Timing Model Predictions for DISH Network Quarterly Customer Acquisitions: Absence Makes the Desire to Subscribe Grow Fonder

1. Executive Summary

In this project, 1996-2016 quarterly data for DISH Network customer acquisitions was provided for analysis. The objective was to build a compelling model for fitting customer acquisitions. Additional covariates were obtained and built, including seasonality and unemployment, in order to improve the analysis.

Many models were built and evaluated, including those that followed Weibull and Weibull-gamma regressions. Latent class models were also built, and upon these models, heterogeneity was introduced on combinations of various covariates. In this paper, seven of these models are featured. Each model was evaluated on BIC and MAPE, as well as the interpretability of the model parameters. The final model selected is the Weibull-gamma with all five covariates and two segments with heterogeneity on the company disputes and unemployment covariates. Though it was not the most parsimonious, it had the lowest BIC, which indicates that the improvement in log-likelihood was worth the additional parameters. In addition to this model, the latent-class Weibull model with heterogeneity on seasonality is considered, as it has the lowest BIC value among the Weibull models and is more parsimonious than the final model.

Model	LL	BIC	R^2 Incremental	MAPE
2 Seg WG w/ 5 Covariates	-254102.9	508263.4	0.981	7.17%
2 Seg Weibull w/ 4 Covariates	-254218.4	508481.1	0.966	8.48%

Applications of these models are demonstrated. First, by analyzing model parameters, a compelling argument can be made that the particular customer acquisition pattern was partially due to product awareness. Secondly, both models were able to capture the acquisition impact of covariates like seasonality. These models exemplify the practicality and effectiveness of this approach to modeling timing data.

2. Initial Observations and Covariate Selection

Below is the quarterly acquisition data for DISH Network (Figure 2.1, 2.2).

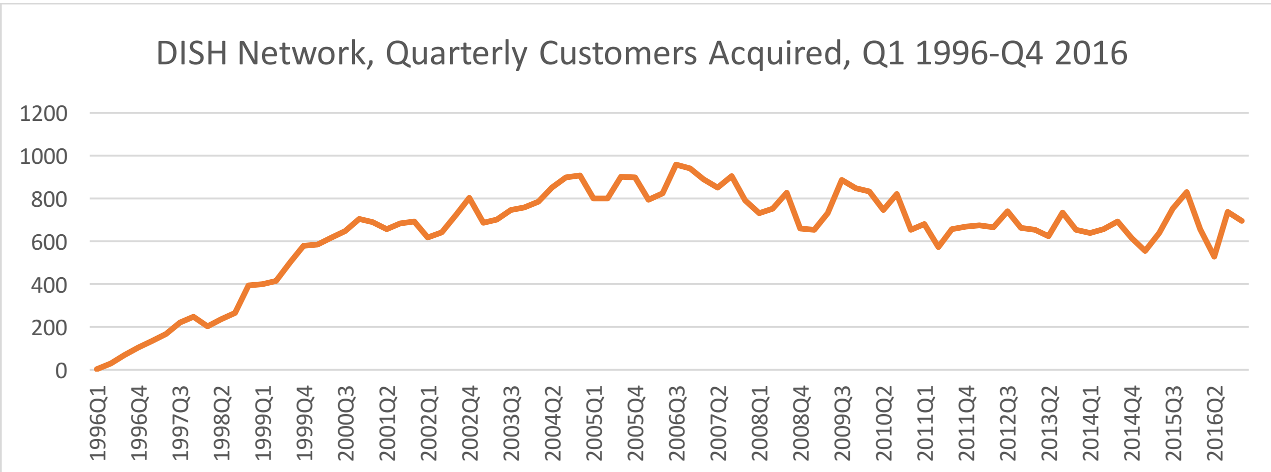

Figure 2.1: Incremental customer acquisition per quarter. Fluctuations are more evident in last two years.

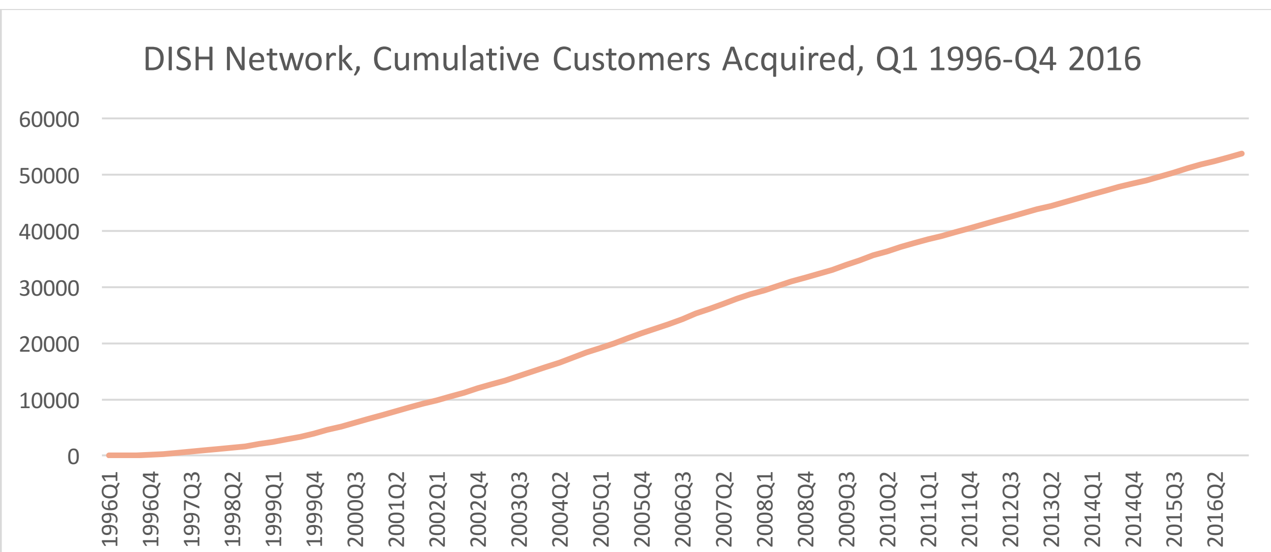

Figure 2.2: Cumulative customer acquisitions. Growth at the beginning looks exponential and eventually looks more logarithmic.

Following 2006Q3, the number of quarterly subscribers acquired began to slip and the incremental subscriber acquisitions became bumpier, particularly between 2009 and 2011 as well as between 2015 and 2016.

Through research, several predictors were incorporated into constructing the models:

1. Seasonality. Based on the academic calendar and its implications for months in which families are together as well as Christmas season, an indicator variable was included to treat Q1 and Q2 as the low season and Q3 and Q4 as high season.
2. Consumer Confidence Index (CCI).[1] CCI is defined as the amount of economic optimism expressed through spending activities.
3. Unemployment.[2]
4. Quarterly percent change in subscriber acquisition cost.[3] Subscriber acquisition cost represents the cost of convincing a potential customer to buy a service.
5. Disputes and dropped channels.[4][5] Since the early 2000s, DISH Network has been involved with many high profile disputes that resulted in dropped channels. An indicator variable was included for the quarters in which such disputes occurred.

A caveat to the models is that acquisitions are considered as one-time acquisitions. Preferably, acquisition data could be partitioned into first-time subscribers and subscribers who have unsubscribed before since repeat and trial purchasing is different. For the purposes of this analysis, the assumption is safe because most people aren't repeat subscribers.

3. Model Building and Selection

Seven models were built for assessment that incorporated segment and covariate effects:

1. Weibull-Gamma with 5 covariates. The Weibull-Gamma model assumes that the individual time to subscription is Weibull-distributed with parameter λ and λ is gamma-distributed with parameters *r* and alpha.
2. Weibull-Gamma with 5 covariates, 2 segments. In the two segment Weibull-Gamma, there is heterogeneity both within each segment and between the segments.
3. Weibull-Gamma with 5 covariates, 2 segments, $\beta_{unemployment}$ and $\beta_{seasonality}$ heterogeneous. The β values vary across the 2 segments.
4. Weibull-Gamma with 5 covariates, 2 segments, $\beta_{unemployment}$ heterogeneous.
5. Weibull with 4 covariates. In the case of subscribers, we want to use the Weibull to capture duration dependence; that is, given that a purchase hasn't occurred by time *t*, the probability of making the purchase between time *t* and *t** changes.
6. Weibull with 4 covariates, 2 segments. In this model, there are two subpopulations which have Weibull-distributed individual time to first purchase but the λ is static within each group.
7. Weibull with 4 covariates, 2 segments, $\beta_{seasonality}$ heterogeneous.

The model parameters are included in Figure 3.1. Evaluation metrics are included, such as Log-likelihood, BIC, R^2, and MAPE.

---

[1] https://fred.stlouisfed.org/series/UMCSENT/
[2] https://fred.stlouisfed.org/series/NROU
[3] http://dish.client.shareholder.com/sec.cfm?DocType=&Year=&CIKPassed=
[4] https://en.wikipedia.org/wiki/Criticism_of_Dish_Network
[5] https://www.tvchannellists.com/List_of_LineUp_Changes_on_Dish_Network_from_the_2000s

	1	2	3	4	5	6	7
	WG_5Cov	WG_5Cov_2seg	WG_5Cov_2se	WG_5Cov_2seg	W_4Cov	W_4Cov_2seg	W_4Cov_2seg
r_1	623.33621	62.97579387	54.6230697	63.72391592			
r_2		1598.187893	2019.42995	2136.205114			
α_1	21549147	2.242E+13	2.2799E+14	2.32526E+14			
α_2		12002870.05	23418912.4	20461410.5			
c_1	2.2952635	5.967289683	6.77834727	6.706721013	2.291935199	3.114969211	2.167957802
c_2		2.165137362	2.13633793	2.130078092		2.172543085	3.08276857
λ_1					2.66802E-05	1.00E-06	0.000210686
λ_2						0.000222627	0.000001
π_1		0.20102835	0.12735819	0.147725229		0.683046933	0.303240685
π_2		0.79897165	0.87264181	0.852274771		0.316953067	0.696759315
β_UE_1	-0.006379	-0.009217527	-0.24964217	-0.187606385			
β_UE_2			0.03124443	0.025452368			
β_S_1	0.0973645	0.073693284	0.22254797	0.061503038	0.09689507	0.075739354	0.003261814
β_S_2			0.04873738				0.07492546
β_D_1	-0.102888	-0.097084592	-0.10538314	-0.091721788	-0.103372085	-0.084317735	-0.085595265
β_D_2							
β_CC	0.1293516	0.050336198	0.0754058	0.063161915	0.135840734	0.08209918	0.09300583
β_AC	-0.160171	-0.178988448	-0.31425509	-0.352375307	-0.181366782	-0.311947569	-0.455695289
LL	-254352.4	-254146.2774	-254100.126	-254102.8943	-254351.6816	-254280.6151	-254218.4154
BIC	508740.29	508345.7246	508262.283	508263.3891	508729.948	508601.1075	508481.1389
R^2 Inc	0.9544669	0.973775165	0.98180205	0.981061721	0.954686759	0.960100847	0.965680702
MAPE Inc	10.661729	8.924694469	7.17453785	7.137900779	10.46155278	9.179907823	8.476071497
#params	8	12	14	13	6	9	10

Figure 3.1: Model parameters and metrics

To pick a final model, both BIC and MAPE are used, with MAPE more indicative of in-sample fit and BIC more indicative of out-of-sample fit since it adjusts for over-fitting. To assess the significance of including segments, and heterogeneity on the betas, the following LRT test results on the Weibull-Gamma and Weibull models are shown in Figure 3.2:

LRT Test	1\2	2\4	4\3	5\6	6\7
LL_small	-254352.4239	-254146.2774	-254102.8943	-254352	-254281
LL_big	-254146.2774	-254100.1257	-254100.1257	-254281	-254218
2*diff	412.2928781	92.30346998	5.53716644	142.133	124.3994
df	4	2	1	3	1
p-value	6.13785E-88	9.04807E-21	0.018616703	1.31E-30	6.89E-29

Figure 3.2: LRT results between model numbers specified in Figure 3.1

In both the Weibull-Gamma and Weibull models, adding in a segment was significant. Furthermore, introducing heterogeneity on $\beta_{unemployment}$ was significant in the WG models and introducing heterogeneity on $\beta_{seasonality}$ was significant in the Weibull models. Moreover, when we seek to add heterogeneity on $\beta_{seasonality}$, the log-likelihood gain is ~2.75. Though the LRT found this to be significant, the model is rejected because it's less parsimonious and the MAPE for a 2012Q1-2016Q4 holdout period for model 4 was 8.84% whereas it was 9.41% for model 3.

In comparing this same holdout period for model 4 and model 7, it's evident that model 7 had a worse MAPE during the period, as its MAPE was 13.66%. Besides the parsimony-lacking model 3, model 4 has the lowest MAPE and lowest BIC of the models presented. Model 7 is a close runner-up. Both models have parameters with logical interpretability. Though model 7 has better parsimony, the BIC for model 4 is lower, and because BIC takes number of parameters into account, model 4 is the better model. Additionally, model 3's fit for 2015Q1 onwards misses the actual quarterly acquisition count by a large margin, which indicates that there may be problems in forecasting future periods.

4. Model Results

For the Weibull-Gamma model, the *r* and *alpha* parameters are very large for both segments, indicating that for each segment, there's a homogeneous distribution of λ parameters throughout the subpopulation; everyone in the subpopulation has an equal rate of joining. In this regard, it seems appropriate to run the model using a Weibull model with segments and covariates instead. This was not done because the interpretation of this model, one where there's homogeneity across subpopulations, does not actually happen in reality. This model takes that heterogeneity across subpopulations into account, so segments and covariates were fit on the Weibull-Gamma model. For the Weibull-Gamma model, there are two pretty large segments of the population- one 14.8% and the other 85.2%. Both segments feature time dependence variable $c > 1$, which indicates that over time, people become increasingly service-aware and likely to subscribe.

In contrast, the Weibull model featured two segments with different λ and *alpha* parameters. In both segments, the duration-dependence variable $c > 1$ and λ is small. In this case, this means that subscribers have a low initial acquisition propensity but that propensity grows quickly over time. This makes sense for DISH Network subscriptions, as people may not have an interest in purchasing the subscription initially but over time, that interest grows as people feel like they're increasingly missing out by not having the channels afforded by the subscription.

Next, we consider the regression coefficients for both the Weibull and Weibull-Gamma models. For the Weibull-Gamma model, unemployment was broken down into two segments. In the first segment that is 14.8% of the population, increased unemployment has a strongly negative effect on subscriber acquisitions. In the other segment of the Weibull-Gamma model, increased unemployment has little effect on acquisitions. In the first segment, the more the economy takes a downturn, the less likely these people will subscribe to DISH. However, in the second segment, the exacerbating economic conditions and rising unemployment have little effect.

In the Weibull-Gamma model, the high seasons (Q3 and Q4) resulted in higher subscriptions. This makes sense, since Christmas season is in Q4 and sales for services tends to increase during that time. In Q3, families and students settle down for the new academic year, leading to higher subscriber acquisitions.

In the Weibull model, seasonality was broken down into two segments. In the first segment of size 30.3% of the population, high seasonality had a tiny, positive effect. In the other segment, high seasonality had a much larger positive effect. This is reasonable, as the first segment could feature a subpopulation that is much less swayed by deals and promotions that happen more often in the second half of the year. They could also be more stationary and not dependent on the academic calendar.

For both models, CCI had a positive relationship with acquisitions. This is rather obvious; as the degree of economic optimism expressed through spending activities increases, it is more likely that consumers will make more retail and service purchases.

In both models, DISH disputes had a negative relationship with acquisitions. Since the early 2000s, DISH's customer satisfaction declines every time there is a high-profile case of dropped channels. Nonetheless, this negative relationship makes sense, as potential subscribers are much less likely to join the subscription if they are aware of dropped TV channels, which they may find unprofessional or if they are heavy viewers of the dropped channel. One example case of this was in 2016Q2, when DISH Network's dispute with Tribune broadcasting led to dropped channels and resulted in a blackout in DISH's broadcasting of the historic 2016 NBA finals as well as the 2016 Stanley Cup Finals. Many potential subscribers who were looking to join DISH during this quarter may have rescinded their decision to do so because of lacking coverage.

Lastly, in both models, quarterly percent change in customer acquisition cost had a negative relationship with subscriber acquisitions. For potential customers who are looking to switch cable subscriptions, a higher customer acquisition cost means that it costs more to convince them to make that switch to DISH. As this cost rises, it means that potential DISH subscribers need more convincing to join. Thus, this negative relationship with acquisitions makes sense.

Lastly, let's evaluate model performance. The incremental subscriber acquisitions are shown (Figures 4.1-4.2). The incremental fits for both models are still very impressive. Both models capture most of the incremental data's nuances and spikes, though the Weibull model was thrown off by the high volatility in 2015 and 2016. Note that the customer acquisitions numbers are in 100,000s.

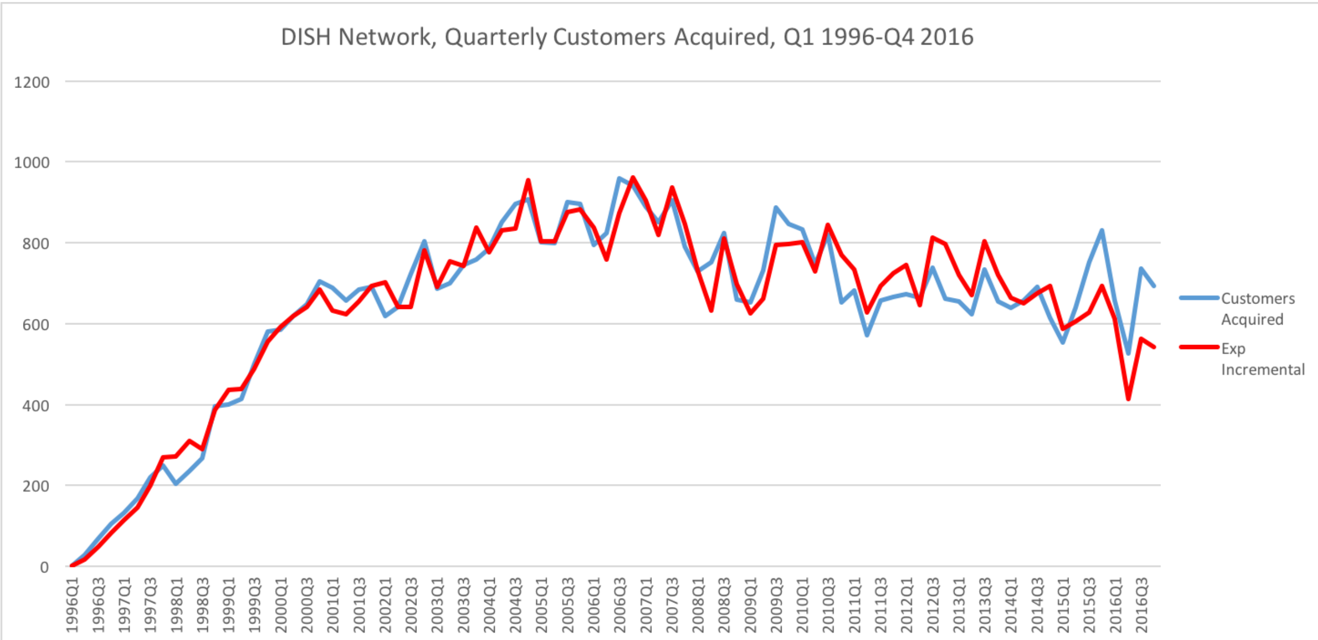

Figure 4.1: Incremental fit for Weibull. A very good fit except for the last two years.

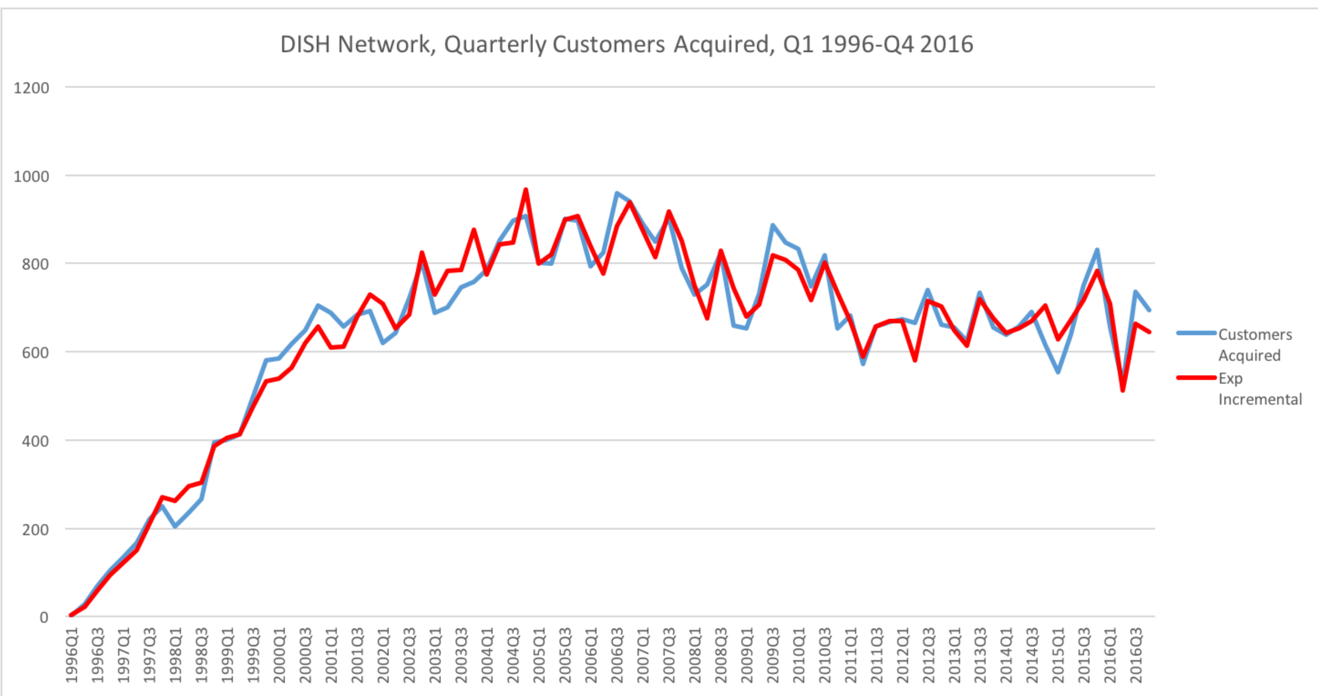

Figure 4.2: Incremental Fit for Weibull-Gamma. The last two years is captured very well.

5. Conclusion

The duration-dependence variable from the Weibull-Gamma and Weibull models has major implications for DISH Network. In both models, c is greater than 1, indicating an absence-makes-the-heart-grow-fonder relationship between DISH's subscription services and potential customers. The interpretation is that service-awareness is high, which explains why for individuals, the probability of trial given that they haven't yet subscribed is affected by time. This phenomenon is something that DISH should've taken advantage of with marketing campaigns in post-recession years, when economic growth improved and potential customers who were already service-aware became even more service-aware.

Other conclusions can also be drawn on the impact of the covariates on acquisitions. DISH Network has long advertised itself as a low-budget cable subscription service. In times of economic downturn, the Weibull-Gamma model seems to suggest that it may be better to appeal towards the segment of people who are less affected by downturn. Moreover, further analysis can be done on other covariates, such as the growth of other digital media platforms like Netflix.

Lastly, a forecast is generated for each of the 4 quarters in 2017 in Figure 5.1. [678] Percent change in customer acquisition costs was found by averaging the quarterly % change in the last 6 years.

	season_high	CCI	disputes	% change in CA cost	unemployn	W Incremental	WG incremental
2017Q1	0	9.72	1	0.0213	4.7	442.9063091	522.7753929
2017Q2	0	9.53	0	0.1223	4.7	430.8317788	514.7743167
2017Q3	1	9.5	0	-0.04511	4.8	471.1970781	533.2709587
2017Q4	1	9.2	0	-0.06451	4.8	432.8345725	483.6325262

Figure 5.1: Forecast for each quarter in 2017 for Weibull model and WG model

The forecast for 2017 customer acquisitions continues at a downward trend, though the Weibull model has a larger drop in acquisitions than the Weibull-Gamma model does. The quarterly acquisitions of the last 5 years seem to suggest that the drop will be subtler and not severe, as the Weibull model predicts.

[6] http://deadline.com/2017/03/hearst-television-warns-long-term-impasse-dish-network-dispute-1202042926/
[7] http://www.tradingeconomics.com/united-states/consumer-confidence/forecast
[8] http://www.tradingeconomics.com/united-states/unemployment-rate/forecast