Numerator

Are COVID Shopping Behaviors "For Real?"

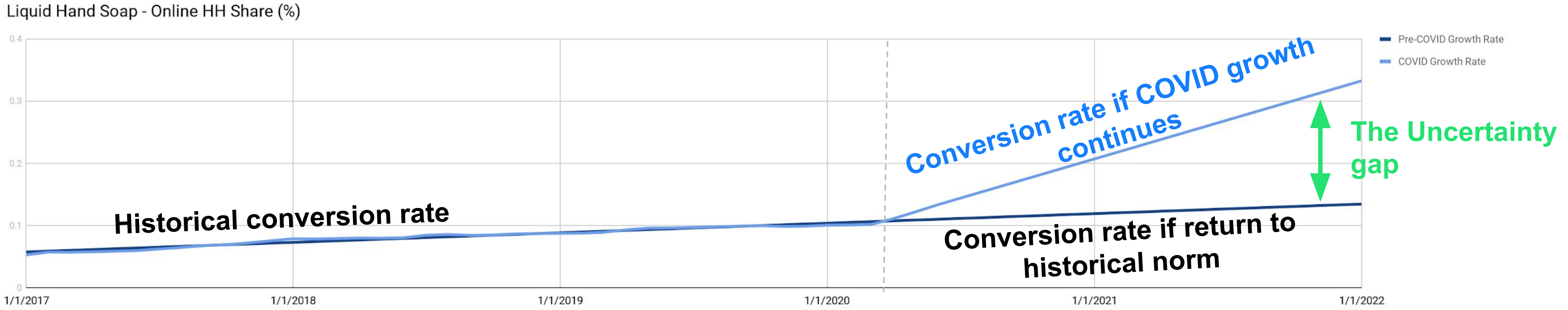
Improving manufacturer demand forecasts with a flexible modeling approach, including COVID parameters

James Fontaine, Jason Goldrosen, JB Helton

Background: Virus has upended consumer behavior

“We are about to be put into a position that is unprecedented and something that we will have to **prepare for and react very quickly** to” - Amit Singh, CEO Chewy.com

- Shoppers converting their shopping to purchasing online has accelerated quickly due to the emphasis on contactless shopping and retail closures - this has caused havoc on most forecasting models
- Retailers and Manufacturers are increasingly concerned about how “permanent” this new behavior is or if consumer behavior will return back to historical norms. Previous work only focused on % of Households impacted



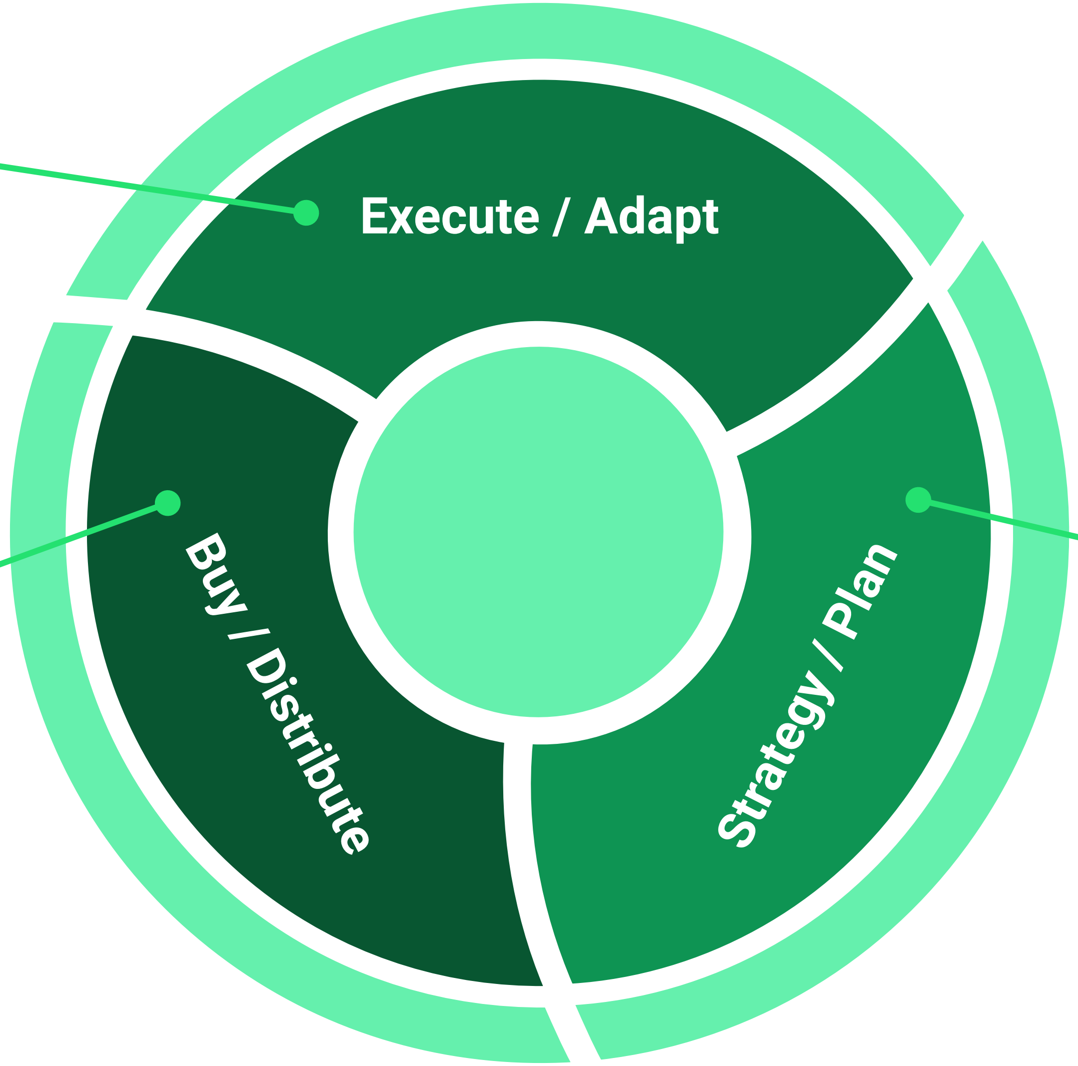
Source: Numerator White Paper: The Rate of eCommerce Acceleration
https://www.numerator.com/resources/white-paper/rate-ecommerce-acceleration-why-it-matters-and-what-next

Numerator

Problem: Will the shift to Online stick and become the “new normal”?

Monitoring performance vs historical norms is a challenge and no benchmarks exist

Buyers are increasingly cautious about early ordering for distribution & promotions as inventory build starts months prior



COVID has broken the traditional cycle between Manufacturers and Retailers

Retailers and Manufacturers are trying to plan 6-12 months with uncertainty about assortment and online / offline channels

Numerator’s unique view into Omnichannel behaviors and deep consumer knowledge positions us well to develop a predictive consumer model that can take both historical as well as future factors into account

Impact

- Demand forecasting is an extremely important discipline for most Manufacturers. Disruptions to supply / demand can result in major impacts to P&L **[Clear Value]**
- COVID, or even other more gradual consumer behavior shifts, are cases where modeling can help manufacturers plan for uncertainty **[Clear Need]**
- With this model, Manufacturers can have a set of models that predict baseline consumer behavior changes & introduce additional models for unknown events to improve their current demand forecasts **[Flexible to additional use cases]**

Financial Impact

- Calculated 3-year NPV: $4.5M
  - High: $11.2M - Low: $2.3M
- Numerator's Current Business
  - ARR: ~$50M (Consumer Panel Product only)
  - Avg Contract Value: $300k
- Assumptions:
  - $30,000/annual fee for packaged analysis for total market with 1 category breakout
  - Years 2 and 3 transition into other consumer behavior impact (econometric modeling)
  - # of clients (category) volume assumptions in Year 1:
    - Churn prevention - 3
    - New Logos - 5
    - Cross Sells / Upsells - 20

Numerator

Approach: Bring together multiple sources in a modular way to maximize predictivity

Numerator Omnipanel Purchase Data

- Largest consumer purchase panel in America with over 450,000 households
- Analysis revolved around 30,000 US households with 28 months of continuous data

+

State-Level Distancing Policies Data

- Tracks social & business restrictions including: stay-at-home orders, gathering sizes, school/bus. closures, eat-out restrictions
- Broader Latent variables created
- Developed and maintained by researchers at University of Washington

+

State-Level COVID-19 Case/Death Count Data

- Daily new case counts and deaths. Tracking started with the first reported US COVID-19
- Used as inputs to develop a case and death score
- Developed and maintained by The NY Times

Data Sources combined to build a predictive model of COVID Behavior “Stickiness”

Various transformations and time lags of each variable were considered and final model was determined through art and science, considering ease of use, predictive power, and statistical integrity

Numerator

Approach: Two Stage Modeling Approach

Historical Prediction

Purchase behavior predicted by separately modeling the distinct behaviors that contribute to overall purchase behavior

1. % of Panelists who make purchase
2. Expected number of trips among purchasers
3. Expected Spent per Trip among purchasers

- Models individual-level behavior that can be aggregate up to market level
- Modeling approach could be used to power market simulations of purchase behavior

COVID Adjustment

Multivariate linear regression for weekly panelist level spend prediction as a factor of Covid concern (daily input by panelist state)

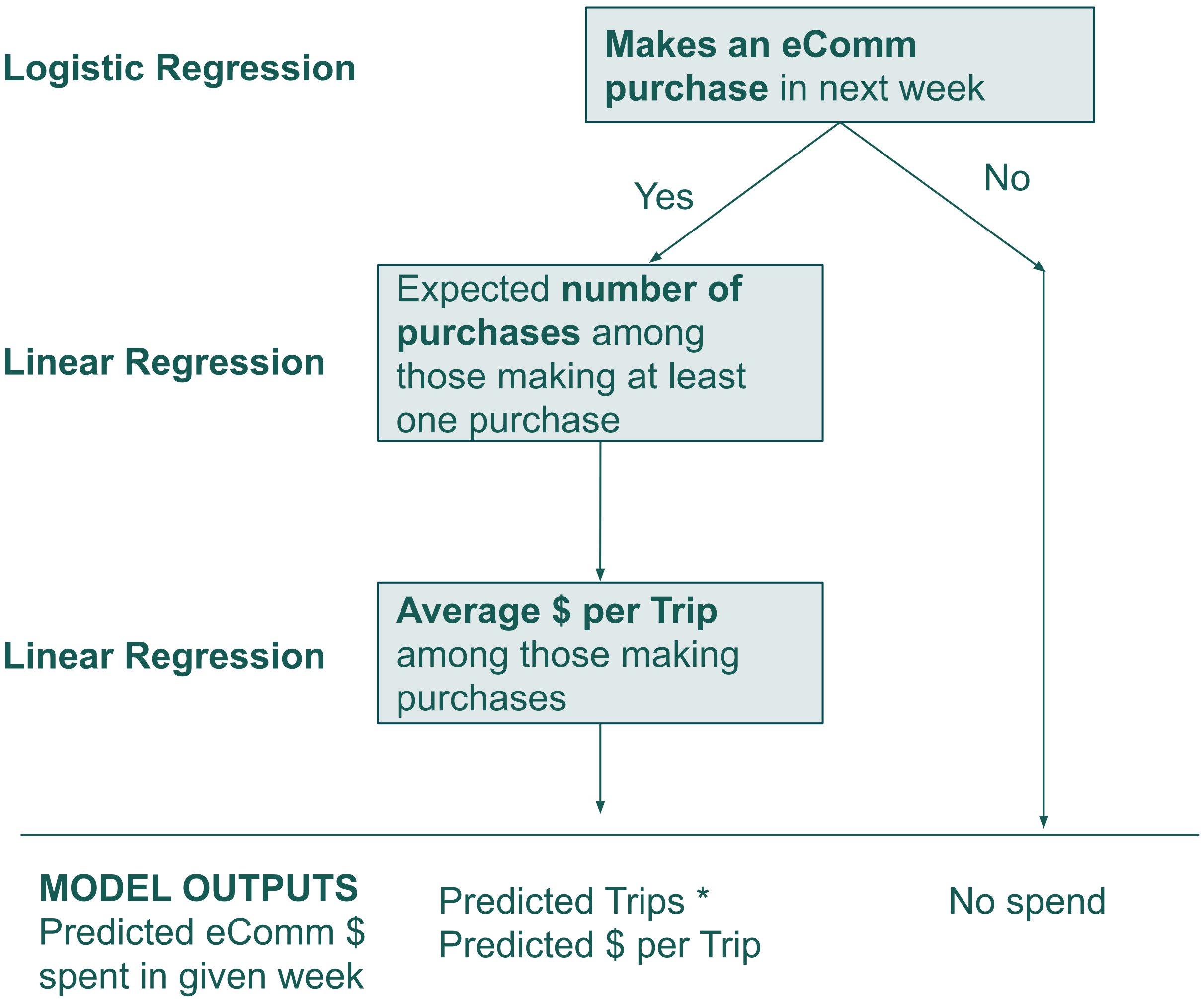
1. **COVID-19 Case Count and Death Score**: Adjusted to reflect impact of large increases and 1st time occurrences
2. **Gathering Score**: Magnitude of restrictions on social gatherings including the size of the group restriction, mandatory or recommended, and if restriction has been eased
3. **Closure Score**: Measures the amount of closures across locations accounting for mandatory / recommended & easing

- Isolate COVID level of influence at the individual level that can be aggregate up to market level
- Forecast spend based on anticipated level of COVID concern

Historic Model Details

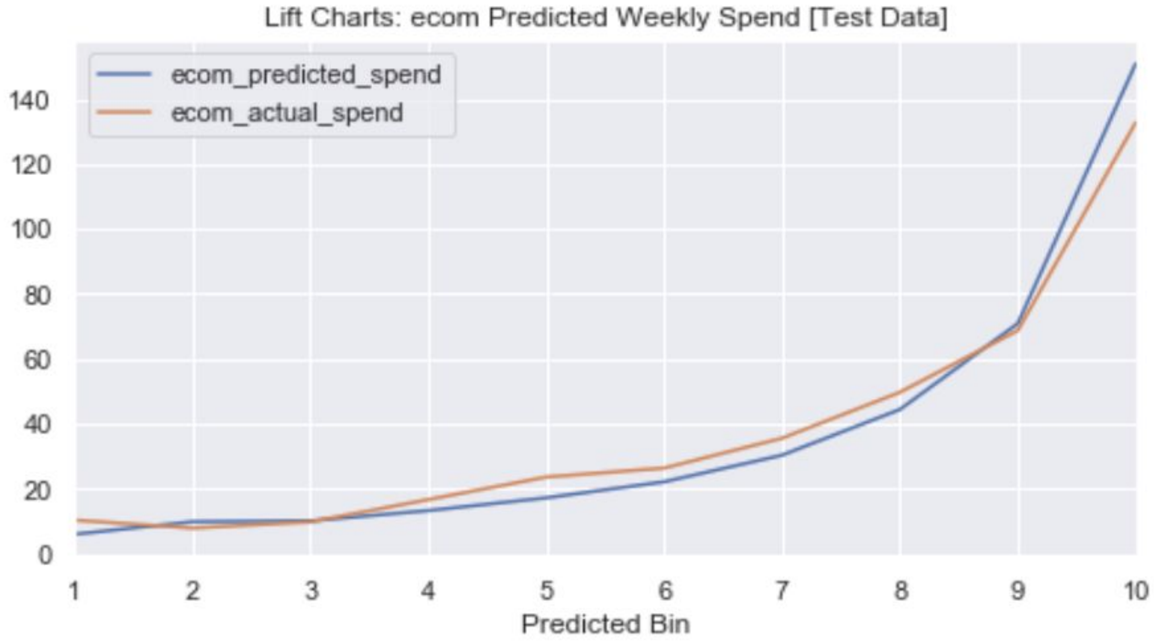
- Household purchase behavior can be mathematically decomposed into three key behaviors
  - Percent of households who make a purchase
  - Number of purchase among those who purchase
  - Spend per Trip among those who purchase
- We took approach of modeling these behaviors individually
- Inputs
  - Recent purchase behavior in last 4 weeks
  - Seasonal effects to control for Prime Day, Holidays
  - Time trend to account for gradual rise in eCommerce behavior behavior over time
- Outputs
  - Predicted eCommerce and FMCG spend for each panelist on weekly basis

Modeling Framework



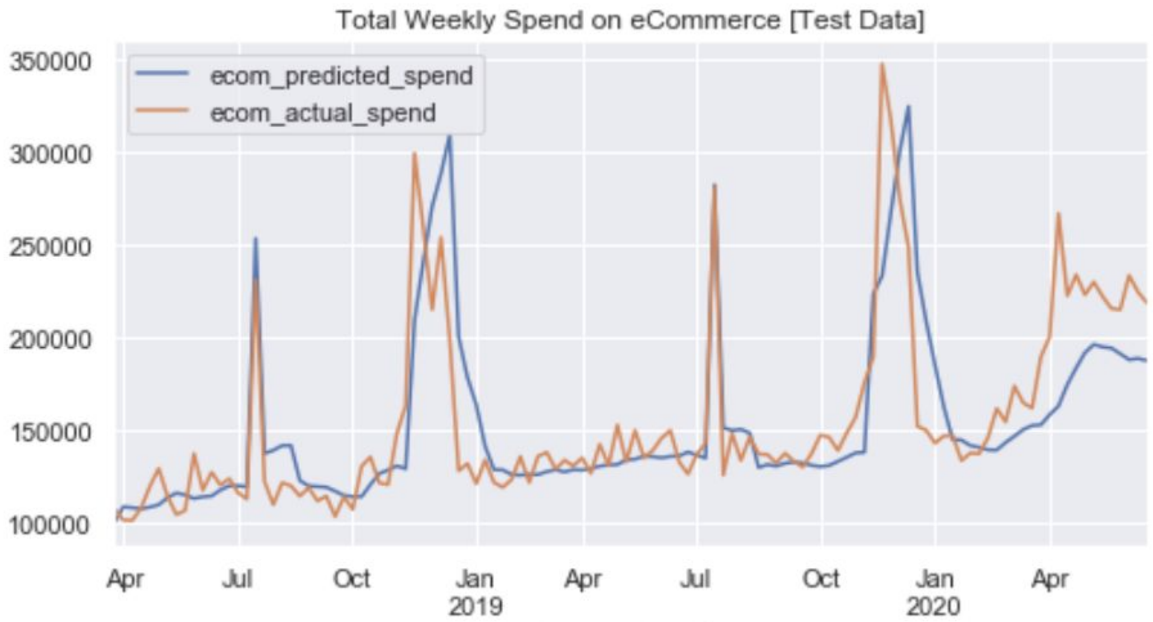
Historic Model Details

- Despite using simple individual models, the framework yields a strong predictions at both a panelist level and an aggregate market level



Strong Foundation w/ Potential to Improve

- Can be extended to any type of purchase behavior (Category, Brand, Store)
- Could be improved with more feature engineering or incorporating more sophisticated ML models
- Could be used to power market simulations of purchase behavior



Models trained on 66% of sample on dates between Mar 2018-Dec 2019. Only holdout panelist data is shown in the charts above.

COVID Model Details - Concern Score Breakdown

Presence of state level gathering restrictions/stay at home orders on a given day weighted by magnitude of the restriction as well as whether the policy was a mandate and/or eased at some point

Business Closures

School, Restaurant, & Bar Closures

Presence of state level closures on a given day (business closures and combination of school/rest/bar weighted equally)

Gathering Score

Closure Score

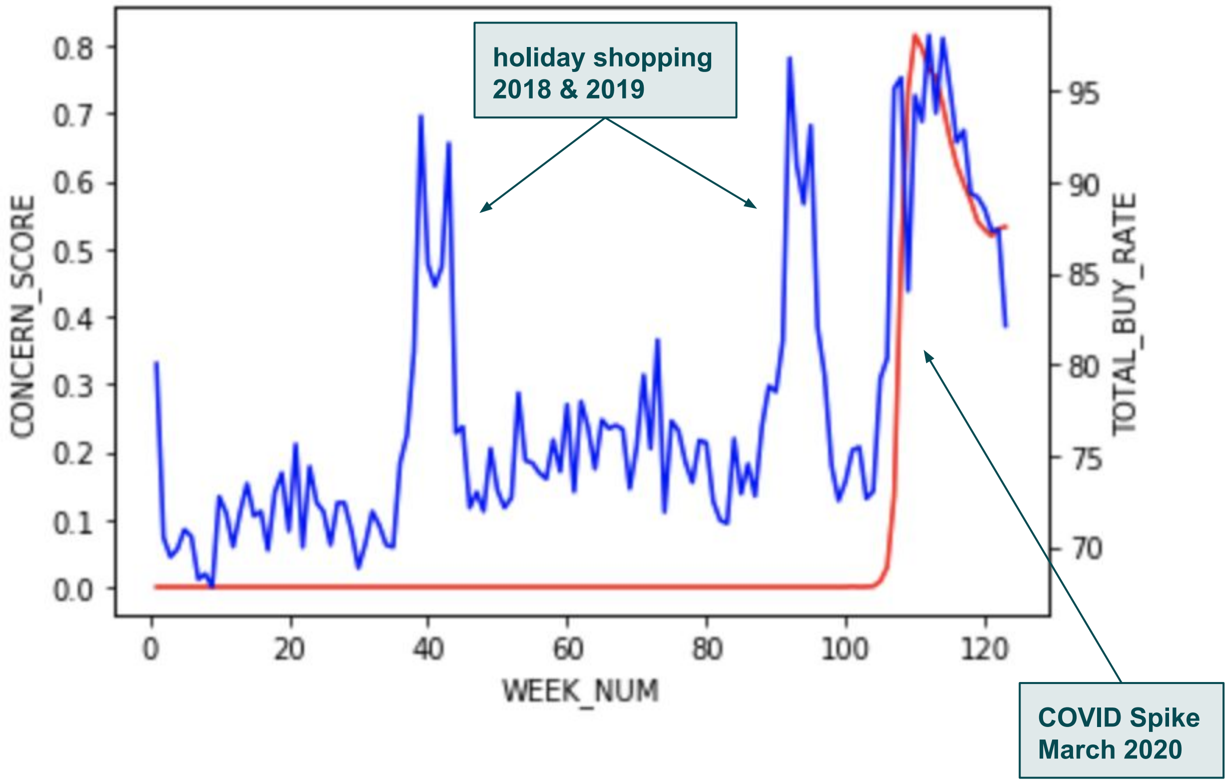
Concern Score

Case Score

Death Score

New cases/deaths by state by day calculated by taking the cumulative z score of new cases/deaths on the given day. Mean/standard deviation evolves over time to allow for a fatigue effect - modeling the fact that cases/deaths were more of a shock to HHs early on in the pandemic

Correlation with Buy Rate: Apr 2018 - June 2020 Data



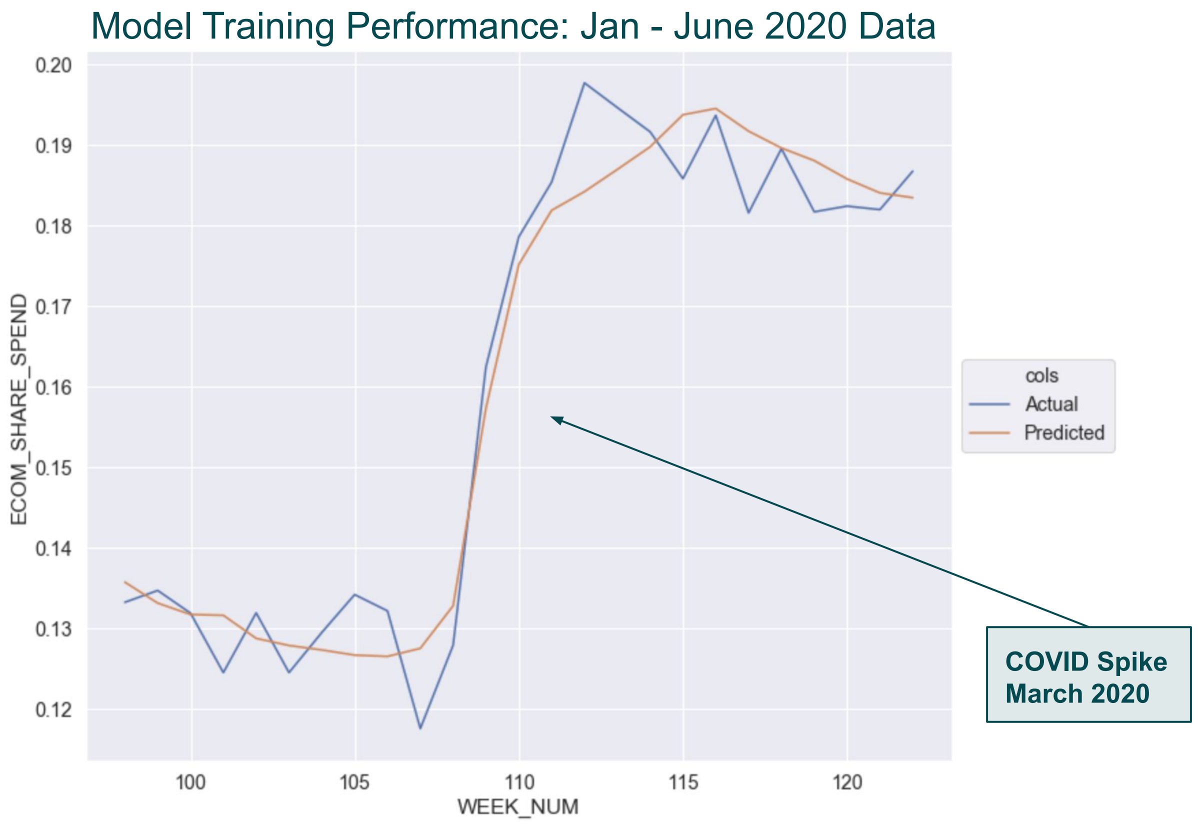
Concern Score is highly correlated with spike in HH Total Spend/Ecom Share at a National level during COVID

Note: only state-wide policies were considered for gathering/closure scores

Numerator

COVID Model Details

- Multivariate linear regression model
- Inputs:
  - COVID concern score (daily input)
    - input both as current value and 1 week lag to account for some delayed behavioral response to COVID
  - spend/share simulated predictions from Historic model, weekly and 4 week rolling (weekly input)
  - demographics/psychographics (HH constant)
  - 2019 historic spend data (HH constant)
- Outputs (weekly):
  - Total Spend Prediction
  - Ecommerce Share of Spend Prediction



Model predicts panelist spend/ecom share as a function of simulated spend (from historic model) and anticipated COVID concern. Flexible enough to accept concern level by day and by state. We can then compare these predicted metrics to initial historical simulated predictions. Thus we calculate an index of COVID's anticipated influence during the predicted time period at the individual HH level.

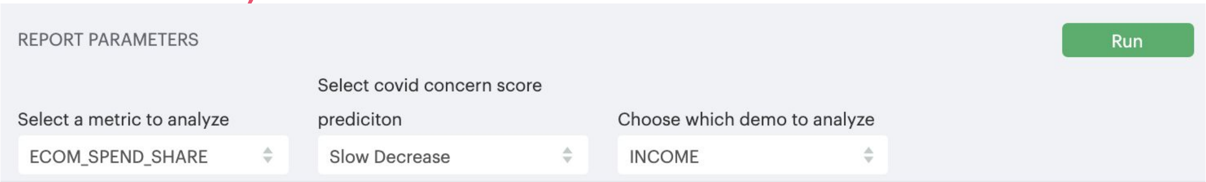
Demos include but not limited to age, gender, ethnicity, income, has children, hh size, urbanicity
Self-reported psychographics include but not limited to work from home, dining out intensity, cat/dog ownership, financial situation last year/this year

Covid Shopping Behavior Tool

General Functionality

1.) Client interactively sets parameters

- **Duration and severity of the pandemic**
- **Purchase metric (% of Spend, Buy Rate)**
- **Demographic split (Income, Urbanicity, Census Region, etc.)**

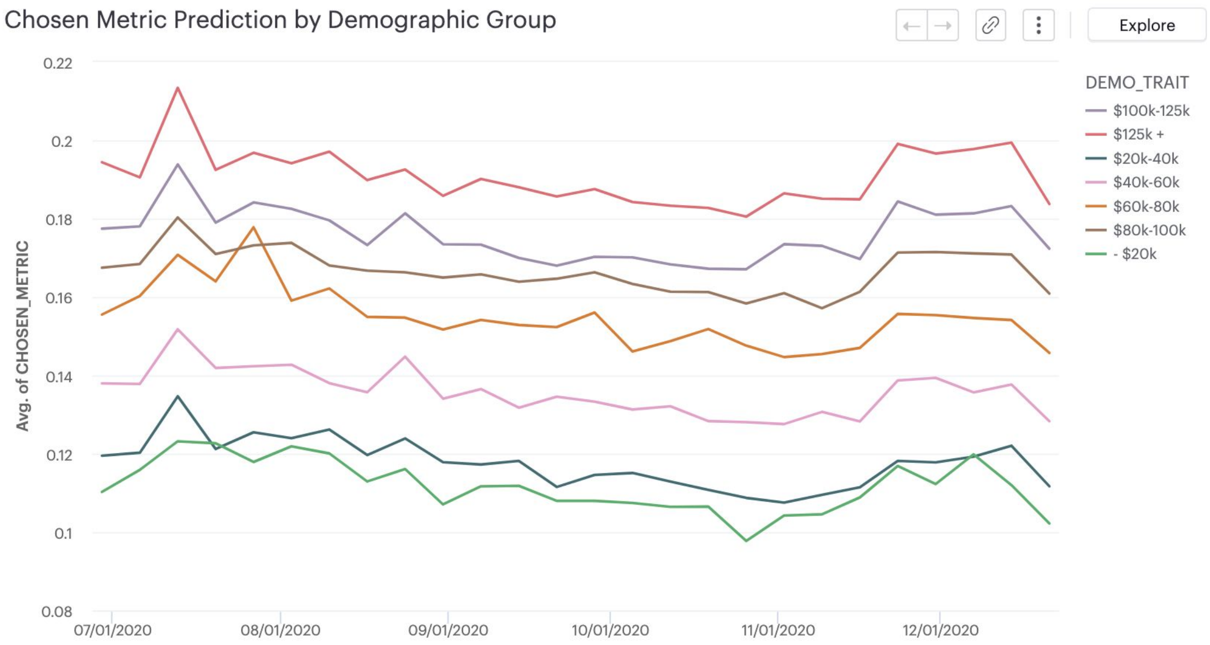


2.) The tool then displays our forecasts for purchase behavior, helping client answer key business questions

- How long do we expect the pandemic to affect purchase behavior in my category?
- Which demographic groups are most affected by the pandemic and how long do we expect that to continue?

Example

- Under a **"Slow Decrease"** simulation (a prolonged but gradually improving Covid-19 situation), we expect **% of eCommerce Spend** to a gradually decrease for all **Income groups** but remain elevated above historical norms through end of year
- By the beginning of Nov, we expect this % of eComm Spend to range from 19% for highest income groups to 11% for lowest income groups
- We expect a spike in Nov & Dec due to Holiday shopping



Conclusion - Thank You!

Current State

- We can predict future behavior in times of crisis using behavior & sales models of the past and present
  - Human behavior is predictable if you can find a pattern and set of replicable variables
  - Modeling Historical Predictions using Crisis Adjustments (based on factors like Level of Concern)
  - Numerator’s ability to quickly survey a broad demographic allows for agile predictions of future behavior, informing strategy and distribution decisions
  - Can immediately start to market our tool for ad hoc “what if” scenarios as well as high level thought leadership

Future State

- Next Steps
  - Adding Retailer and Category Views for deeper granularity - significantly increases client impact
  - Expanding concern score predictions - utilize existing case/death predictions from other sources to hone in on the most accurate possible forecast
  - Continually update model as we move through the pandemic - validate initial predictions and adjust for continual improvement