Paper CI03-2012

# Survival of the Milkman – How Modern Analytics is Helping an Old Fashioned Business Thrive

Bruce Bedford, Ph.D.
Oberweis Dairy, North Aurora, IL

## ABSTRACT

The attainment of customer loyalty is not the end-goal, but it does strongly correlate with building a profitable business. In our current hyper-competitive consumer environment business survival often depends on earning the loyalty of customers. The value of analytics to improve customer loyalty is now broadly recognized. In consumer businesses with multiple channels to market, the requirements to earn customer loyalty in one channel may be substantially different than those of another. Examples of two such consumer marketing channels are (a) bricks and mortar retail store fronts and (b) home delivery services supported by web-based product selection and ordering. Though each channel appeals to different customer segments, common to both is that customer loyalty is earned when the right mix of product selection, quality and service is reliably delivered at a competitive price. This paper presents two case studies in which the use of analytics led directly to improved profitability through increased loyalty. The first case illustrates how mixed model analytics was applied to longitudinal data to help a retail ice cream store chain redesign its in-store menu boards to improve profitability, while increasing speed of service. The second case illustrates how survival analysis was applied to a home delivery service business to drive customer loyalty through improved new customer acquisition program features. In particular, this paper shows how survival analysis provided insights that led to as much as a 35% improvement in customer retention in a milk home delivery business.

## INTRODUCTION

Many business leaders today rely on their employees and customers sharing some common understanding of the concept of customer loyalty to align business activities with business objectives and performance. As broadly used as the term “customer loyalty” has become, its definition remains strangely elusive. After all, is customer loyalty a concept that relates to measures of customer purchase attributes, number of favorable mentions in online customer reviews, stock price movement, all of these, or something else? Similarly, the term “data analytics” is a pervasive term that is commonly used in discussions throughout organizations of various sizes. It is a term that often gets associated with “customer data mining” or “transaction mining”. Within enterprise, the use of analytics is generally deemed valuable, even strategic, with firms often devoting their most talented employees to vast and complex analytical projects in an effort to gain a competitive advantage. Among consumers, however, a corporation’s use of analytics often raises concerns relating to privacy. Like customer loyalty, business analytics is a term that is difficult to define. Depending on the definitional emphasis, the term “business analytics” conjures thoughts and feelings among individuals ranging from pure fascination to suspicion and even fear.

In order to properly frame the topic of this paper, an attempt to define these terms is necessary. Customer loyalty is defined by Toporek as *the continued and regular patronage of a business in the face of alternative economic activities and competitive attempts to disrupt the relationship. Customer loyalty often results in other secondary benefits to the firm such as brand advocacy, direct referrals, and price insensitivity*. While this definition provides context to the core conceptual ideals underlying common usage of the term, it fails to provide a level of utility that firms require when assessing and measuring the degree of “loyalty” expressed among their consumers. In particular, it does not incorporate more quantitative behavioral ideas strongly espoused by marketers, such as recency of purchases, frequency of purchases and monetary value of purchases (RFM). Still, Toporek’s definition leaves room for the incorporation of more qualitative emotional qualities, which some argue are the true indicators of customer loyalty (e.g., see Barnes, 2006). It is emotion that stirs humans into action. Business leaders are attracted to this notion, as it provides constant guidance to employees regarding the behaviors they should exhibit when ambiguities arise and company policy does not adequately address customer needs. In cases like this, an emotion-based definition of customer loyalty allows for guiding statements like, “follow the Golden Rule with our customers and you will never be wrong,” or “always love our customers and they will always love us back.” In other words, a tension embodies the concept of customer loyalty. On one hand there is a desire to describe it in terms that reflect our notion of loyalty as shared between people, full of emotion and personal commitment. On the other hand, firms often require something more amenable to measurement and analysis. To be sure, loyalty for loyalty’s sake does not ensure business success. Customer loyalty must be an *identifiable* contributor to a firm’s financial viability, lest the concept is at risk of being deemed irrelevant. Customer loyalty programs often attempt to keep a running measurement of loyalty. Loyalty card programs, for example, operate by establishing and implementing rules that precisely reward customers for exhibiting desired behaviors, thus allowing an ever ready assessment of customer loyalty. In this context, customer loyalty is deemed in functional terms that place a strong emphasis on quantitative aspects of loyalty. Some firms, e.g., subscription-oriented businesses, take a pragmatic view of loyalty and describe it in a highly functional manner as a measure of retention (or its converse, attrition). Barnes disagrees with this definition of loyalty, declaring, “Retention is a behavioral concept. A focus on retention creates a high-risk situation where a company may think customers are a lot more loyal than they really are. Satisfaction with functional aspects of product and service is sufficient to drive retention. It takes emotionally driven loyalty to create solid customer relationships.” (Barnes, 2006). Business leaders who believe that their firm’s success requires customers to exhibit some measure of “loyalty” are faced with a choice, therefore, between defining customer loyalty in terms that impassion and guide their employees to perform in ways that engender such loyalty among customers, versus defining customer loyalty in quantifiable behavioral terms that unambiguously identify whether or not customer loyalty is present and contributing the firm’s financial success. This paper takes the more pragmatic approach and assesses customer retention as a measure of customer loyalty.

Turning to the term “data analytics,” searchdatamanagement.com defines it as, *Data analytics (DA) is the science of examining raw data with the purpose of drawing conclusions about that information. Data analytics is used in many industries to allow companies and organizations to make better business decisions and in the sciences to verify or disprove existing models or theories. Data analytics is distinguished from data mining by the scope, purpose and focus of the analysis. Data miners sort through huge data sets using sophisticated software to identify undiscovered patterns and establish hidden relationships. Data analytics focuses on inference, the process of deriving a conclusion based solely on what is already known by the researcher.* While it is possible to expand upon the meaning of data analytics and how it may be employed to help companies succeed, it is clear from this compact definition that data analytics logically follows from the behaviorally oriented definition of customer loyalty. As Davenport, et. al. (2010) describe, analytics is a cultural attribute of firms, as much as it is a set of tools and techniques for drawing inferences.

The purpose of this paper is to demonstrate how customer loyalty concepts might inform data analytics efforts and how data analytics feeds back to decisions that drive customer loyalty as defined in the behavioral sense. This paper, therefore, attempts to describe how customer loyalty and business analytics interrelate. This is done through illustration, with two business case studies analyzed in detail. Both cases come from business decisions faced by the leadership of Oberweis Dairy. Oberweis Dairy has three primary distribution channels, each serving a different customer segment. The first channel involves a network of 46 corporate-owned and franchised Ice Cream and Dairy Stores located throughout the Midwest, which serve made-to-order ice cream treats, e.g., hand scooped ice cream cones, sundaes, milk shakes and malts. Additionally, the stores sell fresh bottled milk and other dairy-related grocery items. The second channel is a home delivery business; whereby, subscribed residential customers receive weekly home delivery of fresh, glass-bottled milk, packaged ice cream, a range of other grocery staples and fresh produce. This service entails a delivery fee. The third channel is a wholesale business; whereby, Oberweis sells its products to regional and national grocery store chains for resale to consumers. The first case study focuses on the dairy store channel and illustrates how data analytics (analysis of longitudinal data via the MIXED procedure) may be used in a quick service restaurant (QSR) or fast casual restaurant setting to aid in the redesign of menu boards, where a goal is to improve profitability. The second case study focuses on the home delivery channel and illustrates how a behaviorally defined measure of customer loyalty, i.e., account retention, can be improved in a subscription based home delivery channel via application of survival analysis utilizing the LIFETEST procedure.

## CASE STUDY #1: DAIRY STORE MENU BOARD REDESIGN

### BACKGROUND

Oberweis Dairy management recognized a need to modify the layout of its restaurant menu display boards. Figure 1 shows an image of the menu board layout being considered for redesign.


**Figure 1. Original menu display board contemplated for redesign.**

A primary objective was to improve customer experience by reducing order times. This was to be accomplished by reorganizing the presentation of menu items on the menu boards and by providing more menu item imagery in order to increase the speed of customer product selection. In considering which imagery would be most beneficial, it was determined that a selection of the most popular sundaes should be pictured. Further, the layout should encourage the path of a typical customer's gaze across the menu to reflect the most efficient way to order an item. Not only did the number of possible product images exceed the space available to reasonably display them, the order in which they should appear on the menu boards was uncertain. The time and expense of redesigning the menu boards needed to be associated with a measurable return. If the wrong images were chosen or the wrong image order was selected, the effort could yield no measurable return. Worse, the effort could actually result in decreased customer transaction values if the menu layout encouraged a shift in consumer purchase behavior toward lower valued products. It was determined that four distinct menu board layout options had merit for deeper consideration. A pilot study indicated that each of the four options yielded comparable improvement in order time. What remained unclear was which option would yield the greatest positive impact on profitability over time.

An experiment was implemented to measure the impact of each design on profitability. Details of the experiment have been described elsewhere (Bedford and Raj, 2012). The measurement of interest was "store-level year-on-year change in profit per transaction". Analysis of the experimental results revealed a clearly preferable design for the new menu boards. Figure 2 shows an image of the layout identified as most likely to yield the greatest level of profit improvement.


**Figure 2. Example of one of four layouts considered for menu display board for redesign.**

## NEW MENU BOARD DESIGN

The winning menu board design (Figure 2.) was implemented across the entire Oberweis Dairy store network beginning mid-March 2012. Upon implementation, a longitudinal study was conducted to measure and verify the anticipated profitability improvement. With the new menu boards being installed in every store, a means for measuring results against a control group over the same time period was needed. It so happens that 19 of the corporate Oberweis stores provide drive-thru window service, whereas the remaining corporate stores do not. As the drive-thru menu panels where not altered during the in-store menu board redesign project, it was determined that they could provide a control against which the in-store transaction profit change could be compared over equal time periods.

### Longitudinal Data Analysis

The time period of study was taken as April 9, 2012 – July 8, 2012. The study concludes with July 8, 2012 measurements, as new drive-thru panels began being installed later that week. The baseline period was taken as March 19, 2012 – April 8, 2012. Data acquired throughout the baseline date range are taken as a single period and identified in graphs that follow by the single week-ending date of April 8, 2012. Data covering the period of study were accumulated weekly. Figure 3a shows year-over-year (YOY) transaction profit change by store, by week and Figure 3b shows first degree Loess curves fit to the data of Figure 3a. In both figures, the red lines correspond to the drive-thru control group, with each line representing only drive-thru transactions for each of the 19 stores. The blue lines correspond to the same set of stores as the red lines, but with only in-store transactions represented. Prompted by the earlier test results, the hypotheses formulated prior to implementation of the new menu board design was that in-store transactions would show an increase in profitability over time compared to drive-thru transactions. To account for product mix differences between typical in-store orders versus drive-thru orders, an annual change score is used for measurement. This also has the effect of controlling for differences in product mix observed seasonally. This leads to a longitudinal study, as comparisons between in-store and drive-thru are sought over time. At the store level, correlations in YOY profit change measures over time are expected. In order to test the one-sided hypothesis that the average weekly YOY transaction profit change is larger for in-store transactions (where customers are exposed to the new menu board designs) than for drive-thru transactions PROC MIXED is utilized. When repeated measures of a response variable are to be modeled for analysis, the covariance structure must first be estimated. Then, the treatment and time effects can be assessed. Littell, et. al. (2006) discusses a four step approach to accomplish this:

1. Model the mean structure, usually by specification of the fixed effects.
2. Specify the covariance structure, between subjects as well as within subjects.
3. Fit the mean model accounting for the covariance structure.
4. Make statistical inference based on the results of the mean model.

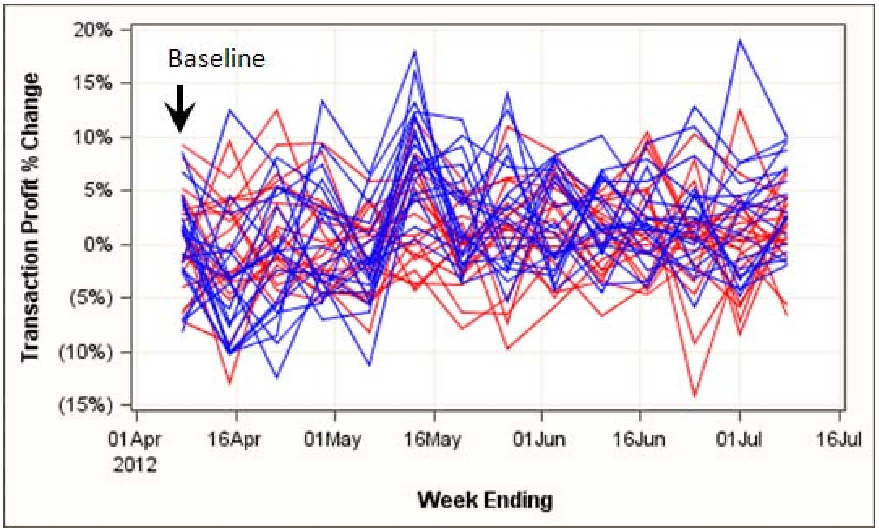


**Figure 3a. Year-over-year (YOY) transaction profit change by store by week. Red lines correspond to control group of drive-thru transactions. Blue lines correspond to test group of in-store transactions at the same set of stores as those corresponding to the red lines. Indicated baseline data cover the preceding 3-week period prior to store exposure to new menu boards.**

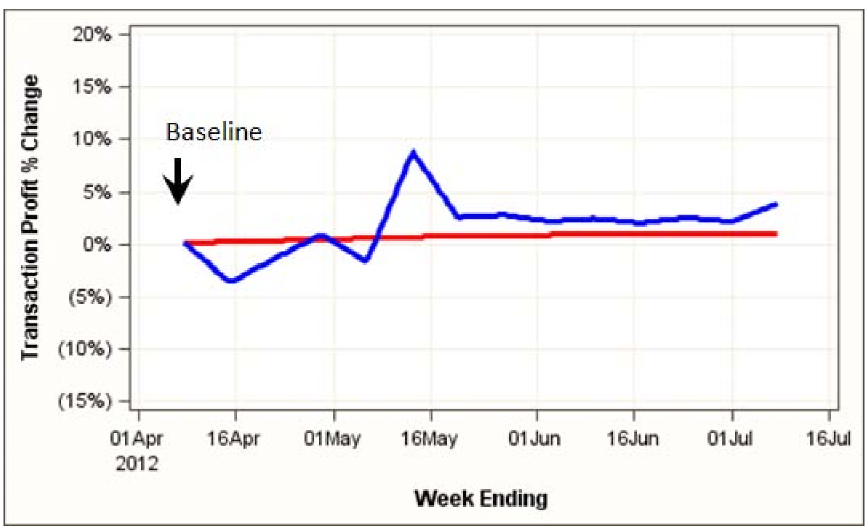


**Figure 3b. First degree Loess curve fits to the raw data shown in Figure3a. As in Figure 3a, the red line corresponds to drive-thru transactions and the blue line corresponds to in-store transactions. The mid-May spike in the blue line occurred during Mother's Day week.**

The Loess curves suggest that a linear or second-order model for the mean structure in time may be reasonable. However, when evaluating fit of covariance models, it is important to understand that the fit depends on the assumed model for the mean response. Therefore, evaluation of covariance model fit should be based on a "maximal" model for the mean. A maximal model is one that minimizes potential misspecification of the model for the mean (Fitzmaurice, et. al., 2004). In this case, a third-order model for time will be used. Therefore, a model of the form:

$$E(\text{Profit\_Change}_{ij}) = \beta_1 + \beta_2\text{Cohort}_i + \beta_3\text{Time}_{ij} + \beta_4\text{Cohort}_i \text{ X Time}_{ij}$$
$$+ \beta_5\text{Time}_{ij} \text{ X Time}_{ij} + \beta_6\text{Cohort}_i \text{ X Time}_{ij} \text{ X Time}_{ij}$$
$$+ \beta_7\text{Time}_{ij} \text{ X Time}_{ij} \text{ X Time}_{ij} + \beta_8\text{Cohort}_i \text{ X Time}_{ij} \text{ X Time}_{ij} \text{ X Time}_{ij} \quad (1)$$

shall be considered first. In equation 1, $E(\text{Profit\_Change}_{ij})$ signifies that the expected value is being modeled and solutions for the coefficients $\beta_i$ are to be calculated. Before fitting the mean model, however, the covariance structure must be specified properly. One may be tempted to simply assume the responses are independent in time, thus naively specifying a diagonal covariance model and then rely on the robustness of a "sandwich" estimator of the covariance matrix to obtain valid standard errors. This is quite easily done by utilizing the EMPIRICAL option in the PROC MIXED statement. This approach is not advisable for data of the type appearing in Figure 3a, however. Though the data correspond to a balanced longitudinal design, the number of subjects across the two cohorts, i.e., in-store and drive-thru, is 38 (19 in each), whereas the number of repeated measures for each cohort is 14. Use of the "sandwich" estimator for balanced longitudinal data is advisable only when the number of subjects is much larger than the number of repeated measures, as the "sandwich" estimator of the covariance matrix is an asymptotic, i.e., large sample, property (Fitzmaurice, et. al., 2004).

PROC MIXED contains a broad library of covariance models. For balanced longitudinal data taken at equal intervals and with no missing values, as is the case with the data shown in Figure 3a, several covariance model options exist and should be evaluated in order to determine the most appropriate to describe the data. For the data of Figure 3a, seven covariance models were investigated before identifying the first-order antedependence model as the best fitting.

To test the fit of this covariance model, PROC MIXED was used with Restricted Maximum Likelihood estimation (REML) employed for the solution method. The following code segment shows how the antedependence covariance pattern model was tested for fit. Though REML is the default solution method with PROC MIXED, it is shown explicitly in the code segment for clarity, as follows:

```
PROC MIXED data=Model_Input method=reml;
      class Cohort Store WeekNum;
      model Profit_Chng=Cohort|Time|Time|Time / ddfm=kr;
      weight ISCYx;
      repeated WeekNum / type=ante(1) subject=Store(Cohort);
RUN;
```

The model statement utilizes equation 1, where Cohort has been specified as a class variable, as has Store and WeekNum. Store provides a unique store id and WeekNum ranges from 0 to 13, where 0 represents the baseline period and 1 is the week-beginning April 9, 2012. WeekNum and time are the same values in the model, but a class variable must be used in the repeated statement to signify the longitudinal nature of the data being modeled, whereas time has been parameterized in the model statement and cannot be a class variable. Denominator degrees of freedom in tests of fixed effects are computed using the Kenward-Rogers method. Response data are weighted by the number of in-store transactions each week of the current year (`weight ISCYx`). This is appropriate to incorporate the reality that some stores consistently see higher guest traffic than others. If this fact is ignored, stores with small relative customer counts will disproportionately influence the analysis.

Table 1 allows fit statistics for the antedependence model to be compared against a variety of other appropriate choices. The antedependence model is a generalization of the autoregressive model.

| | # Parameters | -2 Res Log Likelihood | AIC | AICC | BIC |
|---|---|---|---|---|---|
| Toeplitz | 14 | -1476.8 | -1448.8 | -1448.0 | -1425.9 |
| First-Order Autoregressive | 2 | -1460.6 | -1456.6 | -1456.6 | -1453.3 |
| First-Order Autoregressive + Random Effects | 3 | -1466.5 | -1460.5 | -1460.5 | -1455.6 |
| Heterogenous First-Order Autoregressive | 15 | -1556.5 | -1526.5 | -1525.6 | -1502.0 |
| Heterogenous Toeplitz | 27 | -1577.4 | -1523.4 | -1520.3 | -1479.2 |
| First-Order Antedependence | 27 | -1592.1 | -1538.1 | -1535.0 | -1493.9 |
| Unstructured | 105 | -1709.9 | -1499.9 | -1446.6 | -1327.9 |

**Table 1. Fit statistics for seven covariance model options appropriate for the data of Figure 3a.**

Note that for the results shown in the third row of Table 1, the covariance pattern is set as `type=ar(1)` and a `random Store(Cohort)` statement is added. Of the models attempted, Table 1 illustrates that those suitable for accommodating heterogeneous variances over time yielded the smallest Akaike Information Criteria (AIC). Comparison of AIC for fit with REML estimation is preferred when selecting among non-nested covariance models. As a general (though certainly not universal) rule, use of Bayesian Information Criteria (BIC) is not recommended for covariance model selection, as it entails a high risk of selecting a model that is too parsimonious (Fitzmaurice, et. al., 2004). ANTE(1) is the best fitting model based on AIC and AICC, where AICC adjusts AIC for finite sample size, increasing the relative penalty for model complexity with small data sets. ANTE(1) allows for heterogeneous variances in time. Further, it allows for unequally spaced response measures, which is not a requirement for the data of Figure 3a. Had BIC been used for model selection, a heterogenous autoregressive model would have been selected.

With the covariance structure model identified, the validity of equation 1 can be assessed. In comparing mean response models, however, the appropriate estimation method is maximum likelihood (ML), rather than REML, as shown in the following code segment:

```
PROC MIXED data=Model_Input method=ml;
      class Cohort Store WeekNum;
      model Profit_Chng=Cohort|Time|Time|Time / ddfm=kr;
      weight ISCYx;
      repeated WeekNum / type=ante(1) subject=Store(Cohort);
RUN;
```

Fit statistics for this model are shown in Table 2.

| Fit Statistics | |
|---|---|
| -2 Log Likelihood | -1701.9 |
| AIC (smaller is better) | -1631.9 |
| AICC (smaller is better) | -1626.8 |
| BIC (smaller is better) | -1574.6 |

**Table 2. Fit statistics for third-order time parameterization employing ANTE(1) covariance pattern model.**

The p-value for the term that is third-order time is 0.0388 and the AIC for the overall model is -1631.9. With a second-order model in time, AIC is slightly reduced from -1631.9 to -1632.5, but the p-value for the second-order time term is 0.5818. Thus, the second order model is further reduced to one that involves only a linear parameterization of time, as indicated by the following code segment:

```
PROC MIXED data=Model_Input method=ml;
      class Cohort Store WeekNum;
      model Profit_Chng=Cohort|Time / ddfm=kr;
      weight ISCYx;
      repeated WeekNum / type=ante(1) subject=Store(Cohort);
RUN;
```

Fit statistics for this model are shown in Table 3.

| Fit Statistics | |
|---|---|
| -2 Log Likelihood | -1698.0 |
| AIC (smaller is better) | -1636.0 |
| AICC (smaller is better) | -1632.1 |
| BIC (smaller is better) | -1585.3 |

**Table 3. Fit statistics for linear time parameterization employing ANTE(1) covariance pattern model.**

Table 3 illustrates that the AIC is further improved to -1636.0 with the linear model, indicating that the reduced complexity produces a better model. Moreover, as shown in Table 4, the Time effect has a p-value of 0.0012, indicating strong statistical significance well above the 95% confidence level. The linear model in time with Cohort*Time interaction is used to study the difference between results derived for drive-thru versus in-store transactions. This is accomplished by adding `lsmeans Cohort`. Because the hypothesis is that in-store performance is *increased* relative to drive-thru performance and not simply *different from* drive-thru, a single-tailed test is appropriate and the `diff=controlu` option is employed. This is illustrated in the following code segment:

```
PROC MIXED data=Model_Input method=ml;
       class Cohort Store WeekNum;
       model Profit_Chng=Cohort|Time / ddfm=kr;
       weight ISCYx;
       repeated WeekNum / type=ante(1) subject=Store(Cohort);
       lsmeans Cohort / diff=controlu;
RUN;
```

| Solution for Fixed Effects | | | | | | |
|---|---|---|---|---|---|---|
| **Effect** | **Cohort** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > \|t\|** |
| **Intercept** | | -0.00109 | 0.007181 | 88.7 | -0.15 | 0.8801 |
| **Cohort** | Drive-Thru | 0.003392 | 0.01016 | 88.7 | 0.33 | 0.7391 |
| **Cohort** | In-Store | 0 | . | . | . | . |
| **Time** | | 0.003017 | 0.000911 | 122 | 3.31 | 0.0012 |
| **Time*Cohort** | Drive-Thru | -0.00239 | 0.001288 | 122 | -1.86 | 0.0659 |
| **Time*Cohort** | In-Store | 0 | . | . | . | . |

**Table 4. Fixed effect parameter values and fit statistics for linear time parameterization employing ANTE(1) covariance pattern model for the data of Figure 3a.**

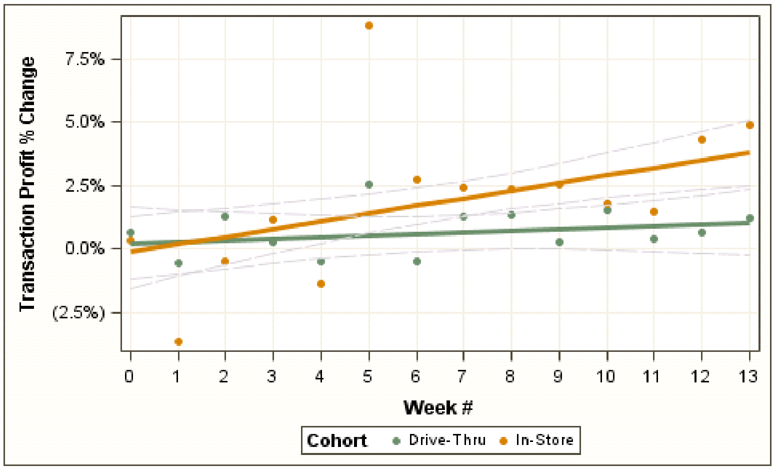


**Figure 4. Plot of model fit results. Lines are predicted values. Dots are weighted-average actual values by week number. Similar to Figures 3a and 3b, week number 0 corresponds to baseline measurements. Dashed lines show 95% confidence intervals for the model fit.**

By incorporating the lsmeans statement in the specification of PROC MIXED, the difference between least squares means for the Cohort effect can be analyzed. Table 5 shows a statistically significant difference between in-store and drive-thru performance, though Figure 4 shows the baseline to be nearly identical for each cohort.

| Differences of Least Squares Means | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Effect** | **Cohort** | **_Cohort** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > t** |
| **Cohort** | In-Store | Drive-Thru | 0.01214 | 0.004938 | 154 | 2.46 | 0.0075 |

**Table 5. Least squares mean difference between in-store and drive-thru performance.**

The estimated mean difference between the cohorts is 0.01214, i.e., 1.2 percentage points, which represents the average improvement in transaction-level profitability between customers across all stores who make purchase decisions based on the new menu-board design, relative to the unchanged drive-thru menu board design.

These results provide strong evidence that customer purchase behavior is influenced by the new menu board designs. Further, that influence results in purchase decisions that improve incremental transaction level profitability by more than a full percentage point for the dairy stores. Given that the previously mentioned pilot work showed that the menu boards also achieve management's initially desired goal of reduced order time, i.e., waiting time for customers, this project is deemed highly successful, with favorable outcomes for customers and the company.

## CASE STUDY #2: PROMOTIONAL OFFER IMPACT ON ACCOUNT RETENTION

### BACKGROUND

Oberweis Home Delivery service provides weekly delivery of farm fresh, glass bottled milk, ice cream and a wide range of other grocery items to residential customers throughout the Midwest. Unlike traditional subscription based services, Oberweis does not require customers to sign a service agreement or any contract obligating subscribers to a minimum number of deliveries or any minimum length of service. Rather, a “standing order” is placed with Oberweis when a customer establishes a new account. The standing order serves as the customer’s continual request for Oberweis to deliver a pre-specified (by the customer) basket of grocery items each week. Customers are placed on a route and assigned a fixed delivery day. Time of delivery for a given customer is approximately the same from week to week, primarily determined by the position of a customer’s address along the route. Delivery times may shift back and forth as new customers are added ahead of an existing customer, or as previous customers discontinue service. A standing order may be modified by any customer in advance, as long as the request is placed prior to the truck loading out (usually by 10:30 PM on the night before delivery). To affect a change a customer may enter a request online, or phone Oberweis Customer Service. New grocery items may be added to or removed from the standing order in this manner. Similarly, items may be added or removed on a one-time basis, with the original standing order remaining intact. Delivery “holds” may be placed in a similar fashion. Holds are used by customers to manage deliveries in relation to vacation plans, for example. With a customer’s account “on hold”, no deliveries are made by Oberweis. All deliveries incur a delivery fee, unless otherwise specified. The standard delivery fee is $2.99 per delivery, regardless of order size. Products are delivered to a “porch box”. Oberweis sells new porch boxes for $24.99 when a new customer account is established. New porch boxes are generally delivered with the customer’s first order, unless requested otherwise by the customer. Customers are responsible for ensuring that their porch box is placed near their doorstep for access by the Oberweis milkman at time of delivery.

Because there are no service agreements or contracts obligating customers to retain service, customers may discontinue their service at anytime by simply contacting Customer Service and requesting that their account be closed. As with any subscription based business, Oberweis assesses account retention as a critical measure of the health of the home delivery business. As discussed in the introduction, retention is viewed as the primary measure of customer loyalty. Oberweis is constantly assessing various elements of the home delivery service to ensure that the influencers of retention are properly managed. Product and service quality are clearly strong influencers of retention. Oberweis maintains a well trained and well staffed Customer Service team to make certain that customer concerns are heard and acted upon quickly and appropriately, ensuring a high level of customer satisfaction in these essential areas.

Still, there is more to retention than simply product and service quality. Oberweis began considering the manner in which new customers were recruited in an effort to better understand influencers of retention. Oberweis attracts new customers to its home delivery business through several different market channels (referred to as “start sources”). These include door-to-door selling (D2D Team), direct mail, online search engine advertising, unsolicited internet based signups (Internet) and telemarketing. To clarify, online new customer sign-ups are either actively sought through paid campaigns or obtained passively as unsolicited sign-ups, e.g., prospects find our website of their own volition and choose to sign up without the enticement of any specific promotional offer. Throughout the remainder of this paper, “Internet” refers to these unsolicited web-based sign-ups and *not* customers derived from a paid online advertising campaign with promotional offers attached. This is an important distinction, as these unsolicited web-based sign-ups (Internet) represent a nearly pure segment of self-selected customers. The remainder of this paper will consider only three start sources in detail: 1. Internet (unsolicited), 2. Direct Mail and 3. D2D Team.

Figure 6 shows retention trends for each of these three start sources. In each case the curves begin with day 0 on the abscissa corresponding to June 1, 2010. With apology, the actual retention values along the ordinate are obfuscated due to the competitive sensitivity of the data. This does not detract from the reader’s ability to make two substantial observations, however. First, the three start sources each produce customer groups with different retention profiles. Second, there is a sharp discontinuity present in the data for both D2D Team starts and Direct Mail starts, whereas no apparent discontinuity exists for Internet starts. Again, Internet starts represent *unsolicited* new customers. The lines plotted in Figure 6 are produced as the survival probability generated by the LIFETEST procedure using the default Kaplan-Meier method, as illustrated in the following code segment:

```
PROC LIFETEST data=LifeTest_Input
       outsurv=LifeTest_Output
       time Duration*Censored(1);
       strata Start_Cat1 / test=wilcoxon adjust=bon;
RUN;
```

The intervals shown are taken as days from account inception. The Start_Cat1 variable identifies the three start sources, i.e., Internet, Direct Mail and D2D Team. PROC LIFETEST allows for convenient testing of differences in survivor functions for each value of the strata variable(s). The Wilcoxon test was selected to analyze the results shown in Figure 6, though no notable difference in test results were observed when the log-rank or other available tests were used.

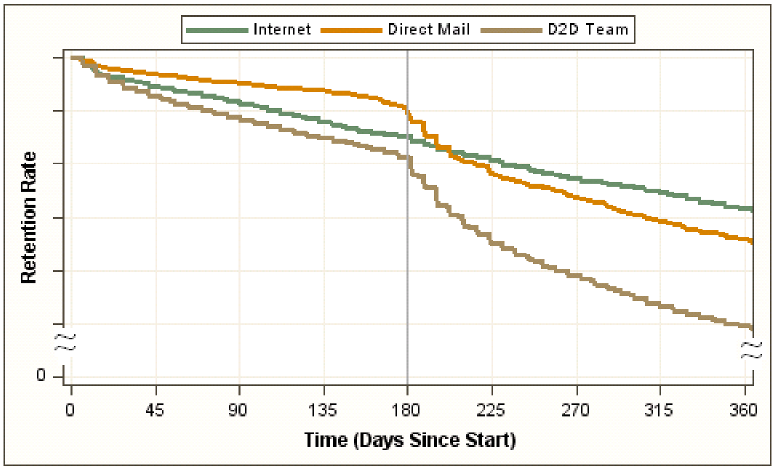


**Figure 6. Retention Rates by Start Source for New home delivery accounts started on or after 6/1/10. Values along the ordinate have been hidden to avoid revealing competitively sensitive data.**

Table 6 shows p-values for the Wilcoxon test comparing each of the three strata values. Because multiple comparisons are tested, a Bonferroni correction was applied to control family-wise error rate. These results provide strong evidence that each start source produces a customer group that exhibits retention behaviors that are statistically significantly different from each of the others above the 95% confidence level.

| Adjustment for Multiple Comparisons for the Wilcoxon Test | | | | |
|---|---|---|---|---|
| Strata Comparison | | | p-Values | |
| Start_Cat1 | Start_Cat1 | Chi-Square | Raw | Bonferroni |
| D2D Team | Direct Mail | 349.6 | <.0001 | <.0001 |
| D2D Team | Internet | 374.4 | <.0001 | <.0001 |
| Direct Mail | Internet | 95.2022 | <.0001 | <.0001 |

**Table 6. Results of multiple Wilcoxon tests across each strata value for the data shown in Figure 6, where the Bonferroni method is used to adjust p-values to control family-wise error rate.**

Based on these results, it is clear that direct mail campaigns beginning on or after June 1, 2010 reliably produced customers with more attractive retention characteristics than those produced by the door-to-door team (D2D) over the entire year following account inception. This is true when comparing direct mail starts against Internet starts through approximately the first six months of the study period. By the 1-year mark, however, the Internet starts exhibited more attractive retention characteristics than either direct mail or D2D. In the case of direct mail and D2D sourced starts, a promotion was offered, whereby new customers received waived delivery fees for the first 6 months of account life, yielding a savings of $77.74 (26 deliveries X $2.99/delivery). Further, for customers remaining after the promotional period, the $24.99 porch box charge was forgiven. However, at the 6 month mark, standard delivery fees of $2.99/delivery began being assessed. Moreover, if a customer closed their account before the conclusion of the promotional period, the porch box charge of $24.99 was also assessed. These features of the promotion put the six-moth milestone in sharp focus for customers. On the other hand, customers signing up unsolicited via the company website received no promotional offers of any kind. For those customers the six-month milestone was not particularly remarkable. As a result, the survivor function of that group shows no discernible discontinuity at or near the 180 day interval mark.

These observations led to the insight that customer retention is influenced by the customer's knowledge that savings offered by a promotion has yet to be fully captured. However, once a promotional period ends, customer loyalty wanes.

## NEW PROMOTIONAL OFFER EXPERIMENT

There is much to be explored regarding the observed behavioral differences for the 6-month promotional offers versus the passive Internet start behavior. For instance, will a different survival function emerge if the promotional period is extended to 12 months, even if the overall promotional value is the same? In particular, it is hypothesized that an extended promotional period will result in smoothing out the survivor function through the six-month milestone, leading to larger survivor function values for the period 6-12 months compared to a promotion whose value drops to zero at the 6-month mark. A direct mail based A/B test was designed to test this hypothesis. The Valpak® Blue Envelope program was chosen as the delivery mechanism. The offers were designed to have nearly identical looks, including the headline offer "Sign Up Today and Save $100". Offer A ran a sub-head that read "That's FREE Delivery for 6 Months", whereas Offer B ran a sub-head that read "That's just 99¢ Per Delivery for 1 Year!" Beyond the promotional code number, the only other differences involved the text describing how the value of the offer is calculated. It is noteworthy, though, that in order to maintain parity in the stated quantitative value of the savings across both offers, the dollar value of the porch box was not claimed in the headline savings value for offer B, whereas the value of $24.99 was included in Offer A. That is, the value of offer A is calculated as the standard delivery fee of $2.99/delivery X 26 deliveries + $24.99 for the porch box; whereas the value of offer B is calculated as simply $2.00 X 52 deliveries, as $2.00 is the amount saved per delivery. Figures 7a-b show the two offers as presented to mail recipients. The two offers were mailed to 1,800,000 randomly selected single-family homes located throughout the Midwest (i.e., 900,000 non-overlapping homes for each offer). The same targeting criteria were applied to recipients of each offer.


**Figure 7a. Promotional advertisement describing an offer for free delivery for 6 months.**


**Figure 7b. Promotional advertisement describing an offer for 99¢ per delivery for 1 year.**

During the active period of the promotion, which began on July 18, 2011, both offers produced quantities of new customers that were statistically indistinguishable at the 99% confidence level. With apology, the number of new customers cannot be published due to the competitive sensitivity of the data.

Due to early indications of success with offer B, the door-to-door team began using the same offer in their selling efforts about 30 days following the direct mail campaign. Additionally, as they had been for over a year, they continued using the equivalent of offer A, as well. In other words, at each salesman's discretion, they were allowed to use offers that mirror those of offers A and B from the Valpak direct mail campaign. This set up a situation that allowed for a simultaneous test of the relevance of the message delivery mechanism (direct mail versus D2D salesmen) in the evaluation of the two offers.

PROC LIFETEST can again be utilized to evaluate differences in strata, where this time both offers and both message delivery mechanisms can be compared against each other and against unsolicited Internet starts. The code segment is nearly identical as before, with the only difference being the input data set and the variable specified in the strata statement:

```
PROC LIFETEST data=LifeTest_Input_99Cents_Sub
       outsurv=LifeTest_Output_99Cents_Sub
       time Duration*Censored(1);
       strata Start_Cat2 / test=wilcoxon adjust=bon;
RUN;
```

Survivor function curves, labeled “Retention Rate” on the ordinate, for each strata value are plotted in Figure 8. In each case the curves begin with day 0 on the abscissa corresponding to July 18, 2011. Again, with apology, the actual values along the ordinate are not shown due to the competitive sensitivity of the data. Still, the discontinuity at 180 days is again observed in the curves corresponding to the “Free Delivery for 6 Months” offers, as anticipated. Further, any notable discontinuity at 180 days is visually absent in the curves corresponding to the “99¢ Per Delivery for 1 Year” offers.

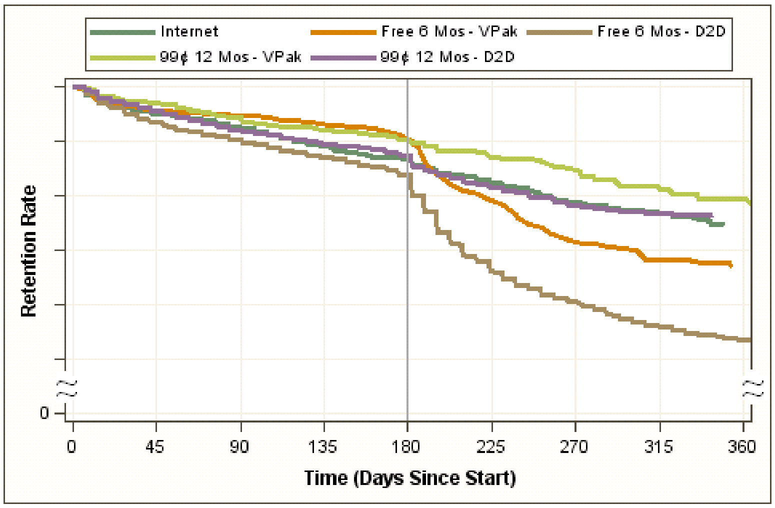


**Figure 8. Retention Rates by Start Source for New home delivery accounts started on or after 7/18/11. Values along the ordinate have been hidden to avoid revealing competitively sensitive data.**

Table 7 shows p-values for the Wilcoxon test comparing each of the five strata values. Again, the Bonferroni correction was applied to control family-wise error rate. These results provide strong evidence that each start source produces a customer group exhibiting retention behaviors that are statistically significantly different from each other above the 95% confidence level, with two critical exceptions. The starts sourced through the door-to-door sales team utilizing the “99¢ Per Delivery for 1 Year” offer compared to unsolicited Internet starts (i.e., “99¢ 12 Mos – D2D” vs. “Internet”) exhibits no statistically significant difference. Similarly, starts sourced through the Valpak direct mail piece utilizing the “99¢ Per Delivery for 1 Year” offer compared to unsolicited Internet starts (i.e., “99¢ 12 Mos – Vpak” vs. “Internet”) also exhibits no statistically significant difference. This is the result that was hypothesized.

| Adjustment for Multiple Comparisons for the Wilcoxon Test | | | | |
|---|---|---|---|---|
| Strata Comparison | | | p-Values | |
| Start_Cat2 | Start_Cat2 | Chi-Square | Raw | Bonferroni |
| 99¢ 12 Mos - D2D | 99¢ 12 Mos - VPak | 9.7876 | 0.0018 | 0.0176 |
| 99¢ 12 Mos - D2D | Free 6 Mos - D2D | 108.5 | <.0001 | <.0001 |
| 99¢ 12 Mos - D2D | Free 6 Mos - VPak | 30.0291 | <.0001 | <.0001 |
| 99¢ 12 Mos - D2D | Internet | 2.1300 | 0.1444 | 1.0000 |
| 99¢ 12 Mos - VPak | Free 6 Mos - D2D | 155.9 | <.0001 | <.0001 |
| 99¢ 12 Mos - VPak | Free 6 Mos - VPak | 15.4966 | <.0001 | 0.0008 |
| 99¢ 12 Mos - VPak | Internet | 2.0940 | 0.1479 | 1.0000 |
| Free 6 Mos - D2D | Free 6 Mos - VPak | 142.7 | <.0001 | <.0001 |
| Free 6 Mos - D2D | Internet | 111.4 | <.0001 | <.0001 |
| Free 6 Mos - VPak | Internet | 15.9907 | <.0001 | 0.0006 |

**Table 7. Results of multiple Wilcoxon tests across each strata value for the data shown in Figure 8, where the Bonferroni method is used to adjust p-values to control family-wise error rate.**

Though ordinate values are hidden in Figure 8, the analysis presented in Table 7 provides strong statistical evidence that the duration of the promotional period has a direct influence on customer loyalty, as measured by account retention behavior. The measured improvements in retention for the longer term offer at approximately 1 year is 15% and 36% for direct mail and door-to-door sales team, respectively. For just the six-month period to the *right* of the vertical line at 180 days in Figure 8, retention improvement has already added $641,000 in incremental revenue.

## CONCLUSION

This paper presented results from two case studies, each drawn from careful analysis of real customer purchase behaviors in different channels of business at Oberweis Dairy. The first case study considered how modifications to menu boards influence customer menu item selections and the impact of that influence on transaction profitability. The second case study considered how the length of a promotional period might influence retention behavior in a home delivery business. In both cases, measures of customer behavior were viewed as indicators of customer loyalty, though it is recognized that customer loyalty has emotional elements that are not captured in this type of analytical work. Still, as a practical consideration in running a business, leaders often rely on measurable surrogates of customer loyalty to make sound business decisions. This paper illustrated how data analytics and customer loyalty concepts intersect in two notably different lines of business in a real company.

## REFERENCES

### Books

Allison, P.D. (2010). *Survival Analysis Using SAS®: A Practical Guide, Second Edition*. Cary, NC: SAS Institute Inc.

Barnes, J.G., (2006). *Build Your Customer Strategy: A Guide to Creating Profitable Customer Relationships*. Hoboken, NJ: John Wiley & Sons, Inc.

Davenport, T.H., Harris, J.G. and Morison R., (2010). *Analytics at Work: Smarter Decisions, Better Results*. Boston, MA: Harvard Business School Publishing.

Fitzmaurice, G.M., Laird, N. M. and Ware, J. H. (2004). *Applied Longitudinal Analysis*. Hoboken, NJ: John Wiley & Sons, Inc.

Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D. and Schabenberger, O. (2006). *SAS® for Mixed Models, Second Edition*. Cary, NC: SAS Institute Inc.

### Online

Bedford, B.D. and Raj, W. (2012). Webinar: *Customer Analytics Drive Customer Loyalty, not Loyalty Programs.* Webinar located at: http://loyalty360.org/conferences/event/customer-analytics-drive-customer-loyalty-not-loyalty-programs.

SearchDataManagement.com. “The Definition of Data Analytics.” provided online at: http://searchdatamanagement.techtarget.com/definition/data-analytics

Toporek, Adam. “The Definition of Customer Loyalty.” *Customers That Stick*, August 27, 2011. http://customersthatstick.com/blog/customer-loyalty/the-definition-of-customer-loyalty/

## ACKNOWLEDGMENTS

I wish to acknowledge Professor Paul D. Allison for graciously agreeing to review the section concerning survival analysis using PROC LIFETEST in this paper prior to publication. His many books and journal publications have formed the educational foundation of countless scientists and statistical practitioners worldwide, both within academia and industry. I am indebted to him for what I have learned through his books and for his patience in reviewing this paper.

Additionally, I must acknowledge the contributions of Oberweis colleagues Dan Rosier and Marijke Shaffer for their project leadership and creative talents in association with the work highlighted herein. Further, I am indebted to Oberweis colleagues Eric McMillan, Elizabeth Craig and Dave Hassler for their exceptional leadership and strong customer-centric focus regarding the project highlighted in Case Study #1. The project would not have been possible without the dedication and talents of each of these individuals.

Also, I would like to acknowledge Ranald MacDonell and Barbara Hacker of Valpak for their outstanding support and patience in working with me to execute the campaign described in Case Study #2, as well as many other campaigns that preceded and followed that one. They are knowledgeable and easy to work with.

Finally, I would like to thank my CEO Joe Oberweis and his father and Chairman Jim Oberweis for their continued support and willingness to test new ideas in true entrepreneurial fashion. Their embrace of analytics as a path toward business growth is essential.

## CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Bruce Bedford, Ph.D.
Oberweis Dairy
951 Ice Cream Drive
North Aurora, Illinois 60542
630-801-6103
Bruce.Bedford@oberweis.com
www.oberweis.com