# Delivering Groceries in as Little as 1 Hour

Modeling the Impact of Coronavirus on Grocery Delivery Services

Jessica Nguyen

Word Count: 2314

## Table of Contents

Executive Summary .......... 2

I. Objective and Managerial Questions .......... 2

II. Data and Covariates .......... 2

III. Building a Model .......... 4

a. The Story and Hypotheses .......... 4

b. Other Considerations While Model Building .......... 6

c. Model Selection and Preliminary Results .......... 8

IV. Final Model .......... 11

a. Interpretations from Final Model .......... 12

b. Validation .......... 13

c. Limitations and Future Areas of Improvement .......... 15

V. Answering Managerial Questions .......... 15

Appendix A: Covariates Exploratory Data Analysis .......... 17

Appendix B: Plot of Incremental Downloads .......... 19

Appendix C: Checking Robustness by Estimating on Subsets .......... 20

## Executive Summary

A Burr XII model with covariates and time-shifted latent classes is used to better understand segments of Instacart downloaders and some factors that are correlated with downloads. The model suggests that there are two distinct segments of consumers. One segment would have downloaded the app regardless of the coronavirus. The other segment potentially downloaded the app as a response to the epidemic. This "coronavirus" segment responds more to coronavirus-related covariates such as number of U.S. cases and Google Trends. Other covariates including signals of economic health (S&P 500) and government restrictions (ban on European travel) are also explored.

## I. Objective and Managerial Questions

The objective of this project is to build a model that would help managers understand how coronavirus has impacted downloads and who consumers are. The model answers:

1. Is there a new segment of consumers who have started downloading since the epidemic started?
2. How alike or different are consumers?
3. What are some factors that are correlated with number of downloads?

## II. Data and Covariates

In addition to the five covariates provided in the assignment, four others were considered:

1. **Closing price of the S&P 500[1]:** This variable was explored because declining stock prices may be a signal of emergency situations. In emergency situations, consumers may

[1] https://www.wsj.com/market-data/quotes/index/SPX/historical-prices

be more inclined to download Instacart because they fear lockdowns. One assumption made was that for the weekend where prices were not available, the price was set as the average of the previous and next business days' prices. For example, the price for Sundays was the average of the price on last Friday and the price on next Monday.

2. **Closing Price of Crude Oil per Barrel (in USD)[2]:** The same assumptions were made as for the S&P 500.

3. **Number of state of emergencies (SoE) declared on a day[3]:** A state of emergency may be declared to free-up federal assistance.[4] This variable was only active for a portion of the dataset. The first SoE was declared on February 29. By March 16th, all fifty states had declared a SoE.

4. **Effect of travel restrictions to Europe[5]:** This variable is similar to the "9/11" variable presented in class. It represents the residual effect of travel restrictions. It represents a story that the restrictions placed on March 14th were correlated to increased downloads because the initial news is unsettling to consumers, potentially prompting them to download the app immediately. However, the effects overtime should decay as people become accustomed to the new situation. The γ parameter that represents the "new normal"; the δ parameter that indicates the speed until the new normal is reached:

$$\boldsymbol{x = (1 - \gamma(1 - e^{\delta(t - t_{EuropeBan})})}$$

[2] https://markets.businessinsider.com/commodities/oil-price?type=wti
[3] https://www.businessinsider.com/california-washington-state-of-emergency-coronavirus-what-it-means-2020-3, https://www.wvpublic.org/post/wva-governor-declares-state-emergency-response-coronavirus-threat , https://www.cnn.com/2020/03/15/politics/oklahoma-governor-deleted-tweet-coronavirus/index.html
[4] https://www.businessinsider.com/california-washington-state-of-emergency-coronavirus-what-it-means-2020-3,
[5] https://www.bbc.com/news/world-us-canada-51883728

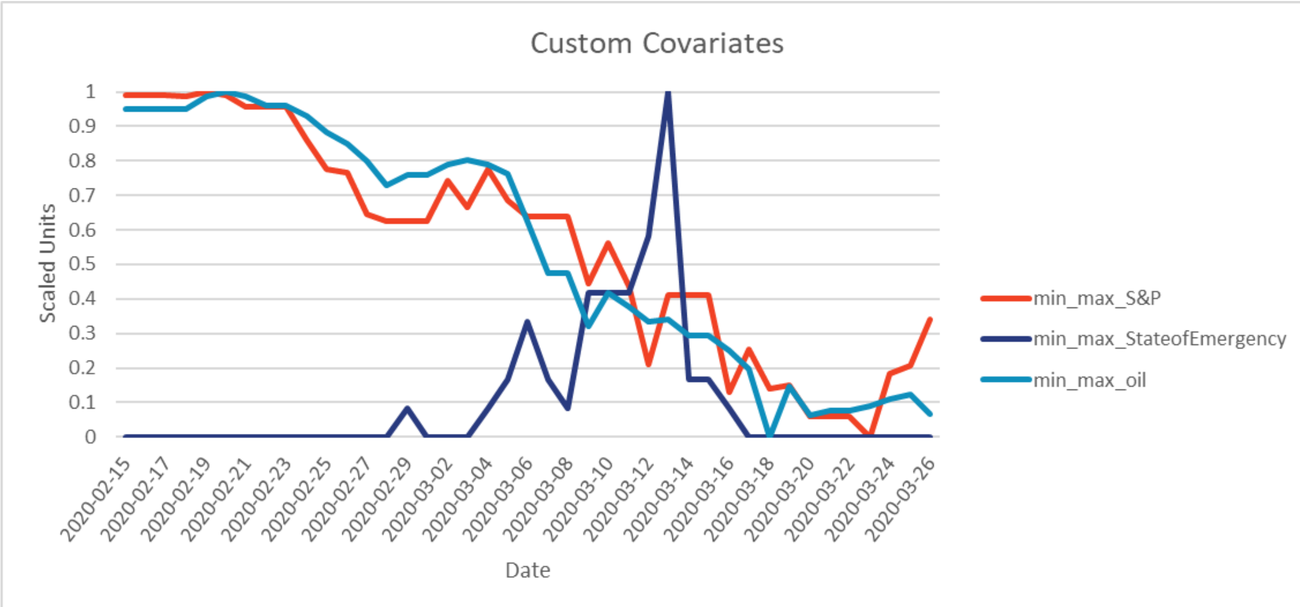



*Figure 1 Custom Covariates*[6]

All covariates were scaled using min-max method:

$$x_{scaled} = \frac{x - min(x)}{max(x) - min(x)}$$

Scaling was done to allow for easier comparison between covariates. Using a min-max method was preferable over a standardization method because it is not assumed that one standard deviation for one covariate is equivalent to the standard deviation of another covariate.

## III. Building a Model

### a. The Story and Hypotheses

The story is that there are two segments of consumers. The first segment includes people who would have downloaded the app, regardless of whether the pandemic happened. The second segment of people only found the app because of the pandemic (coronavirus segment). Therefore, I hypothesized that there would be two segments where one is time-shifted. I expect

[6] Refer to Appendix A for summary statistics and more graphs for covariates.

there to be little heterogeneity within the segments because groceries are needed by everybody and a pandemic has altered the way in which a large majority of the population gets their groceries. I also hypothesize that there will be positive duration dependence for both segments. However, the coronavirus segment will have higher duration dependence because as time goes on, more places institute stay-at-home orders and people are more likely to download the app. The model may show a large amount of duration dependence because it could be masking a social contagion affect. However, there could be many reasons why a model shows duration dependence and we cannot know for certain that there is social contagion.

When estimating coefficients for covariates, I estimated all of them separately for each segment EXCEPT for the economic covariates (S&P 500 and oil prices) and the European travel ban. Since the coronavirus segment downloaded only because of the pandemic, I expect this segment to respond differently to Google Trends and coronavirus cases. In contrast, I pooled some covariates because I believe that the covariates affect everyone in the same way, regardless of segment. I expect π (proportion of population between segments) to be high. I expect most people who have downloaded to be part of the coronavirus segment. I hypothesize (rank of importance listed in parentheses – 1 = most important, 8 = least important):

| | **Segment 1** | **Segment 2** |
|---|---|---|
| **r** | 0.9 | 1.2 |
| **α** | Who cares? | Who cares? |
| **c** | > 1 | > 1 |
| **π** | 0.85 | |
| **Google Trend – Social Distancing** | No effect | + (6) |

| | | |
|---|---|---|
| **Google Trend – Coronavirus** | No effect | + (6) |
| **Google Trend – Lockdown** | No effect | + (2) |
| **U.S. Cases** | No effect | + (1) |
| **World Cases** | No effect | + (5) |
| **S&P 500** | + (7) | |
| **Price of Oil** | +(8) | |
| **State of Emergencies Declared** | No effect | + (3) |
| **European Travel Ban** | No effect | + (4) |

*Figure 2: Hypotheses for Parameters*

I expect the covariates that are most visible to be most strongly correlated with downloads. For example, the media might emphasize U.S. cases and state of emergencies more – prompting consumers to panic and download.

## b. Other Considerations While Model Building

### Multicollinearity

There likely is multicollinearity. To check, I calculated the Variance Inflation Factor (VIF):

| **Covariate** | **VIF** |
|---|---|
| U.S. Cases | 57.14 |
| World Cases | 94.23 |
| "Coronavirus" | 30.39 |
| "Social Distancing" | 25.62 |

| | |
|---|---|
| "Lockdown" | 6.35 |
| State of Emergencies | 4.06 |
| S&P 500 | 20.12 |
| Oil Prices | 17.53 |

*Figure 3 Variance Inflation Factors*

A VIF greater than 10 is concerning. Therefore, I chose to remove the World Cases, “Coronavirus”, and oil prices variables from consideration. Although U.S. cases, “Social Distancing,” and S&P500 have a VIF greater than 10, I choose to keep them in my model because they are crucial to the story. This will be a limitation of my model. The model will be unable to distinguish between the separate effects of the three variables.

## Hardcore Never Tryers

It is possible that there are hardcore never-tryers. These people might be people who do not download apps. I will incorporate this into some of the preliminary models. However, I do not expect there to be many hardcore never-tryers because the coronavirus has affected a large majority of the population in a similar way.

## Out of Sample Validation Set

The dataset is limited because the pandemic is ongoing. Therefore, because the dataset is small and the full effects of the pandemic are not yet, I will not have out-of-sample testing.

## Time-Shifted Segment

One important consideration was determining when the new coronavirus segment kicks in. Rather than testing a bunch of different time periods and potentially p-hacking, I ex-ante chose to build the time-shifted segment starting on March 11$^{th}$. On this day, the World Health

Organization declared COVID-19 to be a pandemic.[7] This is a good day to choose because this announcement was widely covered by media. Therefore, the effects would have been felt for the coronavirus segment.

### March 4th Outlier

I assumed that increase on March 4th is a one-off occurrence because on March 5th, the trend appears to continue as it was before. It could be the case that March 4th is not a one-off occurrence and had residual effects that influenced future days. However, I chose to consider it an anomaly because through outside research, it is not obvious what happened on March 4th (or the prior days before) that would have caused such a spike in app downloads. It could be due some reason internal to the company (a promotion).

## c. Model Selection and Preliminary Results

I did not build and test out every single possible model. Instead, I chose to start with the basic models (Exponential, Weibull) then use LRTs to test out heterogeneity, then duration dependence, latent-classes, and finally covariates. Since I did not exhaustively build all options, it is unlikely that my model is the "best" out of all possible models.

| Model | # of Parameters | BIC | MAPE | MdAPE |
|---|---|---|---|---|
| Exponential | 1 | 21063952.2 | 87% | 91% |
| Exponential + Never Tryers | 2 | 21063970.81 | 87% | 91% |
| Pareto II | 2 | 21063970.77 | 87% | 91% |
| Weibull | 2 | 20712168.6 | 50.19% | 38.29% |
| Weibull + Never Tryers | 3 | 20712287.08 | 50.71% | 37.71% |
| Weibull + all covariates | 13 | 20478427.30 | 9% | 8% |
| Burr XII | 3 | 20831142.66 | 48.58% | 39.14% |

[7] https://www.justsecurity.org/69650/timeline-of-the-coronavirus-pandemic-and-u-s-response/

| | | | | |
|---|---|---|---|---|
| Burr XII + all covariates | 7 | 20478569.44 | 8.55% | 7.87% |
| Burr XII + select covariates + 2 segments + time-shifted | 14 | 20471488.70 | 6.40% | 5.16% |

*Figure 4 Summary of Models*

Heterogeneity

I used a LRT between the exponential and Pareto II models.

| | |
|---|---|
| 2*(LL Pareto II – Plain Exponential) | 0.002 |
| Degrees of Freedom | 1 |
| p-value | 0.967 |

*Figure 5 LRT for Heterogeneity*

The test shows that the models are not significantly different. Although this is aligned with my hypothesis, I hesitate to say that there is no heterogeneity because duration-dependence could also be masking heterogeneity. For these reasons, I will move forward with a heterogeneous model.

## Hardcore Never-Tryers

The LRT test between models with and without hardcover never triers considered there to be no difference. I think this is due to the homogeneity of needing groceries.

| | |
|---|---|
| 2*(LL Exponential-Never-Tryers – Plain Exponential) | 0.00 |
| Degrees of Freedom | 1 |
| p-value | 1 |

*Figure 6 LRT for Hardcore Never-Tryers*

## Duration Dependence

Duration dependence could be tested by performing a LRT:

| | |
|---|---|
| 2*(LL Weibull – LL Exponential) | 351802.1711 |

| | |
|---|---|
| Degrees of Freedom | 1 |
| p-value | 0 |

*Figure 7 LRT for Duration Dependence*

It indicates that there is duration-dependence. We should be cautious in this result because the large amount of duration-dependence may just be masking heterogeneity. However, as per my hypothesis, I do think that duration-dependence is part of this story.

## Covariates

I used a LRT between a model with all the covariates and a model with all but the ONE covariate that I was testing. Therefore, the null hypothesis is that the added covariate has no incremental significant effect with all the other variables.

| | **w/o Lockdown** | **w/o SoE** | **w/o S&P 500** | **w/o "Social Distancing"** | **w/o U.S. Cases** | **w/o Europe Ban** |
|---|---|---|---|---|---|---|
| **2* diff in LL** | 0 | 0 | 0 | 17726.86344 | 45394.49065 | 1961.5934 |
| **d.f.** | 1 | 1 | 1 | 1 | 1 | 1 |
| **p-value** | 1 | 0 | 1 | 0 | 0 | 0 |

*Figure 8 LRT for Covariates*

From the LRT, only "Social Distancing," U.S. Cases, and the European Travel variables are significant. I will include all these as well as S&P 500 in the final model (because it fits the story).

## Latent Classes

I cannot test use an LRT between a 2-segment and a 1-segment model because these models are not nested. However, I can look at the BIC; the two-segment model is superior.

| **Model** | **BIC** |
|---|---|

| | |
|---|---|
| 1-Segment WG + Covariates | 20478569.44 |
| 2-Segment WG + Covariates | 20471488.7 |

## IV. Final Model

I chose the Burr XII with covariates and two-segmented time-shifted model because of:

1. **In-Sample Fit:** The model performs the best in terms of MAPE and MdAPE. The fit is verified by the cumulative and incremental tracking plots (See Validation Section).
2. **Out-Of-Sample Fit:** N/A
3. **Robustness:** The order of the rankings (by BIC and MAPE) of models are the same. I also took subsets of the data and tried to estimate the model to see how the parameters would change (Appendix C). [8]
4. **Story Telling:** This model incorporates in all elements of the story and is compelling to managers because of the covariates and latent classes included.
5. **Parsimony:** This model is not the simplest. However, no simpler model can perform as well (measured by BIC).

The parameters estimated were:

| | **Segment 1** | **Segment 2** |
|---|---|---|
| **r** | 1.199 | 0.356 |
| **α** | 10606.819 | 4670.795 |
| **c** | 1.050 | 1.993 |
| **π** | 0.778 | |

[8] This method was used by Professor Fader in the CBCV paper.

| | | |
|---|---|---|
| **Google Trend – Social Distancing** | -0.004 | 0.182 |
| **U.S. Cases** | 0.001 | -0.030 |
| **European Travel Ban** | -0.061 | |
| **gamma (parameter for European Covariate)** | 0.000001 | |
| **delta (parameter for European Covariate)** | 0.000001 | |
| **S&P 500** | 0.057 | |

*Figure 9 Final Model Parameter Estimates*

### a. Interpretations from Final Model

#### Heterogeneity

Segment one is homogeneous (r > 1) whereas, segment two is more heterogeneous (r < 1). This contrasts with my hypothesis because I thought that both segments would be homogeneous. One reason may be that segment two is the coronavirus segment. Besides the coronavirus, consumers within this segment may differ widely by their grocery shopping habits, willingness to download apps, and location where they live.

#### Duration Dependence

Segment one has little duration; segment two has positive duration dependence. Again, this disconfirms my hypothesis. If segment two is the coronavirus segment, people need to eventually buy groceries and with stay-at-home orders, people would prefer to use Instacart. In contrast, segment one does not necessarily face this same constraint because they would have downloaded regardless.

### Covariates

- Segment one seems to be the non-coronavirus segment. This segment's coefficients for U.S. cases and "Social Distancing" are practically 0 – meaning these covariates were not very correlated with app downloads.
- Segment two is the coronavirus segment. While the coefficient for "Social Distancing" Google Trends is positive, the estimate for U.S. cases is slightly negative. This is puzzling to me because it seems as if both covariates should be positively correlated with one another. However, I will take this result with a grain of salt because there is a lot of multicollinearity for the U.S. cases variable. This multicollinearity may be diminishing the correlation between U.S. cases and other variables. Additionally, the "Social Distancing" covariate seems to matter the most (has highest magnitude).
- The European travel ban variable has a negative coefficient. This does not match my hypothesis. However, I think that this may be due to the high proportion ($\pi$) of segment one consumers. Since I used a pooled coefficient, segment one's lack of reaction to coronavirus may have skewed the coefficient towards zero.
- The S&P 500 is slightly positively correlated – as we would expect. However, I refrain from drawing too many conclusions because the variable was not very significant.

## b. Validation

## 1. Incremental Tracking Plot:

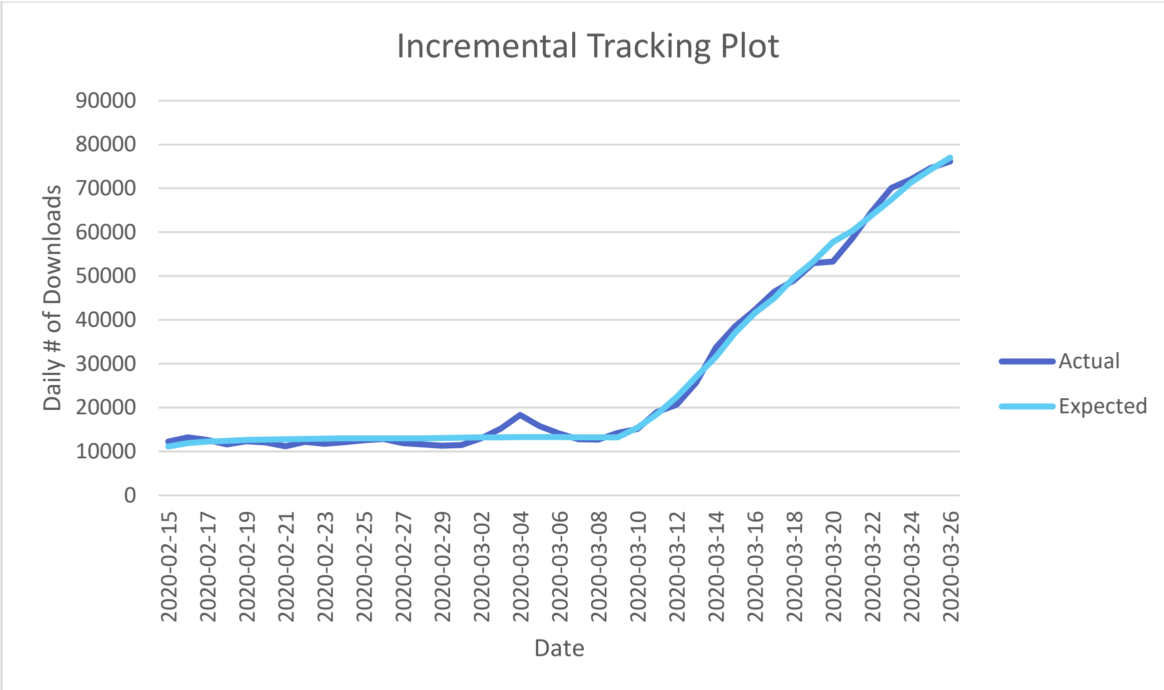



*Figure 10 Incremental Tracking Plot*

## 2. Cumulative Tracking Plot:

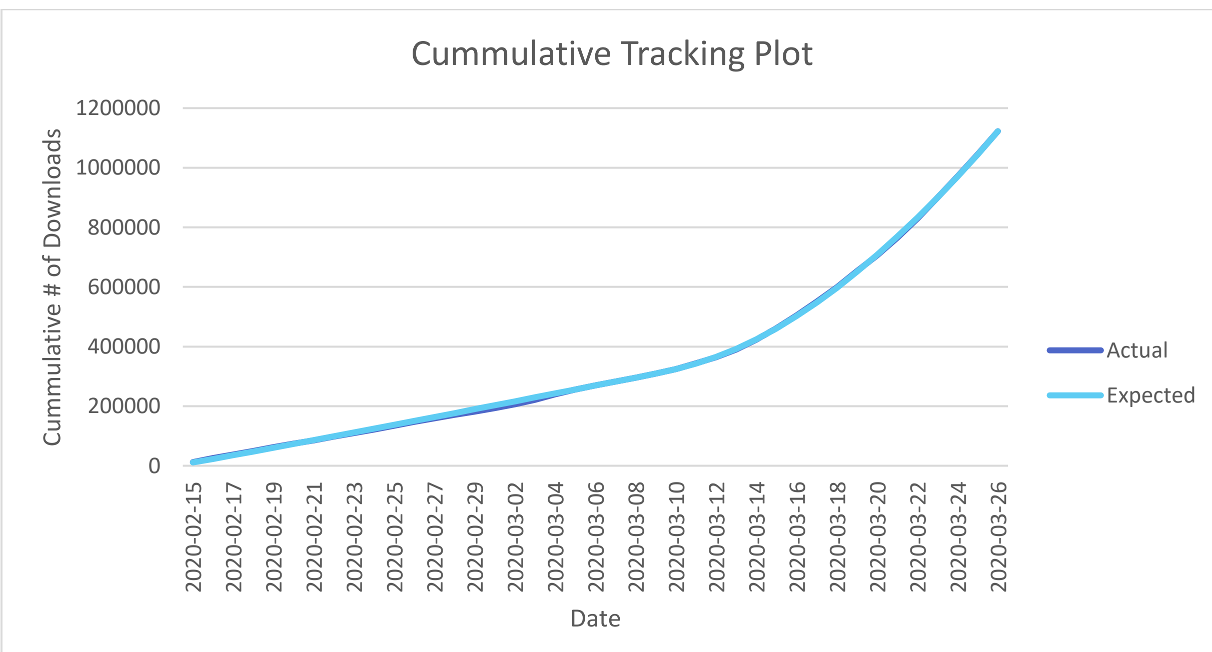



*Figure 11 Cumulative Tracking Plot*

The incremental tracking plot tells us more because the cumulative tracking plot can hide bad performance between periods. The incremental tracking plot is day-by-day, so any bad performance will be more easily visible. The plot indicates that the model works pretty well.

Since this is a single-timing event (there are no repeat downloads), we cannot come up with conditional expectations or in-sample histograms.

### c. Limitations and Future Areas of Improvement

The model relies on a lot of time-variant covariates. This poses a problem for forecasting future values because we would have to build a separate model for what we believe the covariates will be in the future. This makes forecasts more uncertain.

The model also does not account for one-off spikes in the data (March $4^{th}$). Perhaps if the data had more spikes (so that we can come up with a distribution to determine when the spikes would occur and how big) or if we knew of an internal reason (promotion), we could adjust for the spikes. Doing so would allow managers to better allocate resources (server management and customer service) to meet demand.

Finally, there is some multicollinearity among the covariates in the data. Therefore, the model has a harder time parsing out the individual effects of the covariates. We should be cautious about which covariates are actually correlated with downloads.

## V. Answering Managerial Questions

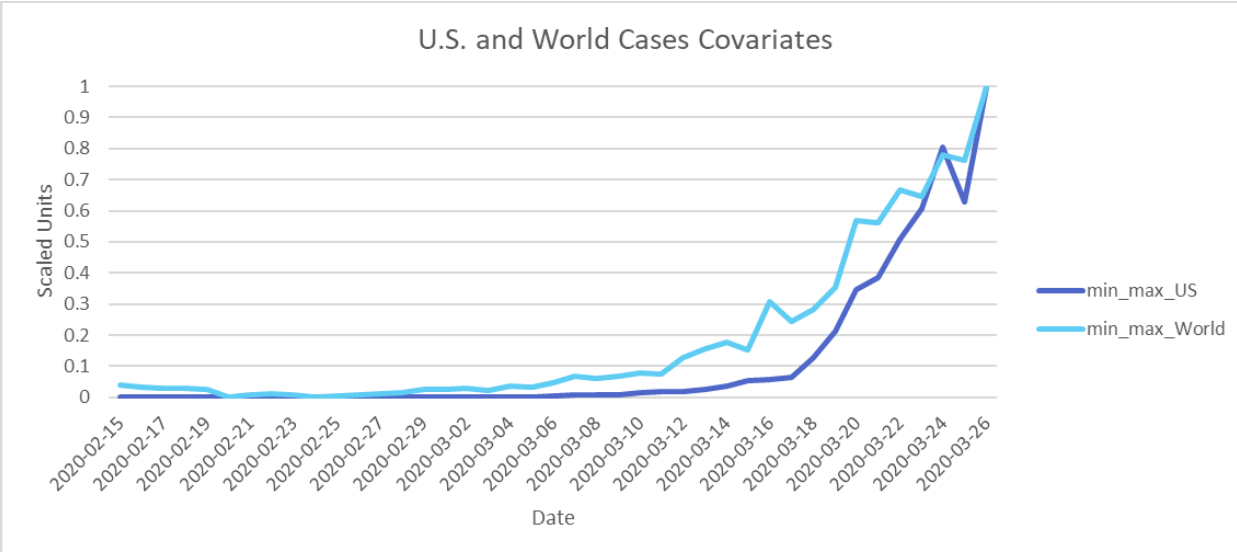
1. **Is there a new segment of consumers who have downloaded since the epidemic started?**

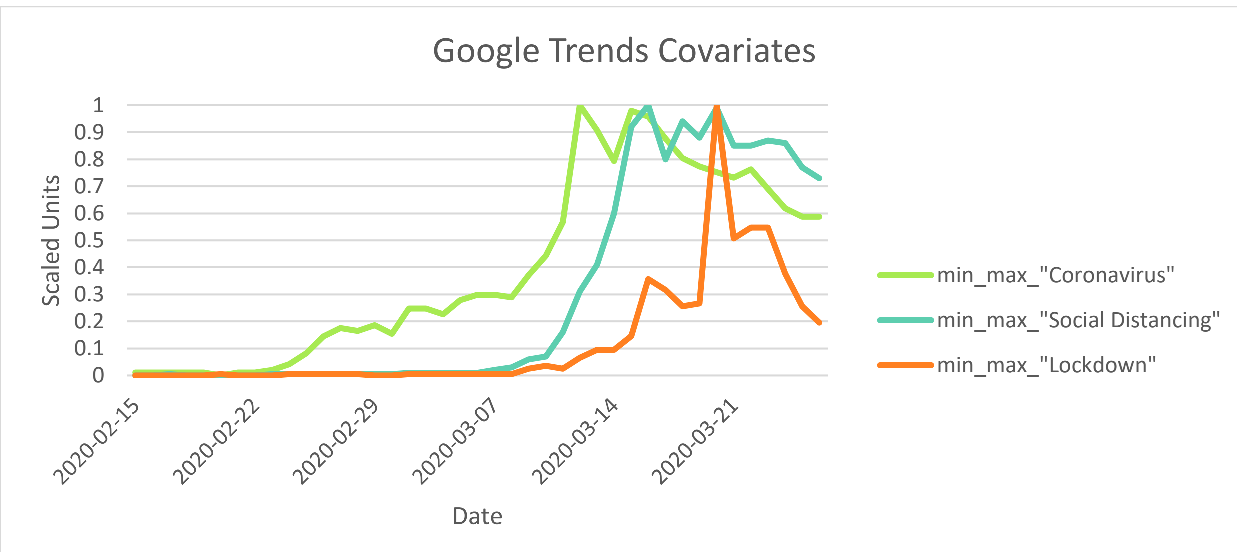
- Yes, the model indicates that there is a new segment of customers that have downloaded around the same time as the epidemic. However, this model constitutes a smaller proportion of overall downloads.

2. **How alike or different are consumers (heterogeneity)?**
    - The segment of customers who would've downloaded regardless of the coronavirus is more homogeneous. These consumers might be early adopters who were interested in Instacart as a niche product.
    - The segment of customers who downloaded during the same time as the epidemic are more heterogeneous. In contrast to the other segment, these people might represent a wider variety (different geographic areas, access to smartphones, incomes, etc.) and they might have only downloaded because it was the only alternative to going grocery shopping in-person.
3. **What are some factors that are correlated with number of downloads?**
    - For the coronavirus segment, coronavirus factors are generally positively correlated with number of downloads. These factors include "Social Distancing" Google Trends.

If Instacart could expand upon the model with additional user-level time-invariant covariates, the model could be used to predict which users are most likely to download. This could help better target advertising and calculate the lifetime value of consumers.

## Appendix A: Covariates Exploratory Data Analysis




**Summary Statistics:**

<u>For cases covariates:</u>

Mean, median, and standard deviation are not applicable because these covariates are cumulative numbers over time.

For Google Trends covariates:

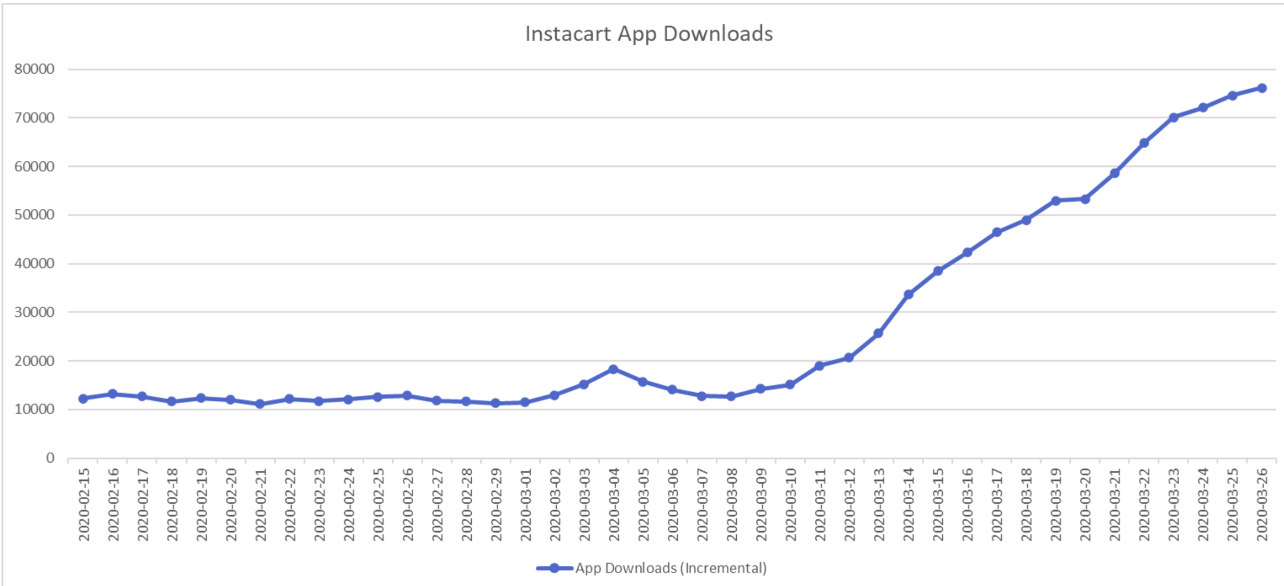
The summary statistics for these covariates are not included because they do not make sense to have summary statistics. These covariates are already scaled so the min and max would align with the scale. Additionally, many unknown values are set at 0, so taking the mean, median, and standard deviation would not give us a good idea of how spread out the data is.

For European Ban Covariate:

This covariate was a constructed covariate representing the decay of the announcement of European travel restrictions. Therefore, it does not have summary statistics.

| | **U.S. Cases** | **World Cases** | **States of Emergency** | **S&P 500** | **Oil Prices (USD/bbl)** |
|---|---|---|---|---|---|
| **Min** | 0 | 527 | 0 | 2237.40 | 20.37 |
| **Max** | 13963 | 50868 | 50 | 3386.15 | 53.78 |
| **Mean** | N/A | N/A | N/A | 2881.79 | 38.14 |
| **Median** | N/A | N/A | N/A | 2954.22 | 36.21 |
| **Standard Deviation** | N/A | N/A | N/A | 369.23 | 11.74 |

# Appendix B: Plot of Incremental Downloads



## Appendix C: Checking Robustness by Estimating on Subsets

Estimated only on days 1 – 21:

| | Segment 1 | Segment 2 |
|---|---|---|
| **r** | 382.359 | 0.980 |
| **α** | 250580.917 | 3000.000 |
| **c** | 0.996 | 2.000 |
| **π** | 0.046 | |
| **Google Trend – Social Distancing** | 26.640 | -3289.035 |
| **U.S. Cases** | 39.866 | 0.000 |
| **European Travel Ban** | 0.000 | |
| **gamma (parameter for European Covariate)** | 0.1 | |
| **delta (parameter for European Covariate)** | 0.1 | |
| **S&P 500** | 0.368 | |

Estimated only on days 22 – 41:

| | Segment 1 | Segment 2 |
|---|---|---|
| **r** | 0.664 | 243.405 |
| **α** | 2270954114.781 | 1482375.341 |
| **c** | 5.684 | 1.836 |
| **π** | 0.020 | |

| | | |
|---|---|---|
| **Google Trend – Social Distancing** | 0.206 | -18.659 |
| **U.S. Cases** | 0.001 | 10.009 |
| **European Travel Ban** | 0.207 | |
| **gamma (parameter for European Covariate)** | 0.000001 | |
| **delta (parameter for European Covariate)** | 0.000001 | |
| **S&P 500** | -0.344 | |