Dish Network Subscriber Acquisition

MTKG776: Applied Probability Models in Marketing

### *2017-04-05*

Contents

Word Count: 1,856

1 Executive Summary

In this analysis we attempt to identify and implement a timing model that will best predict Dish Network's subscriber acquisition in 2017. Our starting dataset is quarterly subscriber acquisition from 1996 to 2016. Using a Weibull distribution to explain a customer's time to subscribe we attempt to account for heterogeneity with segments (finite-mixture models), including hard-core never-acquirers, and a continuous Gamma distribution. We find that a Weibull-Gamma model performs well but alone is insufficient to explain Dish Network's customer acquisition. By adding covariates that describe macroeconomic trends, consumer sentiment, seasonality, and the competitor Netflix's performance, we improve our explanation of Dish' Network's customer acquisition. The plot below presents Dish's actual incremental subscriber acquisition by quarter and our model and forecast (dashed) for 2017.
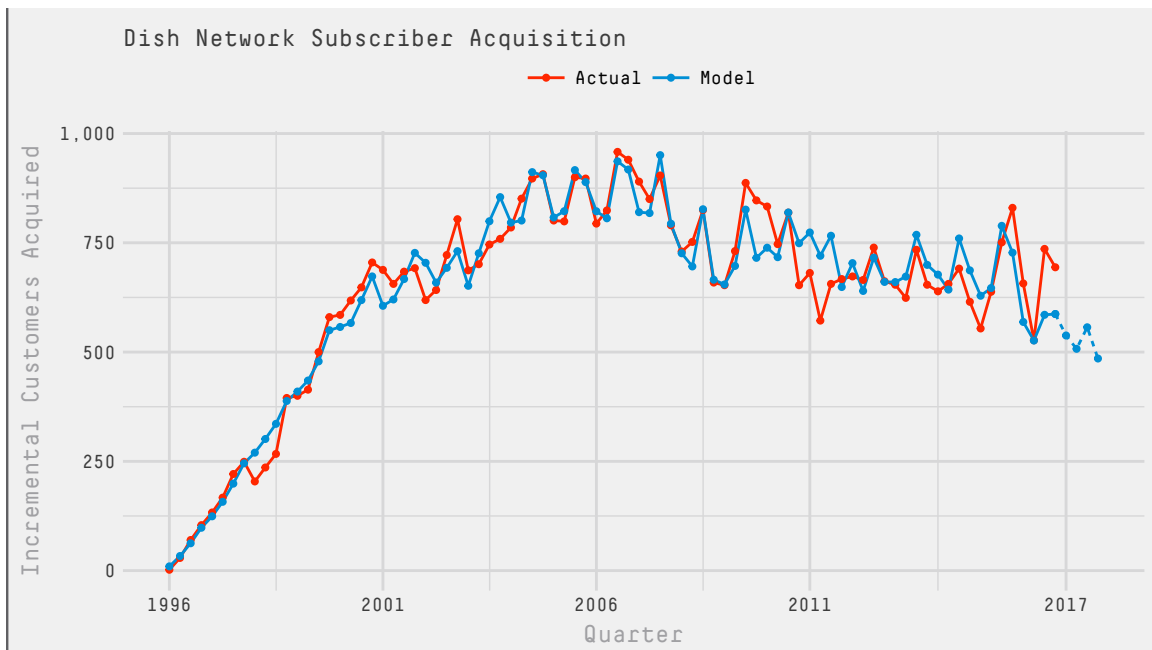
2 Analysis

2.1 Objective

Our objective is to build a timing model to forecast the quarterly customer acquisitions for Dish Network in 2017.

2.2 Candidate Models

The diagram below (Figure 1) provides a framework and an assessment of the timing models considered for this analysis. As a baseline, individual-level model the exponential distribution was not considered because it has no duration dependence. A Weibull was used to allow for duration dependence, the probability of the customer signing up for Dish now, given that they have not signed up yet, to change over time. If heterogeneity were included via a gamma distribution of rate parameter λ, the exponential-gamma distribution (i.e. Pareto II) has a decreasing hazard function which is neither expected for the Dish product nor evident by growth rate in customer acquisition in the data.

The remaining red X's represent models or factors that were attempted but were not selected in the final model. Finite-mixture models of Weibull distributions produced segments with nearly all 60M customers indicating there were not true segments, but rather the customer population was rather homogeneous. In addition, a finite-mixture model of Weibull-Gamma distribution with 2 segments produced one segment with nearly all the customers and another with none. The concept of hard-core never-acquirers was introduced with a vanilla Weibull and a Weibull-Gamma distribution, but both resulted in $\pi = 0$ and thus no evidence of a hard-core never-acquirer segment. Four categories of covariates were implemented: macro-trends, seasonality, firm-specific, and industry-specific. We found that the firm-specific covariate did not capture enough new information for forecasts to warrant inclusion.
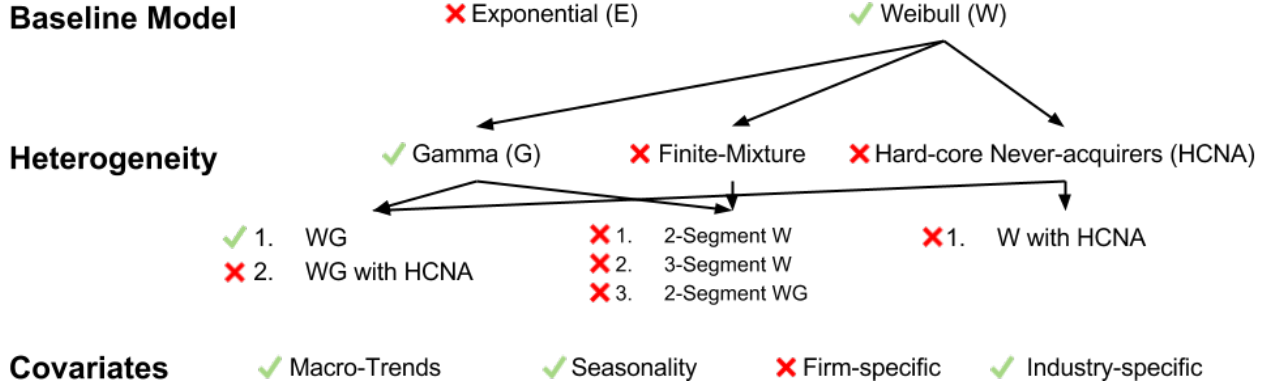

Figure 1: Candidate Models

The resulting model is a Weibull-gamma (i.e. Burr XII) with covariates model that has a cumulative density function given by:

$$P(T \leq t) = \int_0^\infty \left(1 - e^{\lambda B(t)}\right) \frac{\alpha^r \lambda^{r-1} e^{-\alpha \lambda}}{\Gamma(r)} d\lambda \tag{1}$$

$$= 1 - \left(\frac{\alpha}{\alpha + B(t)}\right)^r \tag{2}$$

where

$$B(t) = \sum_{i=1}^{t} \left(i^c - (i-1)^c \right) e^{\boldsymbol{x}(i)\boldsymbol{\beta}} \tag{3}$$

2.3 Covariates

The Weibull-Gamma (WG) model explains customer acquisition as each person in the population having some underlying, unobservable time to buy rate λ and a hazard function that changes over time with a shape determined by c. Furthermore, the model assumes that the rate parameter λ is distributed across the population according to a gamma distribution. Even with this individual-level story and expression of heterogeneity, we have reason to believe that Dish Network's customer acquisition may be influenced by the following external factors:

1. **Macro-Trends** - the performance of the US economy and how consumers feel
2. **Seasonality** - the WG does not distinguish Q2 to Q4, but consumers do
3. **Firm-Specific** - Dish may have taken actions (e.g. product launches) that can contribute to acquisition not governed by the WG
4. **Industry-Specific** - competitive forces or the changing TV environment may influence customer acquisition

So that each external factor, or covariate, can be compared all were scaled appropriately.

2.3.1 Macro-Trends

To capture macro-trends, we explored different indices published on FRED[1], the Federal Reserve Economic Data portal managed by the Federal Reserve Bank of St. Louis. We sought to find metrics that would influence Americans' propensity to start a new subscription service (likely transferring from another television service such as cable). We settled on two indices that capture slightly different phenomena:

1. **CANDH**[2]: This index is a component of the Chicago Fed's National Activity Index, which is "a weighted average of 85 monthly indicators of national economic activity".[3] There are multiple components of the index and CANDH encapsulates the Personal Consumption and Housing data series. CANDH includes data on retail sales, consumption of durable goods, and new housing starts. *It reflects what actually happens in the economy.*
2. **UMCSENT**[4]: This measure is of *consumer sentiment* and is produced by the University of Michigan[5] through a survey of consumers. UMCSENT gauges how people are *feeling*. It is perceived as a leading index of economic activity and does not necessarily reflect reality. The Consumer Confidence Index[6] by The Conference Board would have been preferred measure for this macro-trend but it is not a free dataset.

Below are the two indicators for the 21 years (84 quarters) of this analysis:
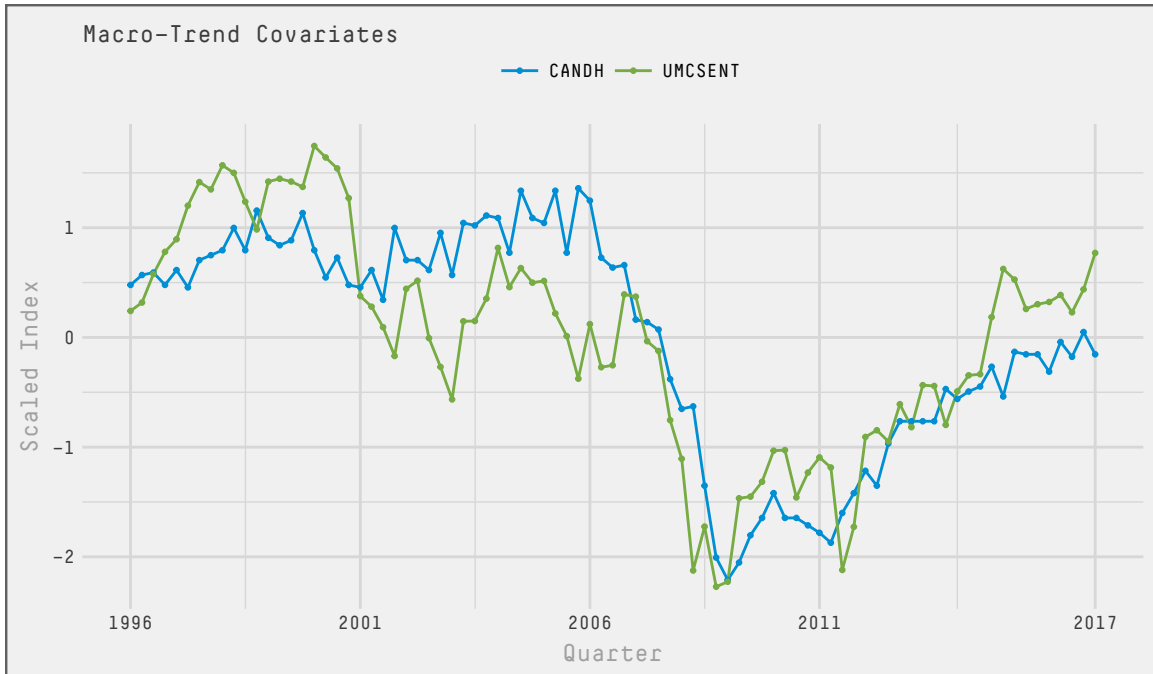
We see the indicators follow a similar trend - there is a significant fall during the great recession of 2007-2009 - but the levels differ over the course the time period.

2.3.2 Seasonality

To account for seasonality in Dish Network sign-ups, we decomposed the customer acquisition time series using STL[7] (Seasonal and Trend decomposition using Loess). Loess[8] is simply a type of local regression used for estimating non-linear relationships. Below is the decomposition into seasonal, trend, and remainder components:
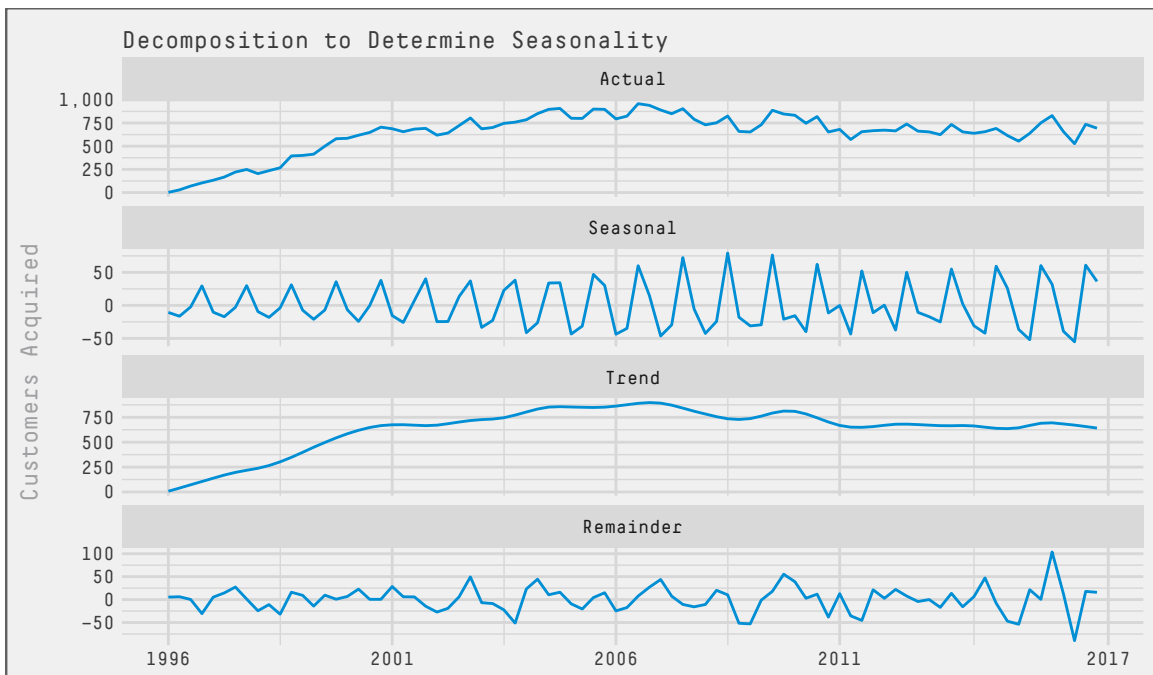
We scaled the seasonal component and used it as a covariate. The seasonality covariate proved extremely helpful in explaining the shape of the Dish customer acquisition series.

2.3.3 Firm-specific

It is reasonable to believe that there were specific actions taken by Dish Network that contributed to the acquisition of customers (or at least the company should hope so), such as product launches or marketing campaigns. There is a specific event that stands out in the time series: the launch of Sling TV in January 2016[9]. Sling TV was the first internet TV service to unbundle ESPN from a typical cable/satellite package and was aimed directly at "cord cutters". In many ways it was positioned as a secondary subscription to complement your Netflix or Hulu subscription. Bloomberg reported that Sling TV surpassed 600,000 subscribers in June 2016 and 1 million by October 2016[10].

To account for the pop in acquisitions during 2015 (that broke with the plateauing or downward previous trend), we created a Sling TV covariate that increased over the four quarters of 2015.

2.3.4 Industry-specific

In addition to actions taken by Dish, competitive forces in the TV industry likely contributed to changes in customer acquisition. A covariate we would have wanted to use for this notion is the number of subscribers to TV streaming services such as Netflix, Hulu, or Sony Vue (see Limitations for more details). However, given the idea that Netflix has stolen TV subscribers, or would-be-TV-subscribers in the case of millennials, from traditional cable and satellite companies we used an easily available Netflix dataset: Netflix stock price. Below is the Netflix stock price averaged by quarter and scaled to conform to the other covariates:
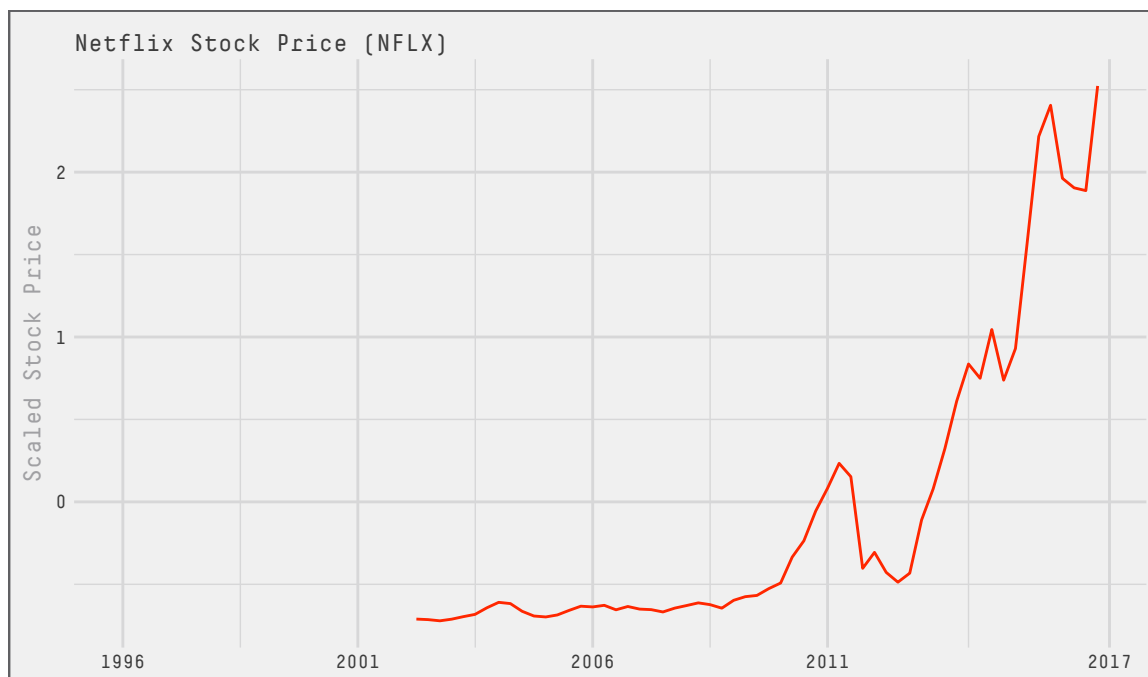

2.4 WG with Covariates

After determining the WG model was the most appropriate, we then needed to identify which combination of covariates, if any, would produce the best model. We ran the WG with coviarates model for all 32

combinations of the five covariates ($2^5 = 32$), which includes no covariates at all. Below are the top 10 models by BIC (a summary of the remaining 22 can be found in the Appendix):

Table 1: Top 10 Weibull-Gamma with Covariate Models by BIC

CANDH	UMCSENT	Seasonality	Sling TV	Netflix	# Params	LL	MdAPE	BIC
X	X	X		X	7	-254,171	0.0513	508,341,669
X	X	X	X	X	8	-254,171	0.0504	508,341,954
	X	X	X	X	7	-254,175	0.0527	508,350,902
	X	X		X	6	-254,175	0.0527	508,350,973
X		X		X	6	-254,196	0.0570	508,391,833
X		X	X	X	7	-254,196	0.0568	508,391,862
X	X			X	6	-254,262	0.0695	508,523,488
X			X	X	6	-254,271	0.0709	508,541,716
		X	X	X	6	-254,271	0.0619	508,542,224
		X		X	5	-254,271	0.0618	508,542,227

First, we note that the log-likelihood (LL), the median absolute percent error ($MdAPE$), and BIC are all relatively similar for these top models. Second, we note that for all 10 models, the Netflix coviarate appears. Third, we see that the seasonality covariate is present in almost all of the top models. Here we find that Sling TV does not add that much more information. For the second model with all covariates compared to the top model with Sling TV, the $MdAPE$ is slightly lower, LL is nearly the same, and BIC is marginally worse. In fact, the first two models and the second two models nearly look the same graphically. As such, the first and third models above are shown in the incremental and cumulative tracking plots below:


6

Cumulative Customer Acquisitions
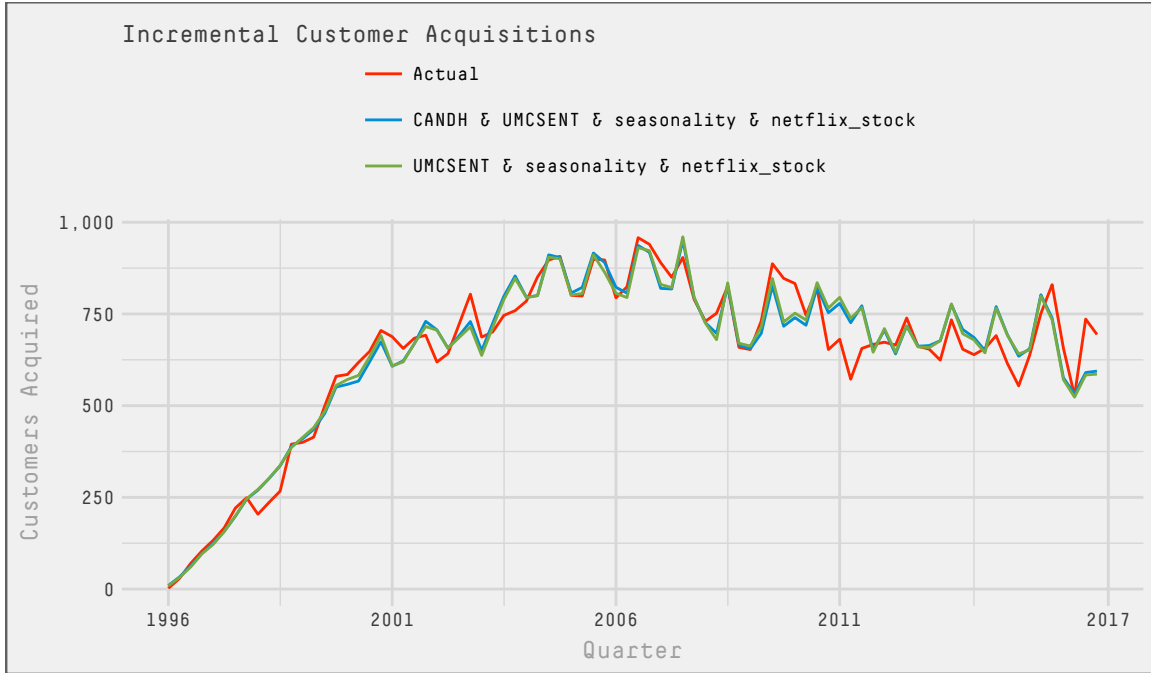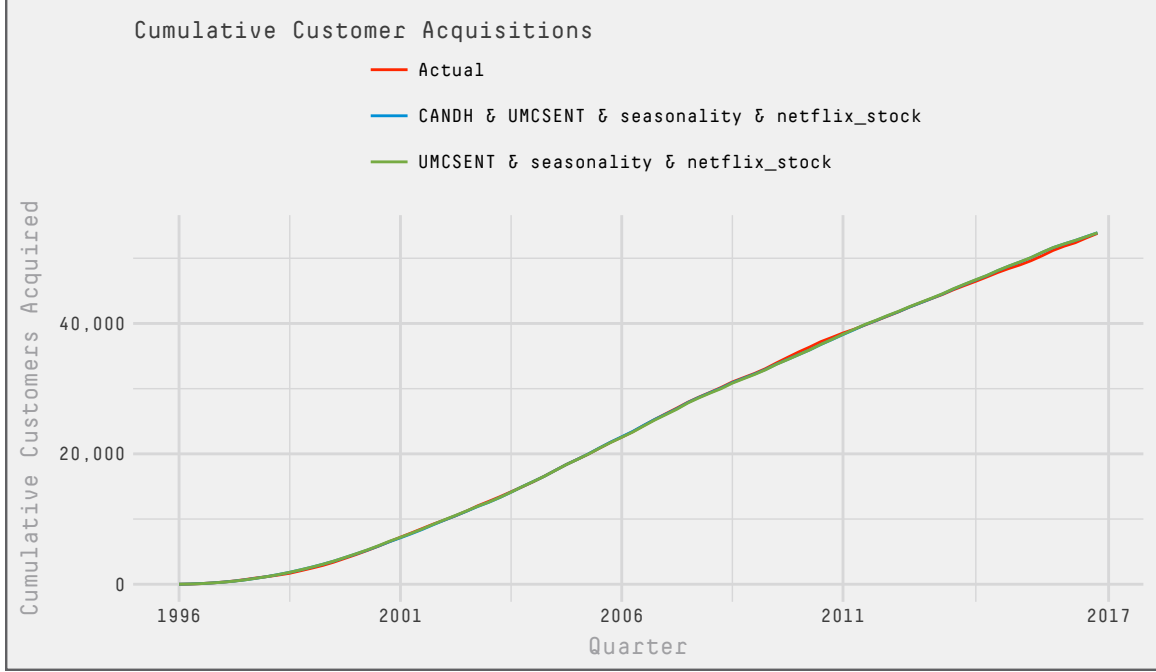
With regards to the model parameters r, α, and c, the table below shows these estimates for the top 10 models by BIC:

Table 2: WG Parameter Estimates for Top 10 Models by BIC

| CANDH	UMCSENT	Seasonality	SlingTV	Netflix	r	alpha	c	BIC
X	X	X		X	5,063	32,072,368	2.158	508,341,669
X	X	X	X	X	315	1,979,103	2.155	508,341,954
		X	X	X	35	234,505	2.177	508,350,902
		X		X	32	215,771	2.179	508,350,973
X		X		X	10,803	58,506,929	2.120	508,391,833
X		X	X	X	2,739	14,954,479	2.122	508,391,862
X	X			X	374	2,262,806	2.148	508,523,488
X			X	X	18,004	97,994,911	2.121	508,541,716
		X	X	X	7	36,954	2.140	508,542,224
		X		X	7	37,165	2.139	508,542,227

We find extremely large values for r and α, which confirms our belief that there is not much heterogeneity in the population and why none of the latent-class models were successful. The plot below shows the distribution of the rate parameter λ for the first model:

Estimated Distribution of Lambda
WG with CANDH, UMCSENT, Seasonality, and Netflix Stock
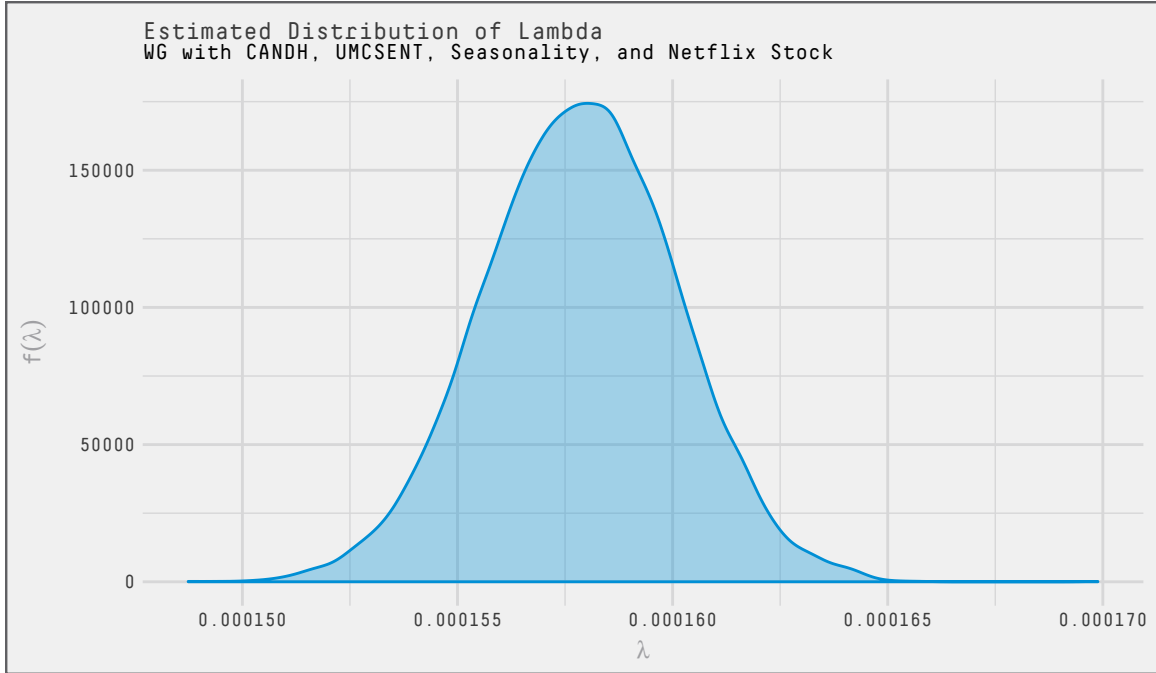
We see that the distribution of λ is nearly symmetric with a large number of people having the same rate parameter value. There is some heterogeneity, but there is are not a significant portion of people with high or low λ's. In other words, we would not consider using a pure Weibull with covariates upon seeing plot, but segments do not appear fruitful either.

Next, we review the hazard function of the WG with covariates model. We are surprised to find values of $c > 2$ as this is rare. A value of $c > 1$ whose hazard function is (monotonically) increasing implies that purchase rate increases over time at the individual level. The functions are not monotonically increasing due to the covariates. The plot below shows the hazard function for the top 9 models, all with increasing hazard functions. The implication is that given you have not subscribed to Dish yet, the probability that you subscribe now increases over time.

Hazard Function for Top 9 Models

8
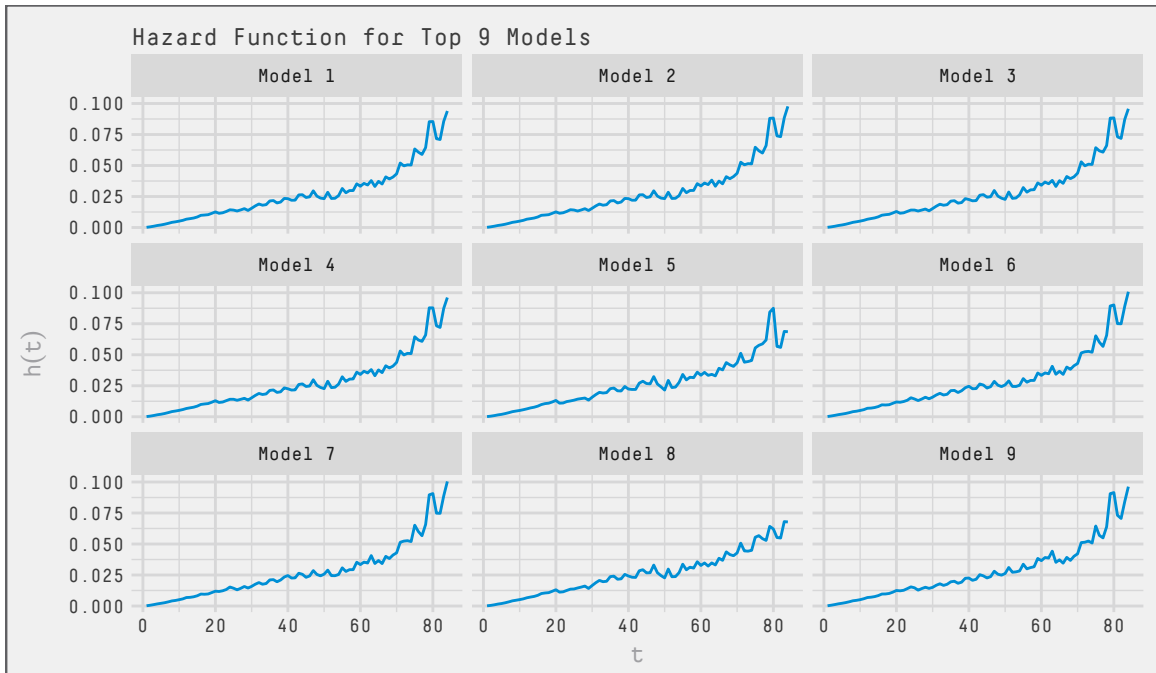
3 Results

3.1 Final Model

As our final model, we select the candidate model with the lowest BIC. To ensure that an additional parameter is necessary, we use the likelihood ratio test ($df = 1$) with the next-best-model that does not include the covariate CANDH. We find a very small p-value and conclude that the models are not the same and thus the model with four covariates is valuable and confirm this as our final model.

Table 3: Likelihood Ratio Test for Next-Best-Model

CANDH	UMCSENT	Seasonality	Sling TV	Netflix	# Params	LL	LRT	p-value
X	X	X		X	7	-254,170.8	9.3213	0.0023
	X	X		X	6	-254,175.4		

This model does not include the artificial Sling TV coviariate. We are glad to remove this coviariate as it will not be helpful for future prediction. Below is a summary of our final model:

Table 4: Summary of Final Model

	Value
r	5,063
α	32,072,368
c	2.1582926
Covariates	CANDH , UMCSENT , seasonality , netflix_stock
Covariate βs	0.0257, 0.0724, 0.0562, 0.1572
LL	-254,170.8
BIC	508,341,669

3.2 Forecast 2017

In order to make predications of Dish Network's subscriber acquisition in 2017, the covariates needed to be carried into the future. The table below summarizes how each covariates values were estimated for 2017:

Covariate	Implementation
Macro-Trends	Index values for CAND and UMCSENT were released for the months in Q1 2017 and were used. For 2017Q2 to 2017Q4, we created an ARIMA model[11] using all the quarterly data from 1996Q1 to 2017Q1 and predicted 3 periods ahead. An automatic method for selecting the parameters of the ARIMA model was used.
Seasonality	Using the STL decomposition, the time series was forecasted 4 periods ahead and the seasonal component was used.
Firm-Specific	Not necessary as not included in model
Industry-Specific	Netflix stock prices from 2017Q1 were used and then an (automatically selected) ARIMA model using all the quarterly stock prices from 1996Q1 to 2017Q1 were used to predict 3 periods ahead.

The plot below shows the forecasts of the four covariates:

Forecasts of Covariates
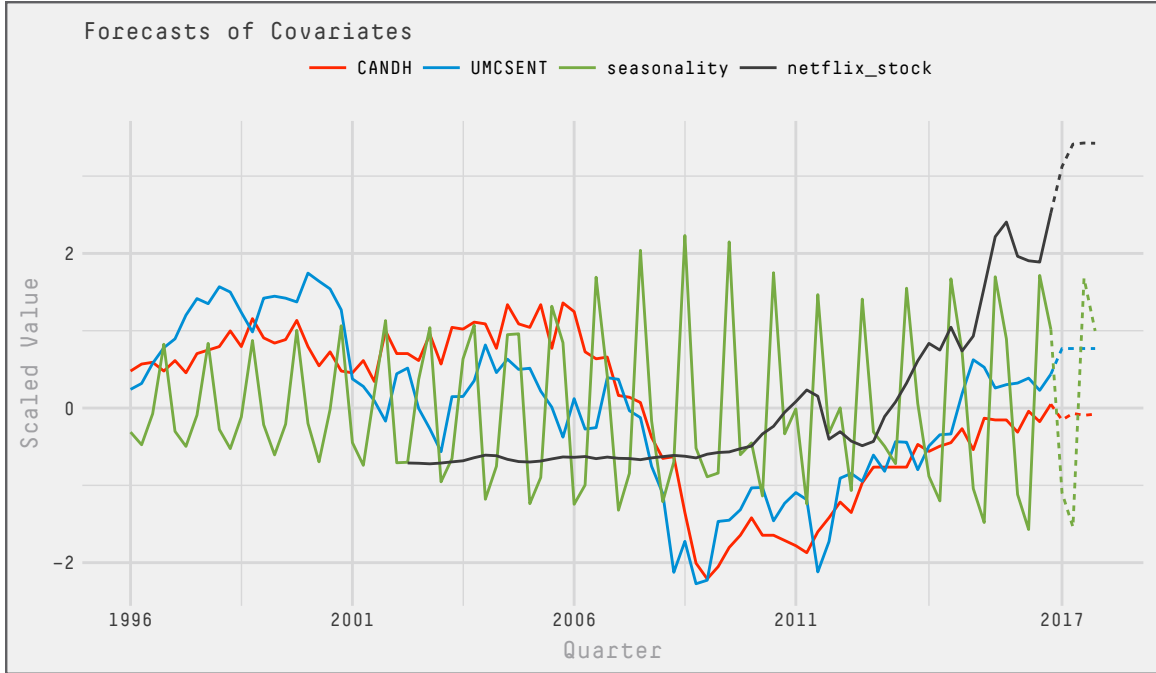
Incorporating these forecasts, we forecast the number of customers acquired in each quarter to be

Table 6: Forecasted Customer Acquisition

Quarter	Customers Acquired
2017Q1	538
2017Q2	507
2017Q3	557
2017Q4	485

The figure below shows the incremental and cumulative tracking plots with the forecasts:

Dish Network Subscriber Acquisition
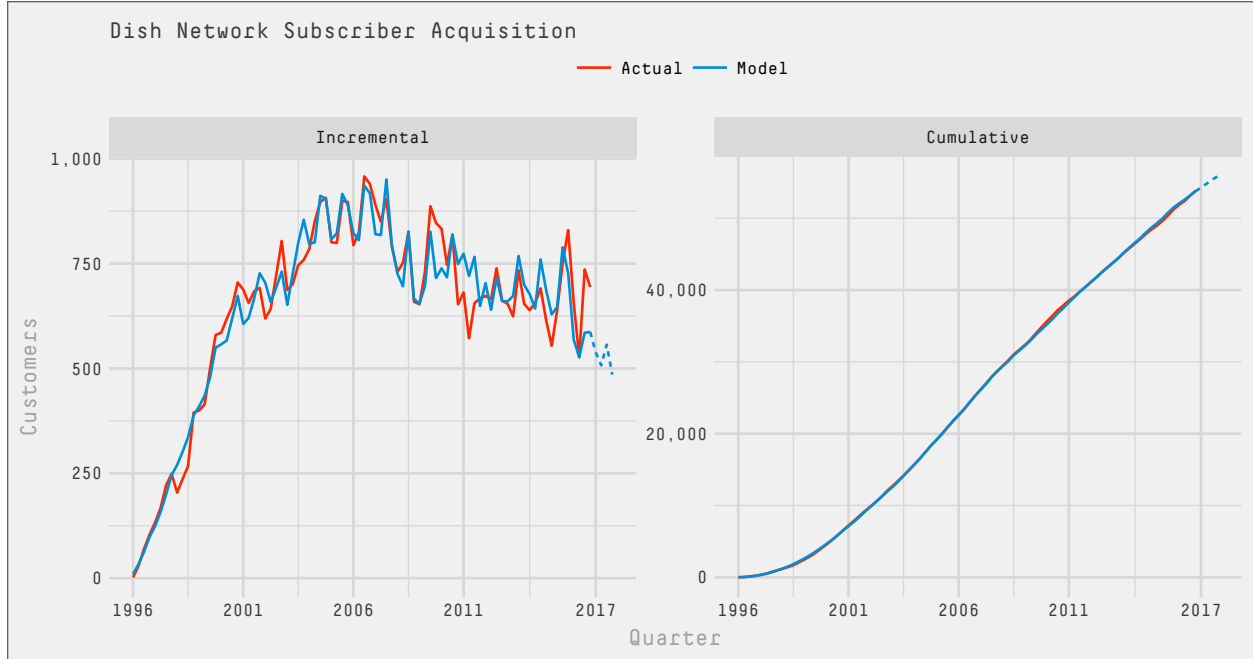
4 Limitations

1. **Population Size**. In this analysis we used the assumption that the overall customer population (N) was 60M. In a subsequent analysis we would attempt to implement (1) a truncated model and (2) vary N to identify its impact on parameter estimates.

2. **Better Metric for Industry-Specific Covariate**. We would have preferred to use the number of subscribers for all US TV streaming services. A secondary measure could have been revenue for Netflix, but this was only readily available for recent time period. A future analysis would implement subscriber information from a market data firm such as Second Measure[12].

3. **Jump in 2016Q3 - 2016Q4 Acquisitions**. The Sling TV covariate sought to capture the impact of subscribers to this new service. The artificial covariate seemed reasonable to cover 2015, but given the drop-off in Q1 and Q2 and lack of product news in 2016 did not seem reasonable to cover 2016. However, the two periods at the end of 2016 indicate a phase shift may have occurred and could cause our forecasts, which are much lower, to be significantly different.

4. **Forecasting Covariates**. The fact that we forecasted covariates is likely to cause issues with our forecast of the three remaining periods of 2017.

5. **Segmenting** c. While we attempted segments with Weibull and Weibull-Gamma models, we did not attempt to segment c in isolation while keeping λ and r fixed.

5 Appendix

5.1 Bottom 22 Models

Table 7: Bottom 22 Weibull-Gamma with Covariate Models by BIC

CANDH	UMCSENT	Seasonality	Sling TV	Netflix	# Params	LL	MdAPE	BIC
X				X	5	-254,271	0.0720	508,543,009
	X		X	X	6	-254,276	0.0680	508,552,138
	X			X	5	-254,277	0.0685	508,553,220

CANDH	UMCSENT	Seasonality	Sling TV	Netflix	# Params	LL	MdAPE	BIC
X	X		X	X	7	-254,314	0.0781	508,627,740
X	X	X	X		7	-254,316	0.0564	508,632,728
			X	X	5	-254,351	0.0693	508,701,777
				X	4	-254,352	0.0697	508,703,386
X	X	X			6	-254,389	0.0623	508,778,411
	X	X			5	-254,418	0.0594	508,835,994
X	X				5	-254,511	0.0866	509,021,819
X			X		5	-254,605	0.1181	509,210,138
X		X	X		6	-254,610	0.0804	509,220,895
		X	X		5	-254,624	0.0933	509,248,431
X	X		X		6	-254,655	0.0756	509,310,904
			X		4	-254,695	0.0898	509,389,507
	X				4	-254,711	0.0839	509,422,333
		X			4	-254,787	0.1032	509,574,612
X					4	-254,806	0.1038	509,612,108
	X		X		5	-254,828	0.0844	509,656,245
					3	-254,894	0.1136	509,788,987
	X	X	X		6	-254,936	0.1045	509,872,875
X		X			5	-255,424	0.1207	510,847,721

5.2 Technical Notes

All MLE optimization of the Weibull-Gamma with covariates model ewas performed using nlminb[13].

The full source code that created this document can be found in the Github repo.

References

[1] https://fred.stlouisfed.org/

[2] https://fred.stlouisfed.org/series/CANDH

[3] https://www.chicagofed.org/~/media/publications/cfnai/background/cfnai-background-pdf.pdf

[4] https://fred.stlouisfed.org/series/UMCSENT

[5] https://data.sca.isr.umich.edu/fetchdoc.php?docid=24774

[6] https://www.conference-board.org/data/consumerconfidence.cfm

[7] https://www.otexts.org/fpp/6/5

[8] https://en.wikipedia.org/wiki/Local_regression

[9] http://www.theverge.com/2015/1/5/7491071/dish-sling-tv-ott-internet-tv-announced-ces-2015

[10] https://www.bloomberg.com/news/articles/2016-10-26/dish-s-sling-tv-service-seen-exceeding-1-million-subscribers

[11] https://en.wikipedia.org/wiki/Autoregressive_integrated_moving_average

[12] https://secondmeasure.com/

[13] https://stat.ethz.ch/R-manual/R-devel/library/stats/html/nlminb.html