THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF MARKETING

Promotions and their Dynamic Effect on Major League Baseball Attendance

MARCUS JACKSON
SPRING 2014

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Marketing
with honors in Marketing

Reviewed and approved* by the following:

Wayne DeSarbo
Distinguished Professor of Marketing
Thesis Supervisor

Jennifer Chang Coupland
Clinical Associate Professor of Marketing
Honors Adviser

* Signatures are on file in the Schreyer Honors College.

ABSTRACT

Major League Baseball (MLB) has the largest inventory of games of the four major sports leagues in the USA. In order to help increase attendance, most MLB teams employ different types of promotion to assist them. Based on a Koyck Distributed Lag Regression Model, I will demonstrate that giveaways, entertainment, price, and events promotions have current and carry over (i.e., dynamic) effects for attendance for the majority of MLB teams, and quantify these effects.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank Dr. Wayne DeSarbo, Chris Mckeon, and Ashley Stadler Blank for giving me the opportunity to work as a Research Associate in Smeal's Center for Sports Business Research. Without that opportunity, my love for the sports industry would not have been able to flourish. I would also like to thank John Powell who put me in contact with Dr. DeSarbo. Lastly, I would like to thank my friends and family who have supported me through this project.

Chapter 1

Introduction

The sports and entertainment industry is one of the largest drivers of revenue for the United States' GDP. Last year in 2013, the sports industry reached a total of $470 billion dollars in revenue (Plunkett Research 2013). The four major professional sports leagues in the USA are the National Football League (NFL), the National Basketball Association (NBA), the National Hockey League (NHL), and Major League Baseball (MLB). The MLB, which will be the focus of my thesis, generated a total of $7.7 billion dollars in revenue in 2013 (Plunkett Research 2013). One of the main drivers of that revenue is game ticket sales. Each of the 30 MLB teams plays approximately 81 homes games and a total of 162 total games. These totals far surpass those of the other three major sports leagues in this country. Each NFL team has 8 homes games and both NBA and NHL each play 41 home games. With such a large inventory of home games, it is important that MLB teams have a constant stream of fans to these games to keep their revenues stable. With that said, in order to keep a maintain flow of revenue, attendance at games needs to be constant and continuous throughout the 81 home games. The MLB had a league-wide attendance increase of almost 2% in 2012 and this increase continued into this season by increasing again by about 1.1% (Major League Baseball 2013). In addition, the league recorded the sixth highest season-long attendance ever (Major League Baseball 2013). This upward trend is encouraging as more than half of Major League Baseball's 30 teams saw drops in their respective attendance just two years ago

(Sports Business Journal 2012). This has sparked a concern for MLB because ticket sales often represent the largest revenue stream for teams (Fisher 2011). Ticket sales for some teams make up about 32% of teams' total gross revenues on average (Badenhausen et al. 2013). It can be said that ticket sales are vital to the each of the respective baseball franchises.

In order to improve the bottom line of all 30 MLB teams, it is critical that the teams understand the factors that influence fan attendance in their respective markets. How does one maintain steady streams of attendance across all the games? There are numerous variables that influence fan turnout at MLB games (DeSarbo, Blank & McKeon, 2012). This thesis will look to highlight one of the main factors that drive attendance to Major League Baseball games (in-game promotions), and from that determine if there are effects of the promotions immediately in both the short term and/or long term.

The remainder of this thesis will be organized in the following manner. In the next section, I will look at the marketing literature that pertains to how promotions have affected attendance in MLB and other sports. In Chapter 3, I will describe my research methodology. In Chapter 4, I will describe the data collected and outline the statistical model that was used to analyze the data sets. In Chapter 5, I will describe and report my findings. Lastly, in Chapter 6, I will conclude with a summary of the findings and some managerial recommendations.

Chapter 2

Promotions and Attendance

There are several factors that can be looked at to affect MLB game attendance and there have been a few papers that have looked at these factors. The factor that will be highlighted in this paper will be in-game promotions, but it is to be noted that previous research has looked at a variety of factors such as the team performance, weather, promotion, media, city demographics, pricing, the opponent played, and venue effects on the game attendance.

One prior study suggested that team performance, which is the most studied variable relation to attendance, is positively correlated to attendance at MLB games (Lemke et al. 2010). Performance covers a broad spectrum of variables such as runs, runs against, home runs, pitchers era, and winning percentage. Another factor is the competiveness of the home team's opponent which is also positively correlated to the attendance to the game. This holds true for both large market and small market teams, but is especially true for small market teams.

The most commonly studied variables in regards to weather have concerned forms of precipitation and temperature (Lemke et al 2010). But most literature on weather and its implications on attendance have conflicting findings with some showing an adverse effect on attendance, while others don't show any impacts on attendance (Trail et al 2008).

Media will be defined as any medium other than the live event itself that gains you access to the game which include television, radio, and internet streaming. Just like weather, literature on media has conflicting findings with some showing a positive relationship (Kaempfer & Pacey 1986), a negative relationship (Baimbridge et al 1996), and no significant relationship (Hill et all 1982).

Promotions in the MLB are utilized to increase same game day attendance, but with the goal of having influence on the future attendance of games (i.e., repeat purchase). McDonald and Rascher (2000) were the first to examine in detail the different types of promotions utilized in the MLB and their effect on attendance. They specifically looked at what they called sale promotions, which they defined as a diverse set of incentive tools mostly for the short term (McDonald and Rascher 2000). In addition, they delineated between two types of sale promotions: price vs. nonprice promotions. Examples included all inclusive ticket deals, t-shirt giveaways, bobble head, and etc. A significant positive relationship has been established between attendance and in-game promotions that the teams implement (Lemke et al. 2010). Several studies show increases in the average attendance to up to 20% when a promotion is incorporated at MLB games (McDonald and Rascher 2000; Boyd and Krehbiel 2003). It is suggested that when a promotion is implemented in a game, it provides utility, or additional value, to the games that typically have lower demand, which historically have included games during the week, when the home team's performance is poor, or when the attractiveness of the home team's opponent is low (Boyd and Krehbiel 2003). Note, most of the previous research that has been conducted has focused on the factors that affect attendance that the management of the teams has no direct control over such as weather, team performance, city demographics, presence of stars, and game attractiveness based on the opponent (Domazlicky and Kerr 1990; Madura 1981; Madura 1982). The promotional activities that a team implements are the only factor of attendance that management can control directly. Because they can control the promotion, management can to some to extent affect the satisfaction and utility that one receives. This satisfaction can lead to positive word of mouth (Westbrook 1987), purchase reinforcement, and repeat attendance. More recent research has begun to inspect three factors that management can control related to promotions: the timing of promotions, number of promotions run, and the type of promotion. Boyd & Krehbiel (1999) first examined the effect of promotion timing on MLB attendance. They used data from over four seasons for six teams they studied. They examined the effect of running promotions based on day of the week. Despite the fact that they used a small sample for their study, it showed that the timing of promotions can influence their effectiveness, in addition to the fact that controllable factors by management can make influence the effectiveness of a promotion. A key insight that can be noted is that they found the impact of a promotion was the strongest when it was run during a week day in the daytime (Boyd & Krehbiel 1999).

In another study that looked at the effect of forty independent variables, McDonald and Rascher (2000) centered their attention on whether diminishing marginal returns existed when adding additional promotional items. It was found that promotions caused an average increase in attendance of 14% in the 1996 MLB season (which then consisted of twenty-nine teams). In addition, diminishing returns were found. In a later study, no evidence was present that supported diminishing marginal returns, but that discrepancy could be attributed to a different sample (Boyd & Krehbiel 1999). A 2003 article published by Boyd and Krehbiel claimed that diminishing marginal returns possibly could have occurred because of the different promotions that were run. As well, they reasoned that promotions were being implemented to games that were already seen as attractive by the fans, which contributed to the diminishing returns.

Because of this, it is no wonder as to why baseball management spends so much money on implementing a promotion schedule and using different promotions. Teams typically use promotions that come in the form of give-a-ways such as t shirts, bobble heads, schedule magnets, and other items. Teams also engage in ticket discounts, typically family oriented or all inclusive package deals. The all-inclusive deals would include food, drink, and ticket all in one price. Promotions have also occurred in the form of pregame or post events and entertainment. These events and entertainment typically included firework shows, tailgates with vendors and food, and concerts that occurred post game. It is important to note that attendance can be negatively influenced if a promotion is run too frequently (McDonald and Rascher 2000).

Chapter 3

Methodology

My research goal is to examine if promotions affect recent MLB attendance. Based on how promotions impact MLB attendance in previous studies, specific promotion items/variables will be organized into one of four different promotion categories: Giveaways, Entertainment, Events, and Price Promotions. Table 1 below shows some examples of the type of promotions of the teams employed.

Types of Promotions Implemented
Kids Run the Bases
Firework Shows
Bark in the Park
Military Appreciation Day
Senior Day
College Night
Bobblehead giveaway
Alumni Autograph Day
T shirt Giveaways
Scouts Night
Heritage Themed Carnivals
Family Sunday
$1 Dog Night
Kids Fun Day
Rally in the Alley
Charity Cuts
Replica Jersey
Rally Towel Giveaway

1 Types of Promotions Implemented

The goal is to quantify the impact that each category promotion has with respect to its current, carry-over/dynamic, and total effect on attendance. Current effect is defined as the effect that the promotions have on the present game attendance. Carry-over effect is defined as the effect promotions have on future attendance. The total effect is the combination of the current and dynamic effects of promotion. The MLB has defined total game attendance based as the number of tickets sold for each game (including tickets purchased by season and single-game ticket holders). (Here, the actual attendance was not used due directly in the analysis to the fact that there would be bias based on different stadium sizes.)

The data used for this study was collected from the complete 2013 MLB baseball season. On a daily basis, data was recorded on all the games that were played throughout the season by each team. The data that was collected involved hundreds of different variables measuring all the different factors affecting attendance including team performance, opponent, day of the week, month, type of promotion that was used if there was one, weather, media, etc.. Each of the thirty teams played approximately 81 home games. So in total, data was collected from approximately 2430 games (81 multiplied by 30). Note, some teams did not play complete games because they were postponed. All data was collected from each teams' respective promotional schedule that was found from their respective websites.

Two multiple regressions were then completed for the aggregate of all the teams, and then for each team individually to determine the effects of promotion. The first regression that was run examined the current effects of promotions. This model will be called the *null* model and it is denoted as:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \ldots \beta_p X_{pt} + \text{error term},$$

Where Y is adjusted proportion of stadium capacity filled (to remove stadium and city size) at home game t, and X_{jt} for j=1,2,3,4 are the four independent variables/categories of promotion denoting whether or not that category of promotion occurred at home game t.

. The second regression model was an expansion of the first in that the second model not only looked to capture the current effect of promotion, but it also added a lagged dependent variable term as in the Koyck model (described in the next section) to test if there were both current and carry over effects present. The alternative, second regression will be denoted as:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \ldots \beta_p X_{pt} + \lambda\, Y_{t-1} + \text{error term}$$

and named the *extended* model.

Next, an F distribution nested models test was used to compare these two aggregate models to see which one model was overall more effective. In order to the complete the test the following test was used:

$$F = (SS(Error)_n - SS(Error)_e)/\, s \div SS(Error)_e \,/\, (n\text{-}p\text{-}s\text{-}1),$$

where: $SS(Error)_n$ is the sum-of-squares error of the null model and $SS(Error)_e$ is the sum-of-squares error of the extended alternative model. N denotes the number of observations which was 2430 based on all the games that were played in the regular season. P is the number of categories of promotional explanatory variables which is 4. S denotes the period of lag which is 1. The calculated F from the equation was compared to the critical F value ($p < .05$). Based on this test, an assumption was made that the better model would be used for the rest of the thirty teams.

Lastly, after running the best regression for each team, a hierarchical cluster analysis based on the Ward method was used to group the 30 teams into Clusters. From the resulting dendogram, three cluster were formed. Following the clustering, an analysis of variance was completed based on the descriptives and means plots to reflect the resulting patterns that existed within the different clusters with discussion. All analyses were completed using the SPSS data package.

Chapter 4

The Koyck Distributed Lag Model

In order to model the relationship between promotions and MLB attendance and examine current and dynamic effects of promotion, a Koyck distributed lag model was used. The Koyck distributed lag models are representations in which a regression equation is utilized to predict present values of a dependent variable based on independent variable's current values and lagged values of the dependent variable. A basic Koyck distributed lag model is represented as:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots\ldots\ldots + \lambda Y_{t-1} + u_t$$

Y_t is the value at time period *t* of the dependent variable *y*, α is the estimated intercept term, and β_0 is an estimated lag weight that is put on the value of the number of periods that come before the explanatory variable that is denoted as the X term. Lastly, u_t denotes the error term. In general, the independent or explanatory variable can be lagged infinitely or for a definitive amount of time. If it is lagged infinitely, it is called infinite distributed lag model, while models that have definite number of lag weights are called finite distributed lag models.

For this paper, a special kind of infinite lag model called a Koyck Model was used because it is anticipated that the effect of a promotion will carry over with geometric decay. This type of model is very restrictive in the sense that in its lag structure, the lag weight of the predictor variables exponentially decline. It is assume the weight or the restrictions follow a relation such that:

$$\beta_i = \beta_0 \lambda^i$$

It is assumed that λ is between 0 and 1 for the Koyck assumption to hold. λ is also known as the rate of decay, which causes the lag variable effect to decline exponentially. The infinite lag version of the model (assume one independent variable) can be rewritten to fit the Koyck specifications as:

$$Y_t = \alpha + \beta_0 X_t + \beta_0 \lambda X_{t-1} + \beta_0 \lambda^2 X_{t-2} + \ldots\ldots + u_t .$$

Subtracting the expression for Y_{t-1} from both sides and simplifying, you end up with:

$$Y_t = \alpha + \beta_0 X_t + \lambda Y_{t-1}.$$

In this model, the short-run (current) effect of explanatory variable is the value of β_0, while the long-run (carry-over) effect in the independent variable can be shown to be:

$$\beta + \lambda\beta + \lambda^2\beta + \ldots = \beta_0/(1-\lambda).$$

Here, λ is the carry- over effect of the promotion (Kappe, Stadler Blank, & DeSarbo, 2014).

Infinite distributed lag models do not undergo certain drawback that definite lag models do. First the lag lengths need to be defined before any estimation. Next, with an increase in lag length, fewer degrees of freedom exist. Also, multicollinearity, the occurrence in a multiple regression question where two or more explanatory variables are highly correlated to each other, is not a major drawback of such Koyck models unlike infinite lag models in general.

Chapter 5

Results and Discussion

Aggregate

Tables 2 and 3 show the regression results for the aggregate *null* model. This model looks at how *current* promotion expenditures had an impact on adjusted % of stadium capacity attendance.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.323[a]	.104	.103	.90454

a. Predictors: (Constant), Giveaway, Entertain, Events, Price

2 Null Result 1

Coefficients[a]

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	.498	.029		17.240	.000
Price	-.164	.029	-.111	-5.733	.000
Events	.108	.025	.082	4.276	.000
Entertain	.298	.046	.126	6.536	.000
Giveaway	.460	.036	.249	12.885	.000

a. Dependent Variable: DepVar

3 Null Result 2

These results indicate that the *null* model has an R-square of .104 meaning the model explains about roughly 10% of the variation in the dependent variable amongst these 2430 observations. The t-tests show that all the independent variables are significant. Note, giveaways have the strongest impact on % capacity of attendance. Each variable is statistically significant based on a .05 alpha. The constant term is also significant. Because this constant is significant, it denotes that without the explanatory variables of promotion, people would still attend the baseball games based on all the games that each team played throughout the season. Unfortunately, what makes little sense here is the fact that price discounts have a negative impact on attendance. This might reflect the fact that such discounts are given during games during the week day when playing less desirable opponents.

There is a major limitation with this *null* model. An R-square of .104 is rather low, which may mean the model may not be a very good predictor of what is actually happening. Tables 4 and 5 reflect the regression results of aggregate *extended* model based on the Koyck model.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.728[a]	.530	.529	.63983

a. Predictors: (Constant), LagDV, Entertain, Events, Price, Giveaway

4 Extended Result 1

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.069	.022		3.089	.002
	Price	-.036	.020	-.025	-1.759	.079
	Events	.040	.018	.032	2.242	.025
	Entertain	.364	.032	.157	11.228	.000
	Giveaway	.249	.026	.137	9.600	.000
	LagDV	.660	.014	.675	47.076	.000

a. Dependent Variable: DepVar

5 Extended Result 2

The *extended* model shows a R-square of .530, which means that the model explains about 53% of the variation in the dependent variable over these 2430 observations. Looking at the unstandardized β's, the lagged DV variable has the strongest effect on the model compared to the other explanatory variables (promotions) based on p<.05. In addition, the variables of events, entertainment, and giveaways are also significant. It should be noted that looking at the two models there are two major differences. First, the comparative R squares: The extended model has a significantly higher R-square value. Next, the unstandardized beta for price remained negative, but is not significant at p<.05. But, all the other promotional variables are still statistically significant and positive.

In order to test which model (null vs. extended) is better, an F distribution test was completed to compare overall effectiveness. The calculated F test statistic was 2483.24. This value was compared to the critical value of 3.84 based on the appropriate degrees of freedom. Because 2483.24 is so much larger than 3.84, it can be said that the extended model is the superior model. Thus, all further discussions on the teams will be based on the extended model because of the results of this F distribution test.

Based on the estimated extended model, it is predicted that there are both significant current and carry over effects of the following promotion expenditures: giveaways, events, and entertainment. This means that overall for the aggregate across all the 30 teams, the money they use on these three promotion categories is predicted to have a positive impact on their attendance for the present and the future. Price discounts were not significant in the model and its unstandardized beta is negative. The lagged DV has an unstandardized beta of .660 at $p< .05$ significance level and is the most significant variable in the model. This could mean that the pricing promotion expenditure is only good for when the team is trying to fill the stadium for an unattractive game based on opponent or day of the week. The constant term in the aggregate lag model is also significant and has a positive unstandardized beta. This means that across all the teams and holding the other explanatory variables at zero, i.e., if the teams as a whole did not spend any money on promotions, one can expect people to still show up and attend the games.

These results reflect the general trend, but there are differences amongst the 30 individual MLB teams in regards to how promotions drive attendance. So, to examine this team heterogeneity, I ran separate extended model regressions for each of the 30 teams. Table 6 below shows the standardized betas from those 30 regressions. In Table 6, the numbers that are bold denotes that the coefficients are significant based on $p<.05$, while the numbers that are not bold are not.

Team	Price	Ent	Give	Events	Lag DV	F Anova
ARI	-0.095	**0.317**	**0.367**	**0.226**	**0.167**	12.765
ATL	0.145	0.535	**0.774**	**0.187**	**0.321**	17.376
BAL	-0.19	**0.828**	0.309	0.024	**0.47**	7.028
BOS	-0.226	0	0	0.01	**0.56**	20.169
CHC	0.037	0	**0.25**	-0.182	**0.289**	5.206
CHW	0.041	**0.37**	0.297	0.162	**0.213**	5.534
CIN	-0.046	**0.573**	**0.454**	0.398	**0.497**	16.698
CLE	0.294	**0.451**	-0.119	0.089	**0.342**	20.088
COL	**-0.408**	**0.879**	0.055	**0.231**	0.219	20.07
DET	-0.153	**0.653**	0.058	**0.244**	**0.371**	17.918
HOU	-0.218	0.212	0.12	0.153	**0.434**	7.875
KCR	0.128	**0.723**	0.294	0.057	0.137	9.585
LAA	0.075	**0.349**	**0.211**	-0.006	0.092	5.291
LAD	0	**0.288**	**0.421**	-0.226	**0.308**	6.238
MIA	**0.222**	**0.252**	-0.002	**0.12**	**0.363**	6.434
MIL	0.081	0	**0.56**	0.272	**0.186**	6.584
MIN	0.089	0.304	**0.229**	0.164	**0.44**	6.565
NYM	-0.127	**0.97**	0.148	0.128	0.128	7.015
NYY	0	0	0.02	0.225	**0.481**	8.756
OAK	-0.398	**1.731**	0.483	0.301	0.079	7.013
PHI	-0.023	0.45	0.126	0.04	**0.453**	6.528
PIT	0.273	0.052	**0.964**	**-0.522**	**0.545**	28.58
SDP	-0.192	0.36	**0.723**	0.049	**0.474**	11.13
SEA	-0.076	0.346	0.199	0.191	**0.391**	5.932
SFG	**0.18**	0.057	**0.042**	-0.008	0.207	3.684
STL	**-0.296**	0.066	**0.256**	0.076	**0.491**	9.759
TBR	**0.36**	0.287	**0.844**	**0.514**	0.136	22.869
TEX	-0.047	**0.568**	**0.514**	**0.356**	**0.22**	7.127
TOR	-0.236	0.127	**0.676**	**0.423**	**0.301**	7.995
WSN	0.134	0.186	**0.433**	**0.167**	0.146	5.362

6 Team Unstandardized Coefficients

Based on the cluster analysis of the unstandardized betas from each team's regression analysis (see Table 6) using the extended Koyck model, the 30 teams were split into three clusters based on the resulting dendogram (Appendix A). A dendogram is a tree like structure that is most commonly derived in hierarchical clustering. The Ward method of hierarchical clustering was utilized here to summarize this heterogeneity among the 30 teams as explained by the regression results of each teams' extended regression coefficients.

Across the three derived clusters, the unstandardized regression coefficients for price, entertainment, events and the lagged dependent variables were statistically significant between these three clusters based on an Anova test using a $p<.05$ s (see Table 7). The effect of giveaway promotions between the groups was not significant at .05.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Price	Between Groups	.160	2	.080	3.669	.039
	Within Groups	.589	27	.022		
	Total	.749	29			
Ent	Between Groups	.641	2	.321	39.848	.000
	Within Groups	.217	27	.008		
	Total	.858	29			
Give	Between Groups	.166	2	.083	3.288	.053
	Within Groups	.681	27	.025		
	Total	.847	29			
Events	Between Groups	.183	2	.092	6.393	.005
	Within Groups	.387	27	.014		
	Total	.570	29			
LagDV	Between Groups	.396	2	.198	23.172	.000
	Within Groups	.231	27	.009		
	Total	.627	29			

7 Cluster Result 1

The means of the groups were compared, based on means plots and other descriptives to see which cluster showed stronger effects of each respective explanatory variable. Table 8 exemplifies these strengths of the different clusters. High denotes that the variable had a high mean (effect) on that cluster. Low indicates that there was low mean (effect) on the cluster. The means plots and descriptives of the three clusters can be seen in Appendix A. As noted in the means plot, each cluster expressed their own respective patterns. Table 9(See Appendix A) shows all the teams and in which cluster they belong to. Based on means plots and the descriptives, Cluster 1 showed the strongest effect of giveaways and events. Cluster 2 showed the strongest effects of price and entertainment promotions, while Cluster 3 showed the strongest effects of the carry-over effects. The remaining sections will breakdown each cluster and explore why they show these patterns.

Column1	Column2	Column3	Column4	Column5	Column6
	Price	Entertainment	Giveaways	Events	LagDV
Cluster 1	High	Low	High	High	low
Cluster 2	High	High	Low	High	low
Cluster 3	Low	Low	Medium	Low	High

8 Cluster Strengths

Cluster 1

The teams that are included in cluster 1 are as follows: Arizona Diamondbacks, Milwaukee Brewers, New York Mets, San Francisco Giants, Tampa Bay Rays, Toronto Blue Jays, and the Washington Nationals. This cluster of teams showed a pattern of having high means in the categories of price, giveaways, and events promotions meaning they showed strong effects in these areas. But at the same time they showed low means entertainment and carry over effects. The mean for price was .08171 but did not significantly differ from the mean of price of the other two clusters based on $p<.05$. Cluster 1 exemplified the highest means in giveaways and events categories in relation to the other clusters with means of .37814 and .19329 respectively. Based on this pattern it is predicted that for the teams that compose cluster one, their fans are driven to attend the games because of the giveaways that are implemented. As a whole, Cluster 1 had the lowest mean, at .19671, for lagged DV or carry-over effects in relation to the other two clusters. According to the post hoc test (see Appendix A), this difference was not significantly different when compared to Cluster 2, but was very significant when compared to Cluster 3. In other words, there are current effects of giveaway promotions for Cluster 1, but the implementation of said promotion does not have implications towards future attendance.

Cluster 2

The teams that compose cluster 2 are as follows: Atlanta Braves, Chicago White Sox, Cleveland Indians, Colorado Rockies, Detroit Tigers, Kansas City Royals, Los Angeles Angels, Miami Marlins, Oakland Athletics, and the Texas Rangers. Cluster 2 showed a pattern of possessing high means for price, entertainment, and events, while showing low means for both giveaways and carry over effects. As mentioned previously, the means of price did not significantly differ amongst the three clusters, but Cluster 2 showed the highest mean for price at .09150. Cluster 2 also expresses the highest mean of entertainment at.43160, but the second lowest mean of carry over effects at .25250. Based on these findings as whole, it is predicted that the teams for Cluster 2 possessed current effects of entertainment and pricing promotions with no implications towards future effects. It is predicted that the management potentially are attempting to market their teams and venues as an entertainment alternatives to the more traditional means of entertainment such as amusement parks, movie theaters, and etc. Even though price was not significantly differently between the three clusters, Cluster 2 still had the highest mean. This could be interpreted that fans of these teams are more sensitive to price than the fans of the teams in the other clusters. It was noted that Cluster 2 had the highest mean for entertainment when compared to the other two clusters, and the difference was significant based on $p<.05$ when looking at the post ad hoc multiple comparisons test. Despite their respective team's performance, management of the teams that belong in this cluster could be pushing their ball games to be an inexpensive entertainment alternatives.

Cluster 3

The teams that make up Cluster 3 are the Baltimore Orioles, Boston Red Sox, Chicago Cubs, Cincinnati Reds, Houston Astros, Minnesota Twins, New York Yankees, Philadelphia Phillies, Pittsburgh Pirates, San Diego Padres, Seattle Mariners, and the St. Louis Cardinals. This cluster displayed a pattern where each explanatory variable possessed a low mean low except for the carry over effects variable. The lagged DV, carry-over effect variable had a mean of .45762, which was significantly higher than the means for Clusters 1 and 2 based on $p<.05$. From this, it is predicted that promotions expenditures are not very effective for these teams in the short run. What is a potentially reason or cause behind the fact the promotions are ineffective?

Two reasons are history and performance. All the teams across the clusters have a history, but most that compose Cluster 3 have a special history/tradition that is based on performance. These teams either have historically performed exceptionally well in the MLB or have historical performed poor. If a team performs exceptionally well for an extended period of time, people are going to attended games. It was noted previously that performance has the strongest correlation to attendance. In this instance, that means even if the teams implement promotions, people will still purchase tickets and attend games. Promotions in this instance will still be used but most likely for games that are seen as unattractive. On the flipside, when a team performs poorly, promotions are needed to help push ticket sales and drive attendance. The promotions may have an effect in the short term, but not in the longer term.

Chapter 6

Conclusion

Promotions are predicted to possess positive current and carry over effects on MLB attendance. They can provide that extra incentive to attend a game especially when the home team's opponent is unattractive or when the game the game is scheduled at an unattractive time. Each team is unique based on the results of the regression and cluster analyses. Some teams need the assistance of promotions in order to drive attendance, while for other teams promotions are not essential.

Note, promotions are not the only factor that may influence attendance. Further multivariate studies should look to see how other factors in combination with promotion impact attendance. Performance, as the previous literature has shown, is positively correlated to attendance. Other potential variables that are worthy of investigation include the distance between the two venues, demographics of the home team's city, weather, day of the week, time of day, and whether the game is televised. A study combining all of these variables would provide more insight into what combinations of effects, if any, driving attendance. Further analysis of the effects of promotions can lead to greater efficiency in promotion scheduling and cost savings on promotions expenditures. (DeSarbo, et. al, 2014).

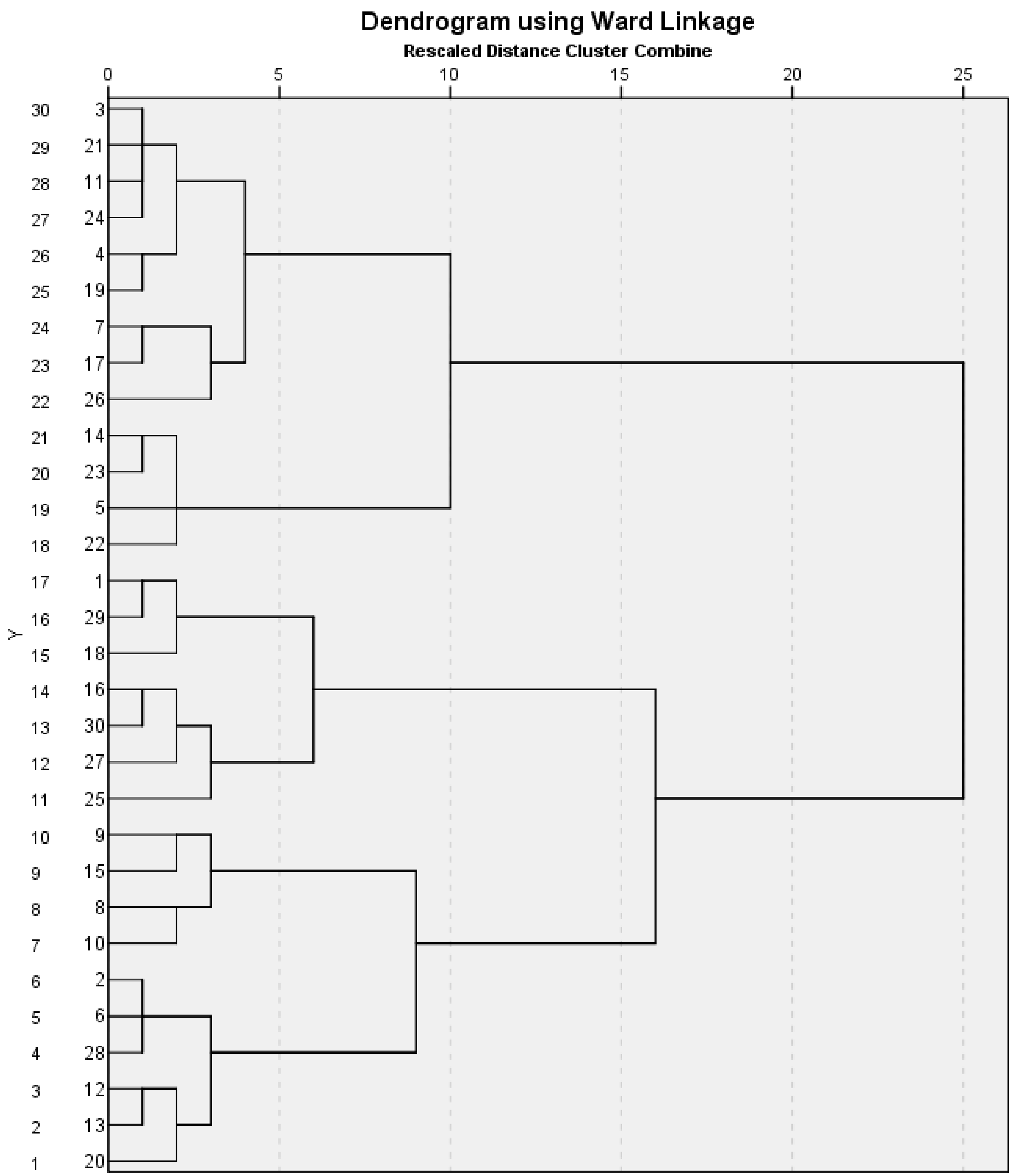
Appendix A

Column1	Column2
Team	**Cluster**
ARI	1
ATL	2
BAL	3
BOS	3
CHC	3
CHW	2
CIN	3
CLE	2
COL	2
DET	2
HOU	3
KCR	2
LAA	2
LAD	3
MIA	2
MIL	1
MIN	3
NYM	1
NYY	3
OAK	2
PHI	3
PIT	3
SDP	3
SEA	3
SFG	1
STL	3
TBR	1
TEX	2
TOR	1
WSN	1

9 Cluster Divisions

The number across the team abbreviation denotes what cluster the team is in.

Figure 1 Dendogram



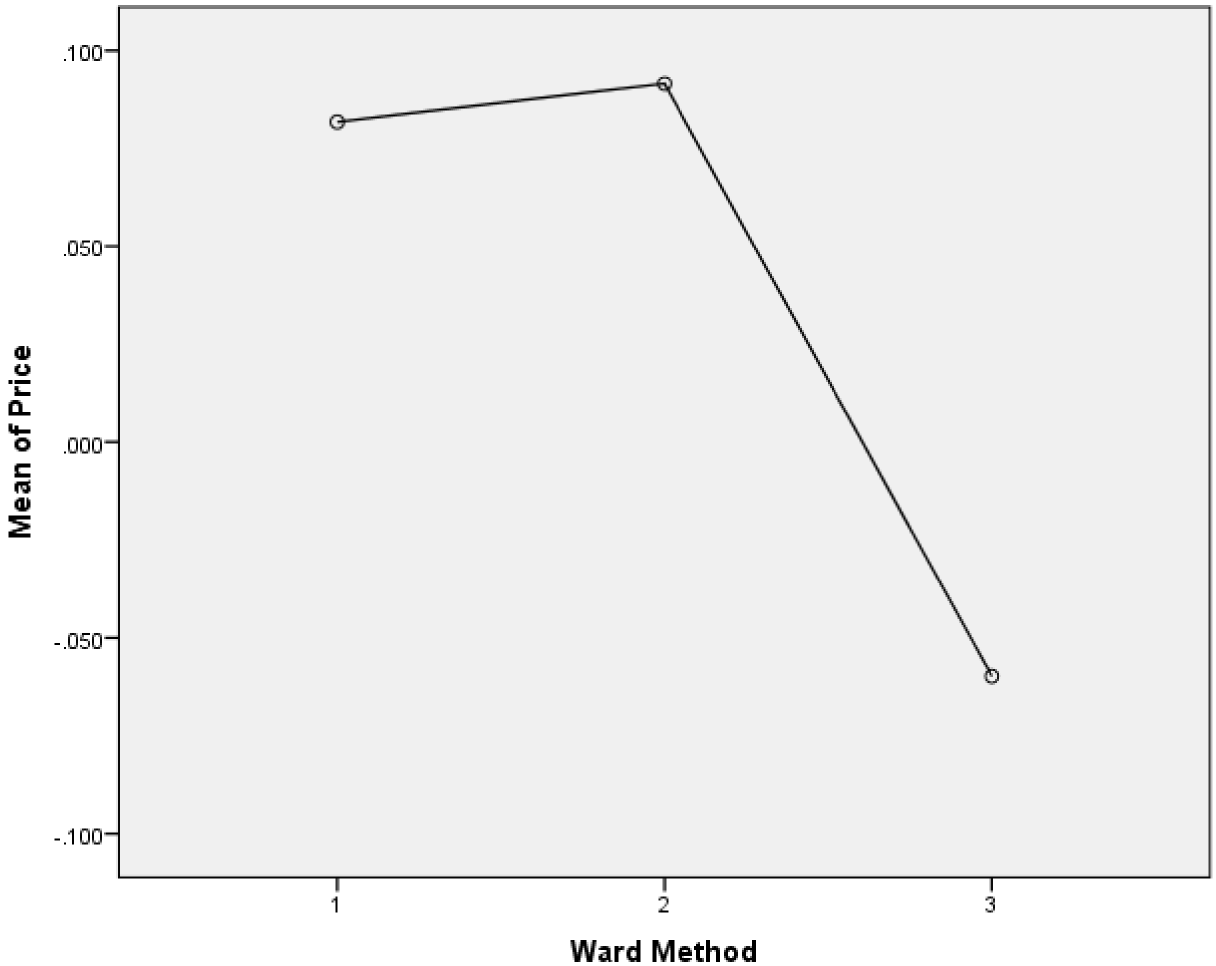
10 Descriptives

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
Price	1	7	.08171	.188179	.071125	-.09232	.25575	-.118	.294
	2	10	.09150	.163767	.051788	-.02565	.20865	-.155	.424
	3	13	-.05977	.106162	.029444	-.12392	.00438	-.280	.110
	Total	30	.02367	.160730	.029345	-.03635	.08368	-.280	.424
Ent	1	7	.11300	.082531	.031194	.03667	.18933	.000	.257
	2	10	.43160	.073727	.023314	.37886	.48434	.268	.534
	3	13	.12638	.103035	.028577	.06412	.18865	.000	.310
	Total	30	.22500	.172032	.031409	.16076	.28924	.000	.534
Give	1	7	.37814	.128380	.048523	.25941	.49687	.222	.580
	2	10	.18060	.168189	.053186	.06029	.30091	-.075	.459
	3	13	.23531	.165255	.045834	.13545	.33517	.000	.544
	Total	30	.25040	.170910	.031204	.18658	.31422	-.075	.580
Events	1	7	.19329	.113278	.042815	.08852	.29805	-.030	.307
	2	10	.17700	.103280	.032660	.10312	.25088	-.009	.281
	3	13	.02646	.133531	.037035	-.05423	.10715	-.193	.186
	Total	30	.11557	.140221	.025601	.06321	.16793	-.193	.307
LagDV	1	7	.19671	.060046	.022695	.14118	.25225	.125	.309
	2	10	.25250	.120623	.038144	.16621	.33879	.082	.413
	3	13	.45762	.080703	.022383	.40885	.50638	.293	.559
	Total	30	.32837	.147015	.026841	.27347	.38326	.082	.559

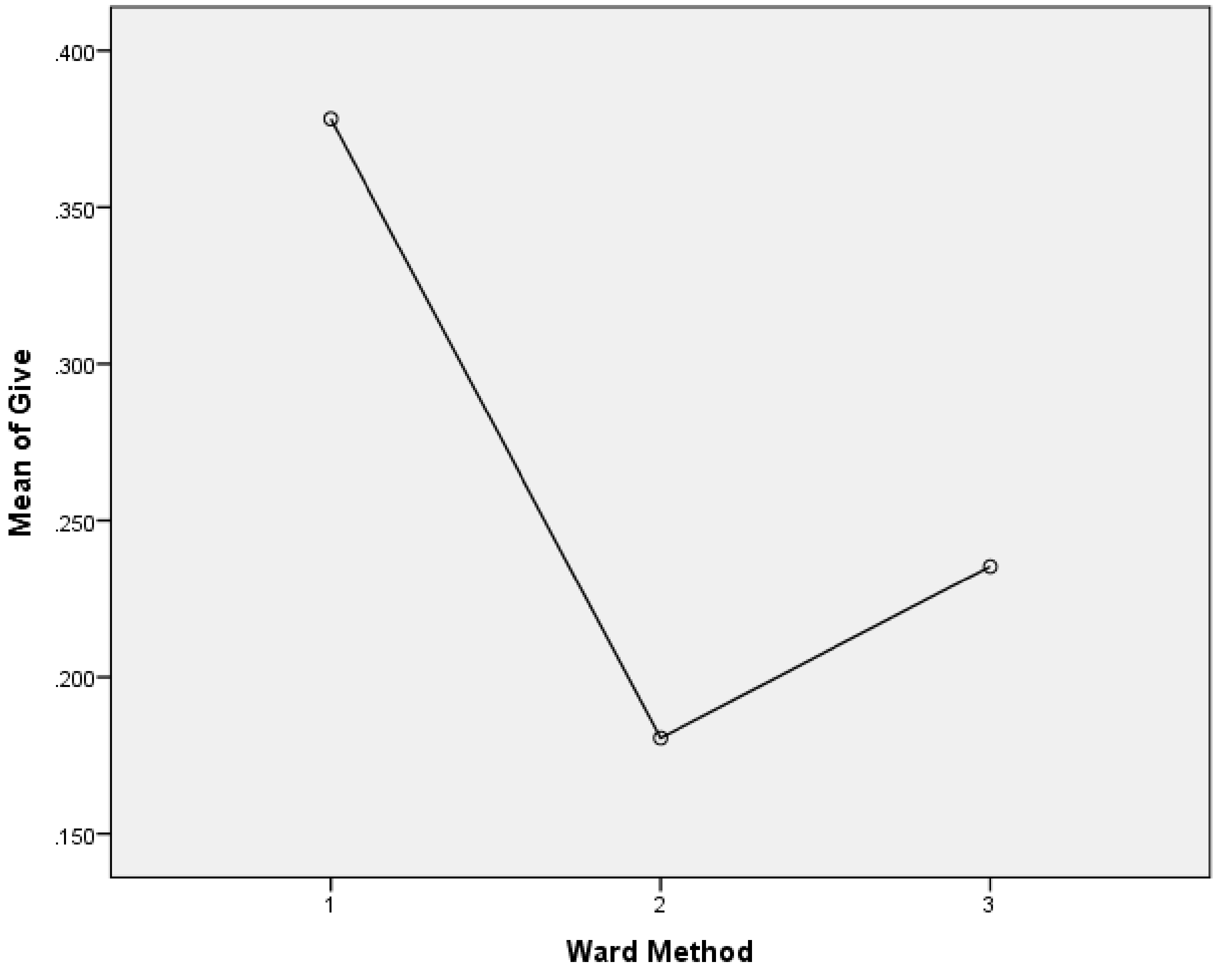
Numbers 1-3 in column denote the cluster number.

Figure 2 Means Plot Price



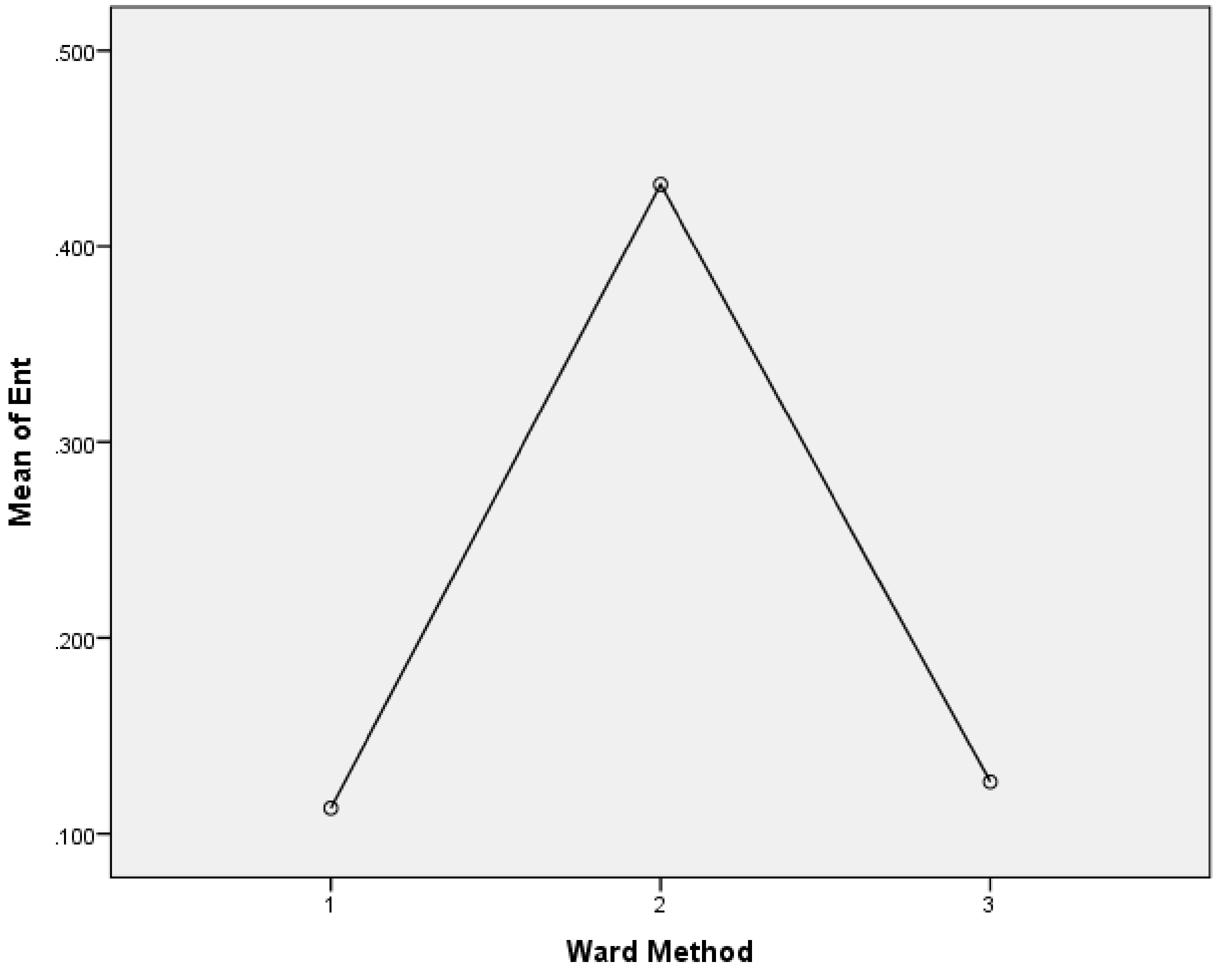
Numbers on the X axis denote the cluster number

Figure 3 Means Plot Giveaways



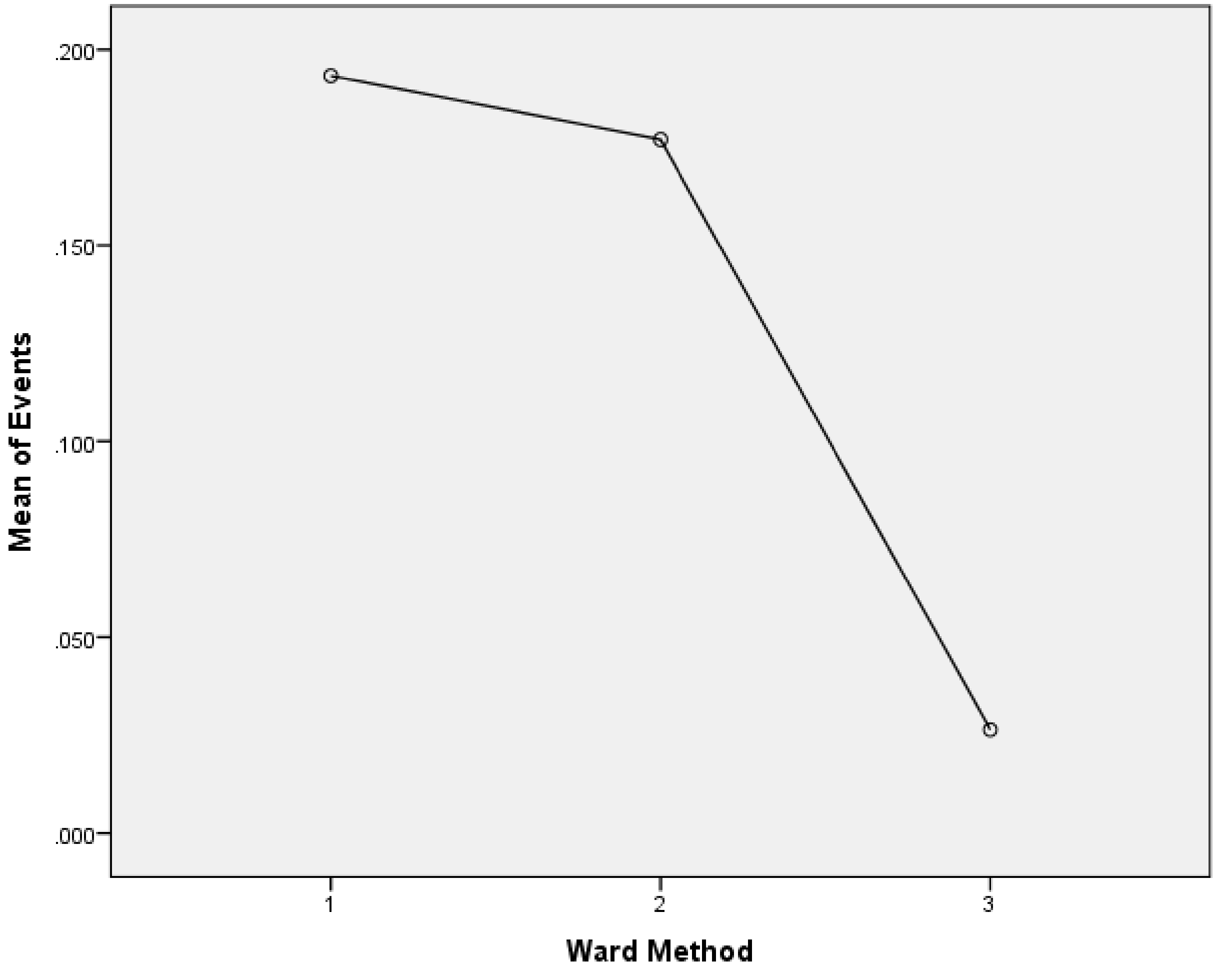
Numbers on the X axis denote the cluster number

Figure 4 Means Plot Entertainment



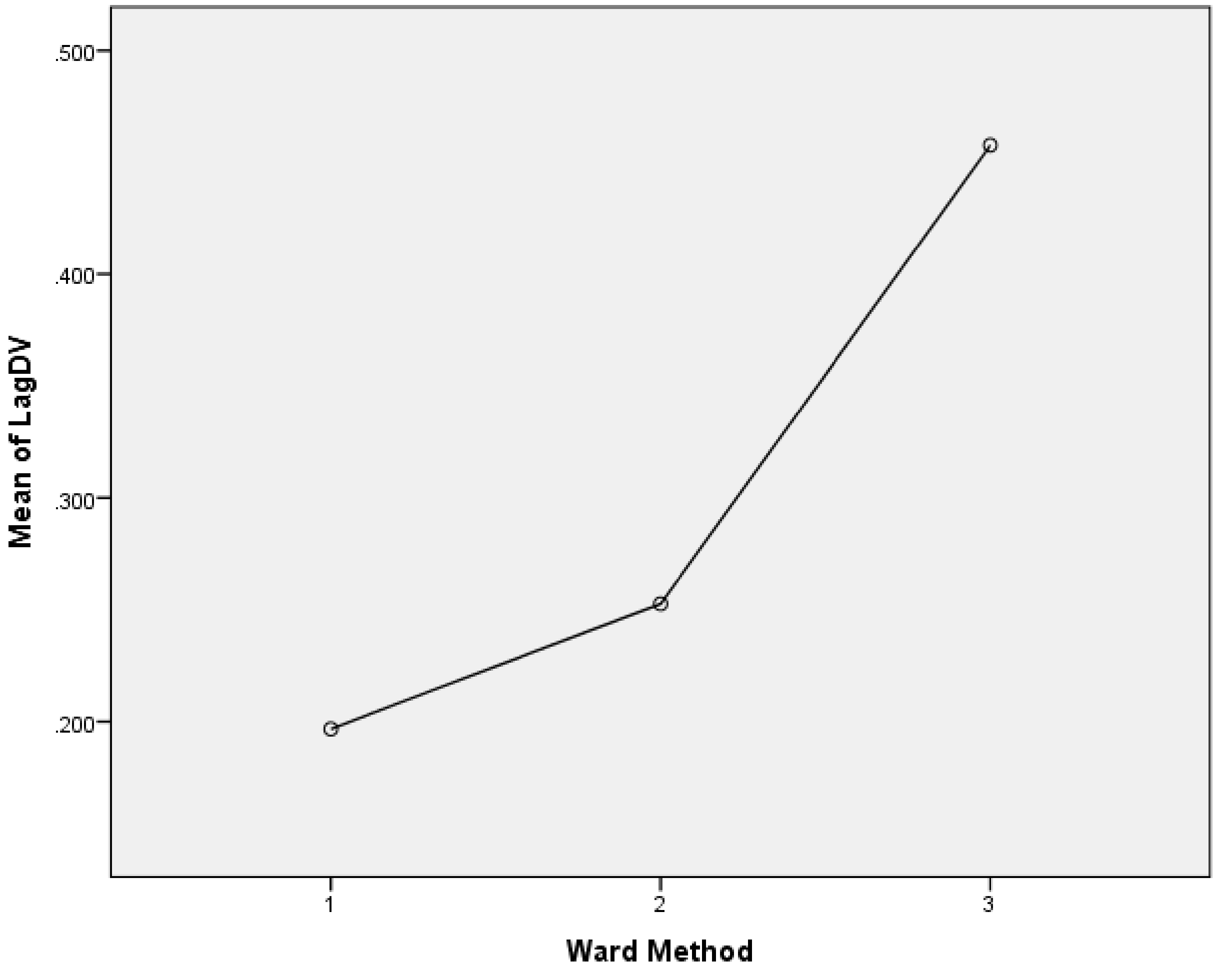
Numbers on the X axis denote the cluster number

Figure 5 Means Plot Events



Numbers on the X axis denote the cluster number

Figure 6 Means Plot Carryover Effects



Numbers on the X axis denote the cluster number

BIBLIOGRAPHY

Badenhausen, K., Ozanian, M., and Settimi, C. (2013), "Mlb team values: The business of baseball," *Forbes* [online], March 27. Available at *http://www.forbes.com/mlb-valuations/list/*.

Baimbridge, M., Cameron, S., and Dawson, P. (1996), "Satellite television and the demand for football: A Whole New Ball Game?" *Scottish Journal of Political Economy*, 43, 317-333.

Boyd TC, Krehbiel TC (1999) The effect of promotion timing on major league baseball attendance. *Sport Marketing Quart.* 8(4):23-34.

Boyd, T. C., and Krehbiel, T. C. (2003), "Promotion timing in major league baseball and the stacking effects of factors that increase game attractiveness," *Sport Marketing Quarterly*, 12, 173-183.

DeSarbo, W. S., Stadler Blank, A., and McKeon C. (2012), "Proper mix of promotional offerings can produce for teams," *Sports Business Journal* [online], May 14. Available at *http://www.sportsbusinessdaily.com/Journal/Issues/2012/05/14/Opinion/From-the-Field-of-Research.aspx*.

DeSarbo, W., Hwang, H., Stadler Blank, A., & Kappe, E. (2014). Constrained extended stochastic constrained redundacnt analysis. *Psychometrika*, (forthcoming),

Domazlicky, B. R., & Kerr, P. M. (1990). Baseball attendance and the designated hitter. *The American Economist*, *34*, 62-68.

Fisher, E. (2011), "Mlb breaks three-year attendance decline," *Sports Business Journal* [online], Oct. 3. Available at *http://www.sportsbusinessdaily.com/Journal/Issues/2011/10/03/Leagues-and-Governing-Bodies/MLB-attendance.aspx?hl=MLB%202010%20REVENUE%20RECORD&sc=0*.

Gitter, S. R., and Rhoads, T. A. (2010), "Determinants of minor league baseball attendance," *Journal of Sports Economics*, 11, 614-628.

Hill, J. R., Madura, J., and Zuber, R. A. (1982), "The short run demand for major league baseball," *Atlantic Economic Journal*, 10, 31-35.

Kappe, E., Stadler Blank, A., & DeSarbo, W. S. (2014). A general multiple distributed lag framework for estimating the dynamic effects of promotions.*Management*, *forthcoming*, Lemke, R. J., Leonard, M., and Tlhokwane, K. (2010), "Estimating attendance at major league baseball games for the 2007 season," *Journal of Sports Economics*, 11, 316-348.

Madura, J. (1981) Factors affecting attendance at sunbelt Major League Baseball games. *Texas Business Review* 55:52-54.

Major League Baseball (2013), "MLB's 2013 regular season attendance ranks as sixth highest ever," MLB Press Release [online], Oct. 1. Available at http://mlb.mlb.com/news/article.jsp?ymd=20131001&content_id=62282120&vkey=pr_mlb&c_id=mlb

McDonald M, Rascher D (2000) Does bat day make cents? the effect of promotions on the demand for major league baseball. *J. Sport Management* 14(1):8-27.

Plunkett Research, Ltd. (2013), "Sports Industry Overview," available at *www.plunkettresearch.com/sports-recreation-leisure-market-research/industry-statistics*

Trail, G. T., Robinson, M. J., and Kim, Y. K. (2008), "Sport consumer behavior: A test for group differences on structural constraints," *Sport Marketing Quarterly*, 17, 190-200.

Westbrook, RA (1987) Product/consumption-based affective responses and postpurchase processes. *J. Marketing Res.* 24(3):258-270.

ACADEMIC VITA

Marcus Jackson
134 E Foster Ave
State College, PA 16801
Mmj5127@gmail.com

Education

Schreyer Honors College, 2014

The Pennsylvania State University, 2014

St. John's College High School, 2010

Honors and Awards

Bunton Waller Academic Scholarship, 2010-2014

Schreyer Honors College Scholar Ship, 2010-2014

Association Memberships/Activities

American Marketing Association, The Pennsylvania State University Chapter, 2011-2014

Professional Experience

2012	Athletics Marketing Intern, American University
2012-2014	Basketball Marketing Intern, The Pennsylvania State University
2013-2014	Undergraduate Research Associate, The Pennsylvania State University
2013-2014	Marketing Intern, Learfield Sports, Penn State Sports Properties