MSIA 421 DATA MINING

Project: Video Strategies for Suddenlink

GROUP 5
Alejandro Avalos Mar
Shiyi Chen
Xiang Li
Steven Lin
Ameer Khan
Igor Uzilevskiy
Sanjeevni Wanchoo

I. Executive Summary

The objective of this analysis was to propose actionable Video strategies that would increase Suddenlink's revenue. The group investigated what attributes of the Video bundle make customers 'sticky', that is, the objective was to determine the attributes associated with higher retention rates for Video subscribers. In order to answer this question, we looked at how the retention rate depended on the Video bundle characteristics, such as whether the customer opted for Premium, TiVo, HDTV, and DVR services, using survival analysis models.

The group analyzed the attributes of Video bundles for customers over a period of two years, from January 2013 to January 2015. In order to accurately estimate retention rates, only customers that subscribed to the Video bundle after January 2013 were included. This approach controls for the effects of left censoring in the data and eliminates the effect of long-term customers on the survival models. A customer is considered active when they subscribe to the Video bundle, and a cancel occurs when a customer drops the Video bundle or switches service providers altogether. The same dataset was used for both survival analysis and customer segmentation.

Although Suddenlink's margin on Premium is low (less than 25%), the analysis showed that having Premium makes customers more sticky. Premium channels are associated with higher retention rates than any other add-on feature of the Video bundle, such as TiVo, HDTV, DVR, etc. We estimated that the overall retention rate of the customers increases by 3.4 percentage points, from 88.1% to 91.5%, when they switch to Premium. Controlling for other factors, the increase in the expected customer lifetime value ranges from $92.26 for low-end Video customers to $157.75 for the higher-end segments when consumers opt for Premium. Scaling these to the number of customers of each type, the expected change in CLV by converting low-end customers to Premium is $365,000 based on a 10% conversion rate. The corresponding figure for the high-end customers is $1,280,000. The number of months a customer subscribes to the service also increases by 2.65 for low-end customers and 2.31 for high-end ones.

To determine the optimal Video strategy, the customers were segmented based on HSD, video and phone prices. For three out of the six segments, the effect of Premium on retention rates is high, indicating that the odds of retaining the customers increase substantially if these customers have Premium. Therefore, it is recommended to target these segments for up-selling Premium.

II. Introduction

The team set out to answer the following questions relating to Suddenlink's Video line of business: What attributes of the Video bundle make customers sticky? What is the best strategy for the Premium service? What is the best overall Video strategy? We began by conducting some exploratory analysis. Of the three services offered by Suddenlink, video prices are the highest, although the margin on Video is not as high as High Speed Data (HSD). Video prices also have the greatest variation. The standard deviation for Video price is about 4 times as large as for the HSD price and Phone price.

The dataset for survival analysis consisted of each row representing a time period-customer combination. In order to accurately estimate retention rates, we included only customers who subscribed to the Video bundle within the timespan of the data snapshot provided. The resulting dataset has over 3.8 million observations for 425,000 unique customers. 423,434 of the observations represent instances of canceling the Video bundle. Customers that were acquired during the period of data will not have any recurring payment information (in table `srv_bundle_mth`) for the first month, January 2013, and customers that cancel the service before the end of the period will not have recurring payment information for the last month recorded, January 2015. If a customer does not have a recurring payment information for both the first and last months in the data, then they joined and canceled during the period of the data. We defined cancellation as dropping Video bundle. We used this same dataset for both survival analysis and cluster analysis.

Breakdown of the customers acquired after January 2013 based on various add-on services to the Video bundle are shown below:

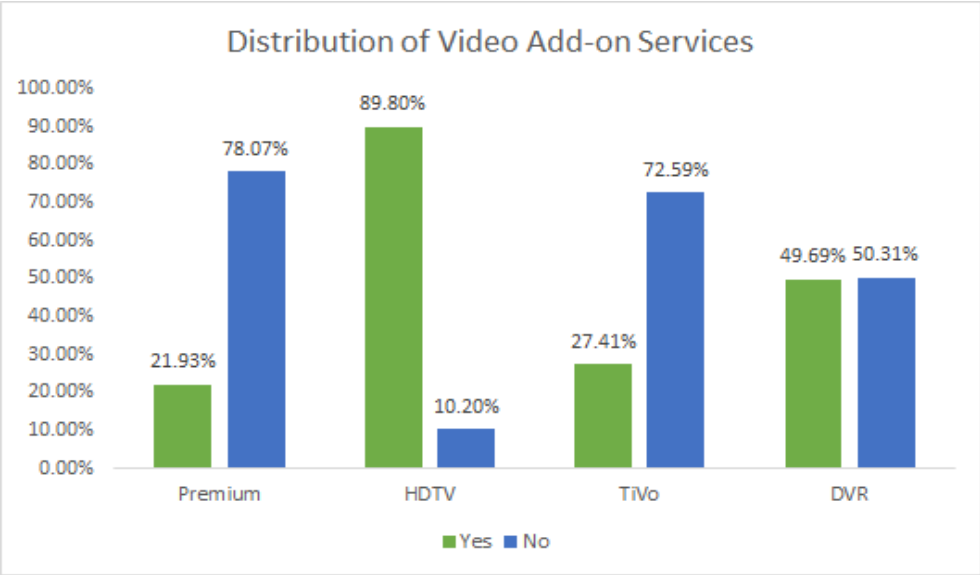

Figure 1: *Distribution of Video Add-On Services among Customers Acquired after January 2013*

III. Analysis

a. Survival Analysis

The team performed survival analysis by fitting a logistic regression model estimating retention rates as a function of various Video bundle characteristics for each customer. The probability of a customer dropping the Video bundle in a given period was modeled against the following attributes:

- `video_t`: Number of months since a customer has subscribed to the Video bundle, ranging from 1 to 23.
- `premium_flag`: 1 if the customer had Premium service in the previous month and 0 otherwise.
- `hdtv_flag`: 1 if the customer had HDTV service in the previous month and 0 otherwise.
- `tivo_flag`: 1 if the customer had TiVo service in the previous month and 0 otherwise
- `dvr_flag`: 1 if the customer had a DVR service in the previous month and 0 otherwise

The model represented mathematically is:

$$\log\left(\frac{\pi_{it}}{1-\pi_{it}}\right) = \beta_0 + \beta_1 video_{t_{i1}} + \cdots + \beta_{22} video\_t_{i22} + \beta_{premium} premium\_flag + \beta_{htdv} hdtv\_flag + \beta_{tivo} tivo\_flag + \beta_{dvr} dvr\_flag$$

where $\pi_{it}$is the probability of cancel for customer$i$ in time period t, and the β's are the coefficients estimated by the model.

The results of the model are shown in the appendix, Figure 1. All variables are statistically significant with small p-values. Negative values of coefficients for a particular service imply a lower probability for a customer to cancel the Video bundle. The coefficients of the first 22 time-periods (`video_t`) are relative to period 23 as the base. The retention rate of a customer for a specific time period, for instance, `video_t` $= 3$ can be calculated as shown below. From the results shown in appendix, we can see that the coefficient estimate for `video_t` $= 3\ is\ -2.3731$, and the intercept is 0.0563. The retention rate $r_{non-premium}$ for a non-Premium customer for their third period is

$$r_{non-premium} = 1 - \frac{1}{1+\exp\left(-(intercept + video\_t_3)\right)}$$

$$= 1 - \frac{1}{1+\exp\left(-(0.0563 - 2.3731)\right)} = 91.03\%$$

Similarly, the retention rate $r_{premium}$ can be calculated for customers with Premium for the same period. From the appendix, the estimate for `premium_flag` is -0.3443. Here, we also take this estimate into consideration, as shown below:

$$r_{premium} = 1 - \frac{1}{1 + \exp\left(-\left(intercept + video_{t_s} + premium\_flag\right)\right)}$$

$$= 1 - \frac{1}{1 + \exp\left(-(0.0563 \ - 2.3731 - 0.3443)\right)} = 92.25\%$$

Using the above retention rates from the model, expected value of the customer lifetime value (CLV) of a customer for Video without any additional services is estimated to be $419.58. For a customer with Premium, and controlling for other add-on services, the expected value of the CLV changes to $511.84, going up by $92.26. The average number of time-periods until a customer cancels the service also goes up by 2.65, going up from 9.77 to 12.42 months. The following figure shows the Video retention rates over time for Premium and non-Premium customers who have no additional services.

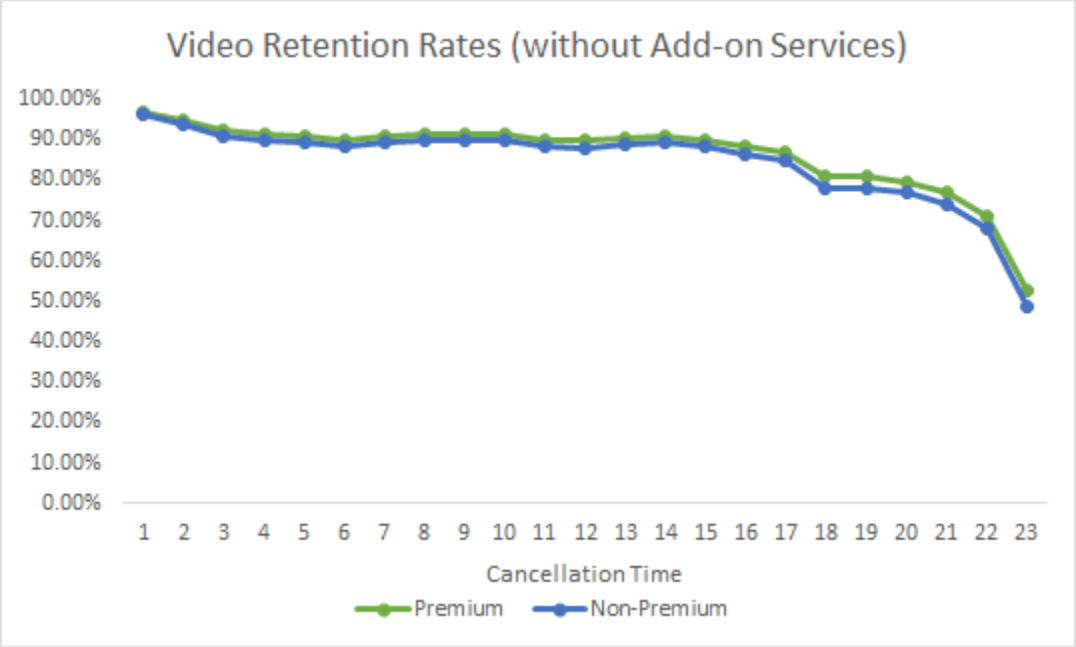

Figure 2: *Video Retention Rates for Non-Premium and Premium Customers who have no additional Add-on Services for Video Bundle*

Similarly, the expected CLV of a customer for Video with additional services such as HDTV, TiVo and DVR, but not Premium, is $721.26. With the addition of Premium, the mean CLV changes to $879.01, going up by $157.75. The average time to cancel also increases by 2.32 periods, from 8.58 months to 10.90 months. The figure below shows

the Video retention rates for customers who also subscribe to HDTV, TiVo and DVR add-on services.

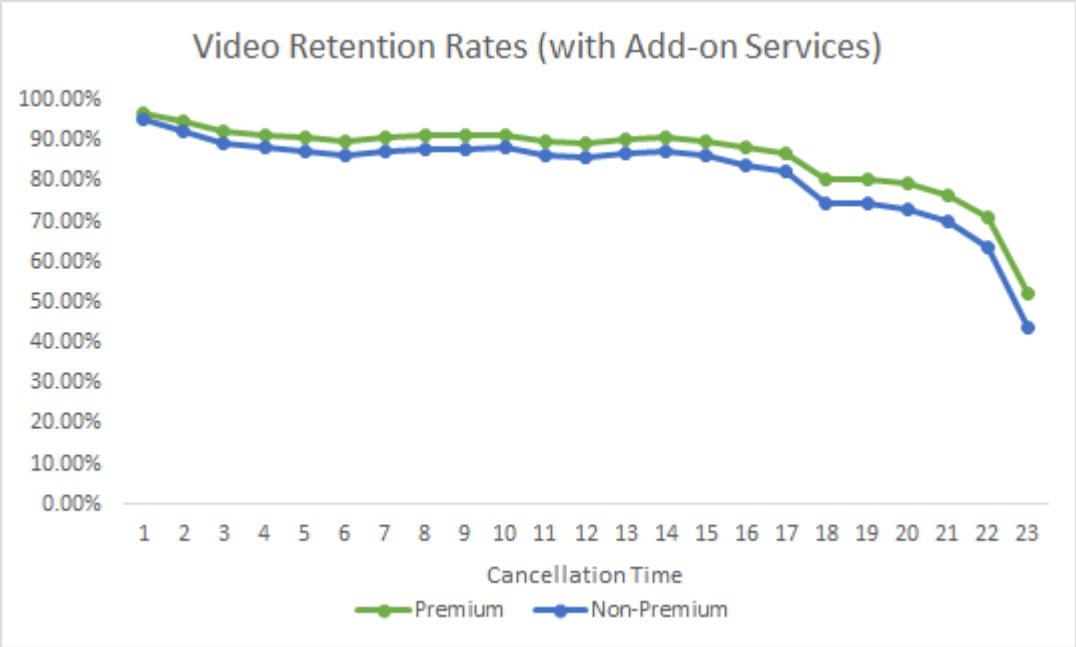

Figure 3: *Video Retention Rates for Non-Premium and Premium Customers with HDTV, TiVo and DVR add-on Services*

All factors being equal, the probability of a customer canceling increases the longer they stay with Suddenlink. Having Premium service is associated with a decrease in the odds of canceling the Video bundle by 29%, whereas customers without the HDTV service have 26% lower odds of canceling the Video bundle.

The model suggests that retention rates are not constant across time periods. More specifically, there is a clear trend of higher retention rates in earlier time periods. The figures above also reveal a sudden drop in retention rates after a customer's 18th month of service indicating external factors as possible causes. This sharp decrease in retention rates coincides with Suddenlink dropping the Viacom network from its Premium service offering. The average price for the Video bundle with Premium also increased by $20 around the same time. The plummeting retention rates beyond time-period 18 can be explained by these changes to the Suddenlink Video bundle.

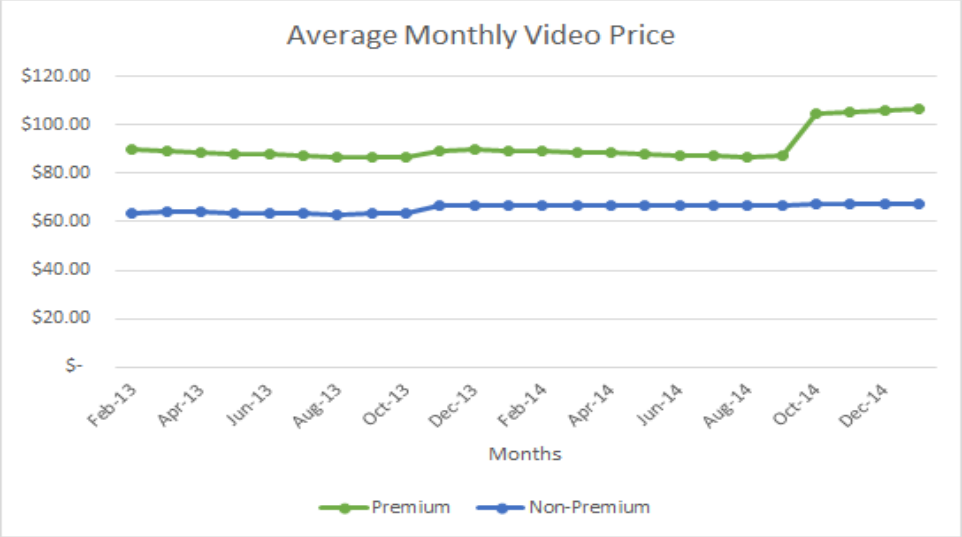

Figure 4: *Average Monthly Video Price for Customers with and without Premium for Calendar Months*

b. Customer Segmentation

Clustering analysis was used to segment the customers with the goal of designing effective and targeted Video strategies. k-Means clustering was used to group customers based on their monthly Video, Data and Phone prices. These variables were chosen because they provided insights on the customer behavior and lead to segments with high/low values on these variables, which facilitated different action plans. Attributes for Data and Phone were included in the Video strategy analysis since they provide a more holistic perspective of a customer. The final number of clusters were determined so that the most actionable clusters could be identified.

Table 1: *Description of Suddenlink Customer Segments*

Segment	Name	Video Description
1	Cord Cutters - Elite	High Data usage, no video bundle
2	Video & Data - Medium	Moderate DVR and low TiVo and Premium
3	Triple Players - Elite	High Premium, DVR and TiVo, moderate Data
4	Triple Players - Medium	High TiVo and DVR, and moderate Premium
5	Video & Data - Economy	Low on everything
6	Triple Play - Economy	Low TiVo and Premium, medium DVR, low Data
7	Cord Cutters - Economy	Low Video and Data

Figure 5 shows the 7-cluster solution, and Figure 6 profiles the segments based on Video add-on services available to the customers. The description of the segments based on key Video and Data attributes is in Table 1.

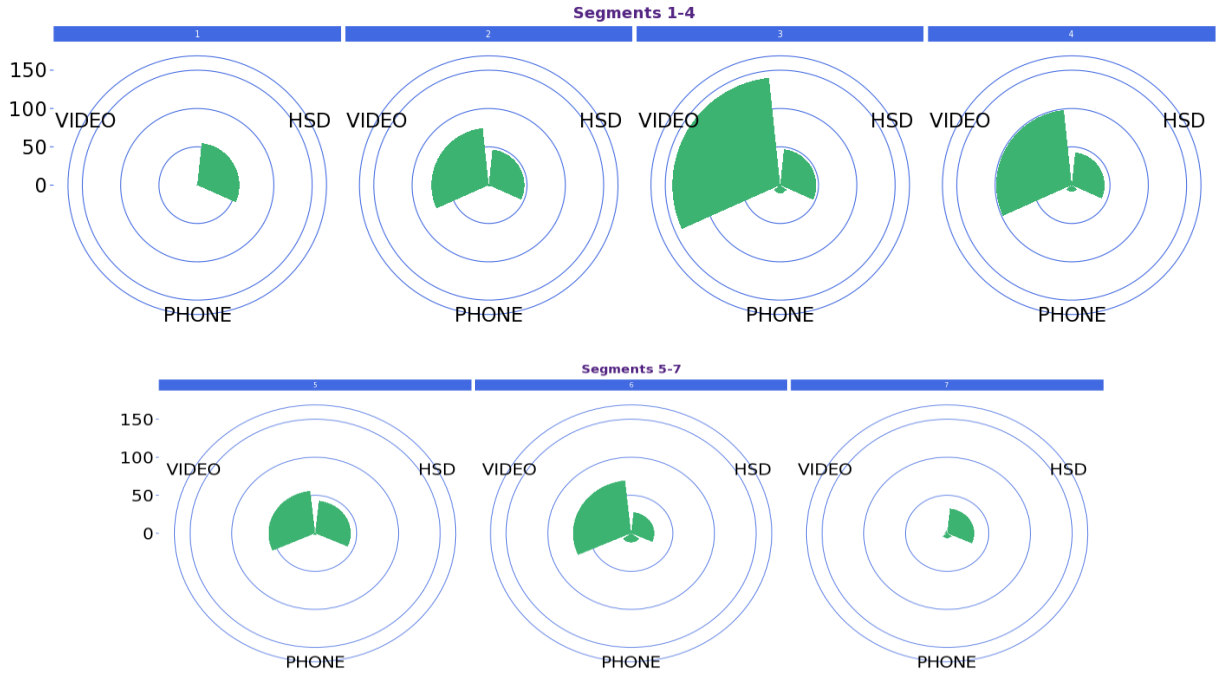

Figure 5: *Visualization of the Customer Segments based on the Clustering Attributes – Video, HSD and Phone Prices*

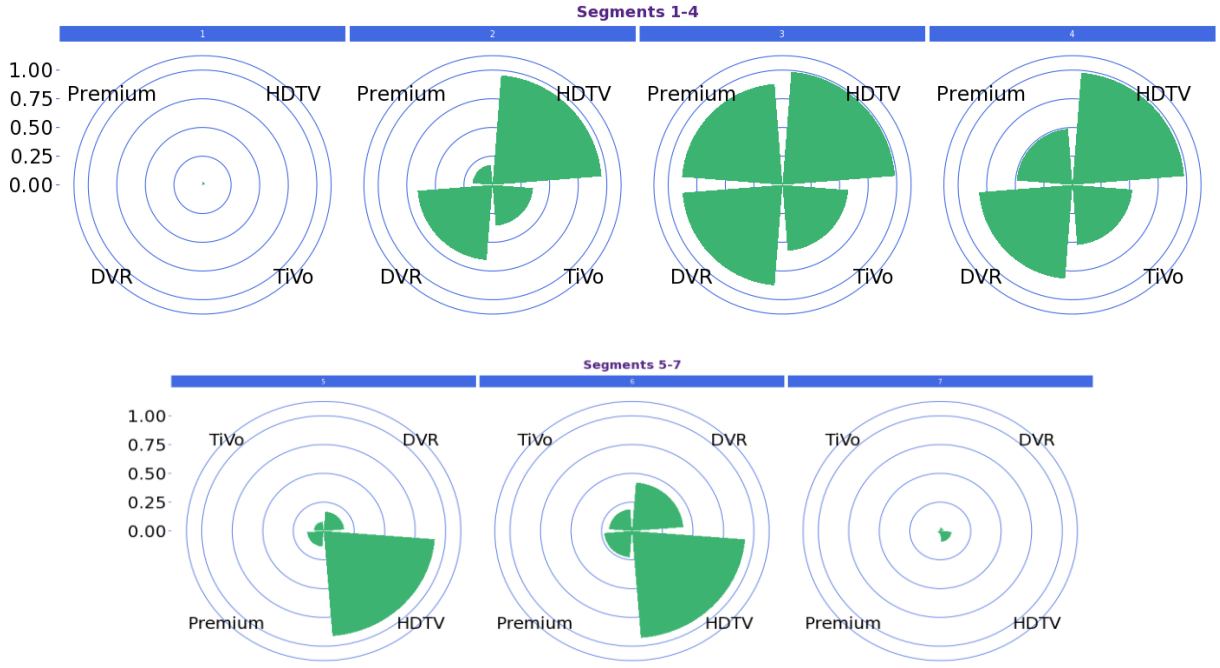

Figure 6: *Profiles of Segments based on their Video Characteristics*

The segments with relatively high Premium subscription (Table 2) are the Triple Players - Elite and Triple Players - Medium (segments 3 and 4). A strategy formed around up-selling the Premium service is feasible for segments Video & Data - Medium, Video & Data - Economy and Triple Play - Economy (segments 2, 5 and 6). Segment 3 is already high on Premium subscription, and the Cord Cutters - Elite and Cord Cutters - Economy (segments 1 and 7), as their name suggests, have no interest in the Video bundle.

Table 2: *Effect of Premium Service on the Retention Rates of Customer Segments*

Customer Segment	Number of Customers	% Customers with Premium	Base Retention Rate (No Premium)	Retention Rate with Premium	Change in Retention Rate
1	167394	0.41%	84.17%	83.87%	-0.30%
2	**107151**	**17.52%**	**89.12%**	**92.20%**	**3.08%**
3	15386	88.17%	91.04%	90.48%	-0.55%
4	63128	48.93%	90.20%	91.66%	1.46%
5	**93691**	**13.71%**	**87.20%**	**92.55%**	**5.35%**
6	**71367**	**22.90%**	**88.84%**	**94.48%**	**5.65%**
7	94108	2.53%	87.68%	85.87%	-1.81%

The estimates of change in retention rates for each segment, when customers add Premium service to their Video bundle are provided in Table 2. A large change in the retention rates is noticed for segments 2, 5 and 6, reaffirming the strategy to up-sell Premium to these segments.

Table 3 gives estimates of the expected CLV increase in revenue by providing Premium service to non-Premium users considering 10% and 30% conversion factors, and without taking into account the costs for Suddenlink to incentivize non-Premium customers to convert. The total CLV increase would solely come from the increase in the expected time periods of the customer. One mechanism to incentivize non-Premium customers to switch to Premium is by offering discounted Premium service. The discount rate would affect the conversion factor, as higher discounts would increase the conversion factor to Premium. This report does not take into account for the cost of offering discounted Premium services to customers, so the total increase in CLV is in terms of revenue. Further analysis would be required to determine the optimal discount and conversion rates to achieve maximum profit in terms of CLV (by adjusting for costs). Refer to Table

A3 for the estimates of increase in expected CLV per customer and non-Premium population for each segment.

Table 3: *Increase in Total CLV by Providing Free Premium Service*

Segment	Conversion Factor	Total CLV Increase (Revenue Only)	Conversion Factor	Total CLV Increase (Revenue Only)
2	10%	$1,967,529	30%	$5,902,587
5	10%	$2,099,324	30%	$6,297,972
6	10%	$2,758,701	30%	$8,276,102

c. Customer Lifetime Value Calculations

Customer lifetime values were computed to estimate the change in the Net Present Value of a Suddenlink's customer of switching a non-Premium customer to Premium service. Because the margins on Premium are small, the focus was only on the change in CLV caused by expected increase in retention rates, and not the increased cash flow per customer. For specific combinations of HDTV, TiVo and DVR services, the period cash flows were estimated by taking the mean monthly Video payments of non-Premium customers.

The retention rates of customers for their first 23 time periods were obtained from the survival analysis performed earlier for a customer with and without Premium service. For periods beyond 23, the retention rates were extrapolated using a simple retention model, which is time period-independent. The CLV is the sum of the discounted cash flows over 500 time periods, weighed by the probability of a customer making that payment (given by period-specific retention rate). The difference between the expected CLV for a Premium and non-Premium customer, controlling for other factors, provides the increase in customer lifetime value due to converting a customer to Premium. The calculations for a range of conversion rates is shown in Table 4. Lower-end customers are those with no TiVo, HDTV or DVR services, while higher-end ones have all three add-on services.

Table 4: *Range of Incremental CLV for Different Premium Conversion Factors and Type of Customers*

Type of Customer	No. of New Customers	% of New Customers	Conversion Factor	Total CLV Increase	Conversion Factor	Total CLV Increase
Lower-End	39,500	9.30%	10%	$364,984	30%	$1,094,951
Higher-End	81,000	19.08%	10%	$1,280,362	30%	$3,841,086

IV. Recommendations

One of the most effective ways for Suddenlink to increase the expected customer lifetime value is to increase the retention rate of their customers by incentivizing them to acquire premium services. This could be done by offering free or discounted Premium service as a promotion for few months. Suddenlink should target the segments 2, 5 and 6 with Premium as they subscribe to basic Video bundles. The effect on cancellation rates is significantly more pronounced for Premium service compared to TiVo, DVR and HDTV for these customer segments.

Another strategy to maintain higher expected CLV is to retain customers for longer time periods. A small change in the retention rates for high retention segments can lead to a large change in the expected CLV. This is illustrated in the figure below. For example, a 1% improvement in the current retention rate increases the expected CLV by almost 18%. Segment 6, which has 94.48% retention for Premium customers, must be targeted to implement strategies to retain Premium customers.

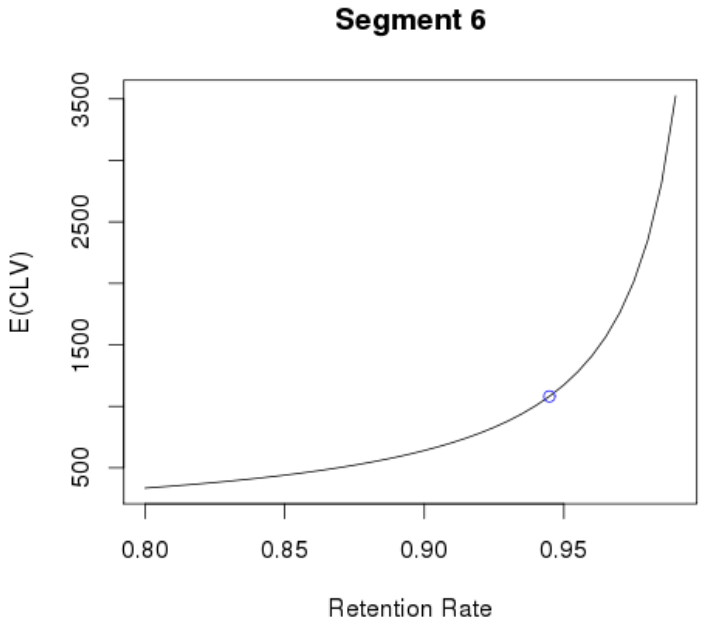

Figure 7: *Change in Expected Value of CLV for Segment 6 Premium Customers against Retention Rate*

A separate analysis can be performed to determine factors that contribute to churn and develop a strategy to prevent churn and a potential large drop in the expected CLV.

Other possible strategies to increase CLV that Suddenlink can pursue include cross-selling other services associated with higher retention rates (for instance, TiVo and DVR), which translates to an increase in monthly revenue as well as expected CLV, and providing incentives for the customers to sign up for longer contracts.

Appendix

Discrete-Time Survival Analysis Results:

$$\log\left(\frac{\pi_{it}}{1-\pi_{it}}\right) = \beta_0 + \beta_1 video_{t_{i1}} + \cdots + \beta_{22} video\_t_{i22} + \beta_{premium} premium\_flag + \beta_{htdv} hdtv\_flag + \beta_{tivo} tivo\_flag + \beta_{dvr} dvr\_flag$$

Table A1: Discrete-Time Survival Analysis Estimates

Analysis of Maximum Likelihood Estimates					
Parameter		**DF**	**Estimate**	**Std Error**	**Pr > ChiSq**
Intercept		1	0.0563	0.0203	0.0056
video_t	**1**	1	-3.2791	0.0211	<.0001
video_t	**2**	1	-2.7647	0.0207	<.0001
video_t	**3**	1	-2.3731	0.0205	<.0001
video_t	**4**	1	-2.2579	0.0205	<.0001
video_t	**5**	1	-2.1876	0.0205	<.0001
video_t	**6**	1	-2.0803	0.0205	<.0001
video_t	**7**	1	-2.1988	0.0208	<.0001
video_t	**8**	1	-2.2444	0.0209	<.0001
video_t	**9**	1	-2.2471	0.0211	<.0001
video_t	**10**	1	-2.2569	0.0212	<.0001
video_t	**11**	1	-2.0880	0.0212	<.0001
video_t	**12**	1	-2.0524	0.0214	<.0001
video_t	**13**	1	-2.1366	0.0218	<.0001
video_t	**14**	1	-2.1845	0.0221	<.0001
video_t	**15**	1	-2.0961	0.0222	<.0001
video_t	**16**	1	-1.9189	0.0222	<.0001
video_t	**17**	1	-1.7884	0.0224	<.0001
video_t	**18**	1	-1.3298	0.0222	<.0001
video_t	**19**	1	-1.3266	0.0229	<.0001
video_t	**20**	1	-1.2519	0.0237	<.0001
video_t	**21**	1	-1.0938	0.0246	<.0001
video_t	**22**	1	-0.8019	0.0257	<.0001
premium_flag		1	-0.3443	0.00441	<.0001
hdtv_flag		1	0.2318	0.00597	<.0001
tivo_flag		1	0.0561	0.00476	<.0001
dvr_flag		1	-0.0876	0.00439	<.0001

Table A2: *Odds Ratio Estimates*

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
video_t 1 vs 23	0.038	0.036	0.039
video_t 2 vs 23	0.063	0.060	0.066
video_t 3 vs 23	0.093	0.090	0.097
video_t 4 vs 23	0.105	0.100	0.109
video_t 5 vs 23	0.112	0.108	0.117
video_t 6 vs 23	0.125	0.120	0.130
video_t 7 vs 23	0.111	0.107	0.116
video_t 8 vs 23	0.106	0.102	0.110
video_t 9 vs 23	0.106	0.101	0.110
video_t 10 vs 23	0.105	0.100	0.109
video_t 11 vs 23	0.124	0.119	0.129
video_t 12 vs 23	0.128	0.123	0.134
video_t 13 vs 23	0.118	0.113	0.123
video_t 14 vs 23	0.113	0.108	0.118
video_t 15 vs 23	0.123	0.118	0.128
video_t 16 vs 23	0.147	0.141	0.153
video_t 17 vs 23	0.167	0.160	0.175
video_t 18 vs 23	0.265	0.253	0.276
video_t 19 vs 23	0.265	0.254	0.278
video_t 20 vs 23	0.286	0.273	0.300
video_t 21 vs 23	0.335	0.319	0.351
video_t 22 vs 23	0.448	0.426	0.472
premium_flag	0.709	0.703	0.715
hdtv_flag	1.261	1.246	1.276
tivo_flag	1.058	1.048	1.068
dvr_flag	0.916	0.908	0.924

CLV for Customer Segments:

Table A3: *Expected CLV per Customer for Selected Segments*

Segment	CLV Increase per Customer	Total Customers	Percentage of Premium	Non-Premium Customers
2	$222.62	107151	17.52%	88381
5	$259.68	93691	13.71%	80844
6	$501.39	71367	22.90%	55021