Tainsky and Winfree (2010) conduct a analysis for MLB using a number of alternative measures of season-level outcome uncertainty. Soebbing (2008), King et al. (2012), and Lei and Humphreys (2013) performed related empirical tests using similar season-level measures of outcome uncertainty. The results in this literature are generally mixed; Lahvička (ress) discusses reasons that measurement issues may explain the mixed findings in this literature. Curiously, no previous research has tested the "League Standing Effect" in the way Neale (1964) conceived of season-level outcome uncertainty. Here, we focus on a strict reading of Neale's paper as a motivation for the empirical analysis.

4.3 Results and Discussion

Two structural econometric models were developed above. Equation (3) includes actual home winning probability estimates based on betting market data. Equation (4) includes an adjusted home winning probability variable that accounts for unobserved quality of the teams in each game. This variable can be used to disentangle the effects of fan preference for quality from fan preference for home team wins, independent of team quality. The parameters of Equation (3) and Equation (4) were estimated using OLS. Similar results were obtained when these models were estimated with the Tobit maximum likelihood estimator to control for censoring in the form of sold out games.

Table 4 contains parameter estimates, estimated standard errors, and summary regression statistics for Equation (3). The dependent variable is log home attendance for each game between home team h and visiting team g. All standard errors have been cluster-corrected at the home team level. This model also included home team, visiting team, season, and team-season fixed effects to control for unobserved heterogeneity; these parameter estimates are not reported. The three models reported on Table 4 each contain a different variable reflecting the "League Standing Effect."

For two of the three models, the estimated parameter on the average ticket price variable is negative and statistically significant, the predicted sign on the demand curve for game level attendance. The parameter on the home team's MSA population variable, which captures the size of the pool of potential attendees, is positive and significant in all three models, as predicted. Among the four game-level team performance variables, home and visiting team runs scored and runs allowed, only home-team runs allowed explains observed variation in attendance; the estimated signs suggest that fans do not like to watch teams that give up a lot of runs, holding other factors constant. There is somewhat weaker evidence that fans like to watch teams that score a lot of runs, holding other factors constant.

The primary parameters of interest are those on the game-level and league-wide outcome uncertainty variables. The parameter estimates on the probability of a home win and probability squared variables are statistically significant at conventional levels and suggest the presence of loss aversion and home win preference, and not the "Louis-Schmelling' paradox in this setting. Recall that, in the context of Equation (3), the "Louis-Schmelling' paradox is the joint hypotheses that $\gamma_1 > 0$, $\gamma_2 < 0$, and $0 < \gamma_1 + \gamma_2 < -\gamma_2$. $\hat{\gamma}_1$ is clearly negative and $\hat{\gamma}_2$ positive; the p-value on the null hypothesis that $\hat{\gamma}_1 + \hat{\gamma}_2 = 0$ is about 0.45 for all three models. The parameter estimates on Table 4 clearly do not support the "Louis-Schmelling' paradox.

The parameter estimates on the three variables reflecting the "League Standing Effect" are not statistically different from zero in all three models. There is no evidence that variation in total daily changes in rank order in the three MLB divisions, cumulative changes in rank order to date, or the standard deviation of winning percentages in each MLB division on each day of the season are associated with changes in at-

Table 4: Empirical Results - Full Empirical Model

	Dependent variable: log(attendance)		
	(1)	(2)	(3)
Average Ticket Price (FCI)	-0.001	-0.001**	-0.001**
	(-1.53)	(-3.23)	(-3.21)
Home Team Metro Area Population	0.035***	0.035***	0.035***
	(4.21)	(4.33)	(4.23)
Home Team Avg Runs Scored	0.171	0.173	0.173
	(1.80)	(1.91)	(1.92)
Visiting Team Avg Runs Scored	0.126*	0.127*	0.126*
	(2.12)	(2.15)	(2.15)
Home Team Avg Runs Allowed	-0.253***	-0.253***	-0.256***
	(-4.02)	(-4.10)	(-4.14)
Visiting Team Avg Runs Allowed	-0.0279	-0.0277	-0.0278
	(-0.74)	(-0.73)	(-0.73)
Observations	12140	12140	12140
R^2	0.7433	0.7434	0.7433

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Empirical Results - Full Empirical Model

	Dependent variable: log(attendance)		
	(1)	(2)	(3)
Prob Home Team Win	-1.171**	-1.160**	-1.176**
	(-3.55)	(-3.51)	(-3.55)
Prob Home Team Win ^{2}	1.216***	1.206***	1.220***
	(4.08)	(4.04)	(4.07)
Cumulative Changes in Rank Order	0.050		
	(0.25)		
Total Changes in Rank Order		-0.004	
		(-1.98)	
Daily SD(wpct)			-0.023
			(-0.35)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

tendance. We also estimated these models with other alternative measures of the "League Standing Effect" including changes in rank, and moving averages of changes in rank order over the previous 3,5,7, 14, and 21 days in each division. None of these models generated evidence supporting the "League Standing Effect" in this sample.

Table 6 contains parameter estimates, estimated standard errors, and summary regression statistics for Equation (4). The dependent variable is log home attendance for each game between home team h and visiting team g and the home team winning probability variable adjusted for team quality. All standard errors have been cluster-corrected at the home team level. This model also included home team, visiting team, season, and team-season fixed effects to control for unobserved heterogeneity. These parameter estimates are not reported.

The estimated parameters on the average ticket price variable, as well as the market and game characteristic variables are basically unchanged in this specification. The estimated coefficient on one of the "League Standing Effect" variables is statistically significant, but the negative sign provides no support for Neale's (1964) prediction. The key parameters of interest are on the adjusted home win probability and probability squared variables. Line in the previous model, $\hat{\gamma}_1$ is clearly negative and $\hat{\gamma}_2$ positive. There is still no evidence supporting the "Louis-Schmelling Paradox". Instead, these parameter estimates suggest that the effect of loss aversion on fans' decisions to attend games still outweighs the effect of preferences for outcome uncertainty, even when the probability of a home team win is adjusted for unobservable game-level

Table 6: Empirical Results - Team Quality Adjusted Model

	Dependent variable: log(attendance)		
	(1)	(2)	(3)
Adj. Prob Home Team Win	-1.328**	-1.339**	-1.328**
	(-3.50)	(-3.53)	(-3.50)
Adj. Prob Home Team Win ^{2}	1.469***	1.480***	1.469***
	(3.78)	(3.81)	(3.78)
Average Ticket Price (FCI)	-0.001	-0.001**	-0.001**
	(-1.44)	(-2.99)	(-2.95)
Home Team Metro Area Population	0.0348***	0.0351***	0.0345***
	(4.18)	(4.31)	(4.20)
Home Team Avg Runs Scored	0.171	0.174	0.173
	(1.83)	(1.93)	(1.94)
Visiting Team Avg Runs Scored	0.129*	0.130*	0.130*
	(2.18)	(2.20)	(2.21)
Home Team Avg Runs Allowed	-0.255***	-0.255***	-0.257***
	(-4.08)	(-4.15)	(-4.19)
Visiting Team Avg Runs Allowed	-0.0299	-0.0296	-0.0297
	(-0.79)	(-0.78)	(-0.78)
Cumulative Changes in Rank Order	0.050		
	(0.25)		
Total Changes in Rank Order		-0.004*	
		(-2.09)	
Daily SD(wpct)			-0.014
			(-0.21)
Observations	12134	12134	12134
R^2	0.7430	0.7431	0.7430

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

heterogeneity.

Again, recall that $\gamma_1 + \gamma_2 > 0$ implies fans get higher utility from games that the home team wins, which we call "home win preference." For all three of the models shown on Table 6, the null hypothesis that $\hat{\gamma}_1 + \hat{\gamma}_2 = 0$ is rejected at p-values of about 0.004. Since $\hat{\gamma}_1 + \hat{\gamma}_2 > 0$, the results on Table 6 support the importance of home win preference, even when adjusting the probability of a home win for game quality. Note that we still cannot reject the hypothesis that $\alpha_{uo} > 0$, the existence of fan preferences for uncertain game outcomes, in this context. But these results strengthen the case that outcome uncertainty is the least important determinant of fan decisions, compared to loss aversion and home win preference.

5 Conclusions

Neale (1964) represents a seminal analysis of the economics of team sports leagues. The ideas in this paper have influenced sports economics research for fifty years. In this paper, we flesh out two key ideas posited by Neale (1964), the "Louis-Schmelling Paradox" and the "League Standing Effect" using a reference-dependent preference model of consumer choice under uncertainty. Both ideas can easily be incorporated into this model, providing new evidence that Neale's (1964) ideas are compatible with modern consumer choice models. This model also identifies two factors affecting consumer choice not considered as important by Neale (1964): loss-aversion and home win preference.

Neale (1964) hinted at the importance of home win preference when he wrote " 'Oh Lord, make us good, but not that good,' must be their prayer. (page 2)" The "make us good" part of this quote refers to home win preference. But he clearly places this in the context of the "classical" UOH that implicitly assumes success beyond some specific point leads to reduced gate revenues. Pure home win preference suggests that gate revenues increase monotonically with team success. Our model provides a clear distinction between home win preference and the classical UOH.

Using a structural econometric model derived from this consumer choice model, we investigate the empirical importance of the "Louis-Schmelling Paradox" and the "League Standing Effect," in the context of attendance at MLB games over the period 2006-2010. Unfortunately, these two ideas receive little empirical support in this setting. There is no evidence that greater turnover in league standings, measured at the daily or cumulated levels, is associated with increases in attendance at MLB games; the "League Standing Effect" may not describe the preferences of baseball fans in North America. Also, the "Louis-Schmelling Paradox", or the "classical" UOH, does not appear to be important in this setting. Instead, loss aversion and home win preference better explain observed MLB game attendance. Since loss aversion was identified long after Neale's (1964) paper was published (Kahneman and Tversky, 1979), he can not be taken to task for this omission. In any event, the concepts raised by Neale (1964) appear to have relevance for current research in sports economics.

Our reassessment suggests several future avenues for research. First, the "League Standing Effect" has a public good aspect, both as originally conceived by Neale, the "Fourth Estate Benefit," in his words, and in light of the importance of loss aversion and home win preference, which imply that individual teams would not want to play in games with uncertain outcomes, since they can draw more fans by playing in games that the home team will either win or lose with a high probability. Public goods will be under provided by profit maximizing firms, so the provision of this "Fourth Estate Benefit" deserved further attention. The lack of empirical support for the "League Standing Effect" further heightens the importance of more research on this topic. Second, like the results in Coates et al. (2014), the results here further underscore the importance of understanding why leagues would want to ensure balanced competition while individual teams might prefer less balance in order to take advantage of fan's loss aversion and home win preference. Until this tension has been addressed, the fundamental issue of the difference in teams' and leagues' incentives uncovered here will remain.

References

Bagwell, K. (2007). The economic analysis of advertising. *Handbook of industrial organization*, 3:1701–1844.

Coates, D., Humphreys, B. R., and Zhou, L. (2014). Reference-dependent preferences, loss aversion and live game attendance. *Economic Inquiry*. Forthcoming.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292.

King, N., Owen, P. D., and Audas, R. (2012). Playoff uncertainty, match uncertainty and attendance at Australian National Rugby League matches. *Economic Record*, 88(281):262–277.

Koszegi, B. and Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, 121(4):1133.

Kuypers, T. (2000). Information and efficiency: an empirical study of a fixed odds betting market. *Applied Economics*, 32(11):1353–1363.

Lahvička, J. (inpress). Using monte carlo simulation to calculate match importance: the case of English Premier League. *Journal of Sports Economics*.

Lei, X. and Humphreys, B. R. (2013). Game importance as a dimension of uncertainty of outcome. *Journal of Quantitative Analysis in Sports*, 9(1):25–36.

Levernier, W. and Barilla, A. G. (2007). An analysis of the home-field advantage in major league baseball using logit models: Evidence from the 2004 and 2005 seasons. *Journal of Quantitative Analysis in Sports*, 3(1):1–22.

Neale, W. C. (1964). The peculiar economics of professional sports. *The Quarterly Journal of Economics*, 78(1):1–14.

Rottenberg, S. (1956). The baseball players' labor market. *Journal of Political Economy*, 64(3):242–258.

Soebbing, B. P. (2008). Competitive balance and attendance in Major League Baseball: An empirical test of the uncertainty of outcome hypothesis. *International Journal of Sport Finance*, 3(2).

Štrumbelj, E. (inpress). A comment on the bias of probabilities derived from betting odds and their use in measuring outcome uncertainty. *Journal of Sports Economics*, pages 1–15.

Tainsky, S. and Winfree, J. A. (2010). Short-run demand and uncertainty of outcome in Major League Baseball. *Review of Industrial Organization*, 37(3):197–214.