

USING PRIVATE DATA TO UNDERSTAND PUBLIC BEHAVIORS

PRIVATE SECTOR DATA CAN DRASTICALLY
IMPROVE SITUATIONAL AWARENESS AND
DECISION MAKING IN A CRISIS



Erik Wetter



Sara Rosengren



Huong Nguyen

This is a preprint from the book "Sweden Through the Crisis", to be published
in the fall by SIR, Stockholm School of Economics Institute for Research.

The novel COVID-19 virus (SARS-CoV-2) and the associated disease (COVID-19) represent one of the biggest global public health crises in a century. Most countries have responded with drastic measures, and by the end of March more than half the world's population were under lockdown or government-imposed mobility restrictions. These decisions are not taken lightly and the economic and societal ramifications in the short and long run remain uncertain.

A key challenge in the current crisis has been the extended time period between infection and disease. According to the WHO, the time between exposure to COVID-19 and the moment when symptoms start is commonly around five to six days but can range from 1 – 14 days. This time-gap makes it hard for governments to assess how spread out the disease is, but also to evaluate different measures taken to prevent it.

This has been especially evident in Sweden, a country which unlike most other EU states did not order a lockdown, instead following a soft approach, issuing recommendations and calling for citizens to 'take responsibility' and to follow government guidelines. While global policies and interventions differ, most policymakers struggle with a lack of timely indicators, specifically with regards to public responses and behaviors.

In an ongoing project we combine data and insights from private sector partners to provide new understanding of public behavioral dynamics during this crisis. In doing so, we highlight the value that private companies can provide in terms of high-resolution insights into public behaviors and responses to government guidelines during crises. Specifically, for infectious diseases such as COVID-19, we can see that private sector data can provide timely and disaggregated insights on different segments of the public, specifically those age groups designated as high-risk and thus considered more vulnerable.

The potential of private sector data

Private sector data sources are growing exponentially with regards

to collection and size and has radically transformed how economic behavior can be measured. There is ample research showing how private data such as online search behavior, media analytics, and official statistics can offer high-resolution, high frequency and/or real-time measurement of economic activity.

The use of company data (for example from mobile operators) has also been found to provide rapid and unique insights to support crisis response after natural disasters such as the earthquakes in Haiti 2010 and Nepal in 2015. Similar data has also been shown to provide more precise models and estimates for combating the spread of infectious diseases, both in terms of modeling and predicting the spread as well as understanding the effectiveness of and population adherence to mobility restrictions.

In the Spring of 2020, we have seen several examples of private companies sharing data to provide insights on public reactions to the crisis. For example, Google has with their Community Mobility Reports contributed a comprehensive cross-country comparison of population mobility, with locations broken down by retail & recreation, grocery & pharmacy, parks, transit stations, workplaces and office locations, and residential locations. In fact, the industry response to the COVID-19 crisis has been overwhelming with numerous initiatives, with over 250 public-private data collaboratives worldwide to improve data collection and private sector insight sharing to support the response¹.

In our project we hope to complement these initiatives by

- Combining various data sources
- Explaining the data sources and methods.
- Producing in-depth research insights.

First, it is evident that deeper insights can be gained by combining datasets and insights from multiple sources. By being able to access and analyze data from several companies as well as relevant public and open data sources, we hope that this project will generate novel insights that go beyond what any one of the data sources could provide.

Second, although the value of high-resolution company data is intuitively compelling, there are a lot of misconceptions and flawed assumptions about the nature of the data and methods, even among policy and research professionals. In providing explanations of the data sources and methods aimed at practitioners, we hope to support more informed debate and use of non-traditional data sources for public good. Many of the data sources in this project are already available for private or commercial use, but most of the analytics and use cases are for operational purposes and decision support. In setting up an ongoing research platform and collaborations, our purpose is to be able to pursue longer term research questions and produce robust research findings on the value that private sector data can provide in understanding and potentially predicting population behaviors and behavioral dynamic in a crisis for the benefit of society as well as for the project partners.

Third, by collecting a range of different datasets on public and economic reactions to the crisis, we hope to build a resource that can be used to create a better understanding of public behaviors, such as different timelines and population segments during a crisis. The project will evolve as we continue to combine datasets, timeframes, and insights. While numerous open data initiatives already exist, some of our project partners already conduct and disseminate their own analytics.

What is required to make it work?

While private sector data has the potential to provide new indicators and analyses to drastically improve public situational awareness and decision making in light of the crisis, it is also surrounded by numerous challenges to data sharing: privacy concerns, intellectual property issues, and commercial sensitivities, to name a few.

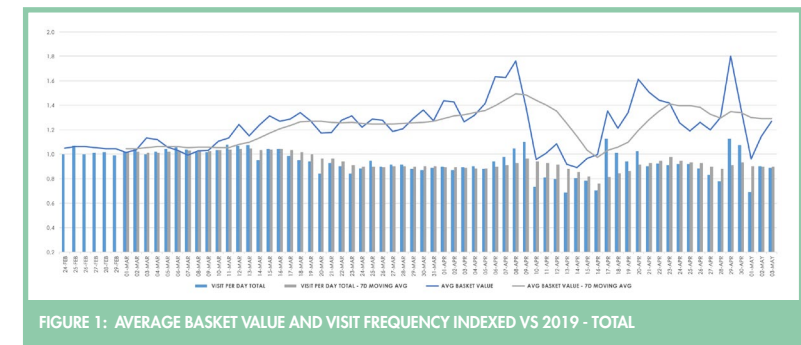
Academia can fill a crucial role in the ecosystem by functioning as a qualified and neutral third-party partner that can manage to link different datasets from several companies, increasing the overall value without putting privacy and competitive concerns in jeopardy.

The case of grocery retail data

As an illustration of how private sector data can be used, let us consider grocery retail data.

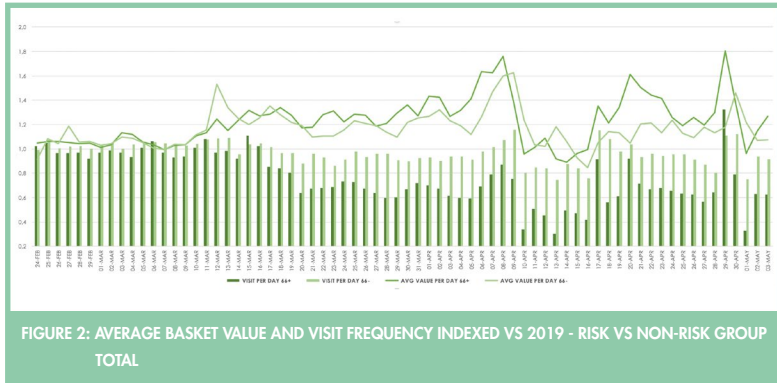
After about two months of the pandemic, data from grocery retail suggests that Swedes are seemingly entering the phase of “the next normal”. To compare with “the previous normal”, we created two indexes based on grocery sales data: number of visits (the columns) and average amount spent on each visit (the lines) where 2020 figures were compared to a benchmark of 2019 using same weeks and same weekday. As a result, value over 1 indicates an increase and value under 1 indicates a decrease versus the same period last year.

Figure 1 reflects a dramatic change in grocery shopping behavior starting mid-March, when the Public Health Agency of Sweden (PHAS) raised the risk of domestic spread to very high. Observably, consumers went to the grocery stores less often at this point, but when they did, they tended to buy more, indicating bulk-buying behavior. Using the store visits as an indicator, it can be seen that generally, Swedes were following the social distancing recommendation from PHAS.

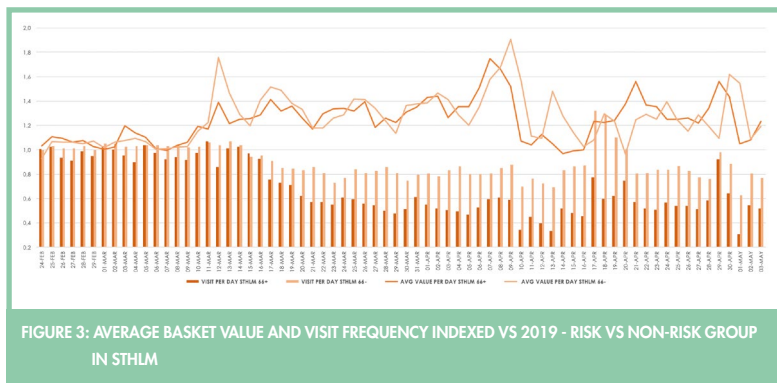


Shortly after the PHAS announcement (March 10, 2020), the non-risk group (people under 66 years old) promptly reacts. However, in the long run, the members of the risk group are seemingly the ones who drive the change. As seen in Figure 2, the risk-group went grocery shop-

ping significantly less often (number of visits decreased by 20% in late March and by 37% in April) and when they did so, they generally bought more (average basket value increased by 25% in late March and by 31% in April).



The bulk-buying pattern is more visibly observable in Stockholm, the area most impacted by COVID-19. Compared to 2019, the number of store visits by Stockholm consumers as a whole decreased by 14% and 16% in March and April respectively. Taking a closer look, the risk-group decreased by 30% and 45% and the non-risk group decreased by 12% and 13%. Interestingly, the average basket value of the risk and non-risk group in Stockholm both increased by approximately 30-35% in late March and April 2020, which may imply a higher level of stocking up among the non-risk group.



These analyses are purely descriptive and can be regarded as a first glance in terms of what is possible from using private data in a crisis. As such, they illustrate the value of private data for understanding public behaviors and dynamics in a crisis context. Whereas mobile data has previously been used to study mobility in a large number of urban and crisis settings, the use of grocery retail data for these purposes is more novel. This indicates how grocery retailers can serve as information resources to understand how people are reacting to information and guidelines from authorities. Retail data has several advantages as it is typically large in volume and variety, but also has high velocity. Sales data is continuously updated and tracked by retailers and as such it could be a valuable source of information for understanding public reactions in a time of crisis.

THE AUTHORS

Erik Wetter is Assistant Professor at the Department of Entrepreneurship, Innovation, and Technology at Stockholm School of Economics.

Huong Nguyen is a PhD candidate at the Center for Data Analytics and the Center for Retailing at Stockholm School of Economics.

Sara Rosengren is the Association of ICA Retailers' Professor in Business Administration, especially Retailing at the Center for Retailing at Stockholm School of Economics.

REFERENCE

E. Wetter, S. Rosengren, and F. Törn (2020). Private Sector Data for Understanding Public Behaviors in Crisis: The Case of COVID-19 in Sweden, SSE Working Paper Series in Business Administration, Stockholm School of Economics, No. 2020:1.