SSE Working Paper Series in Business Administration
No. 2020:1

Private Sector Data for Understanding Public Behaviors in Crisis: The Case of COVID-19 in Sweden

Erik Wetter [a, b]
Sara Rosengren [a]
Fredrik Törn [c]

[a] Stockholm School of Economics

[b] Swedish Defence University

[c] Coop Sverige

SSE Working Paper Series in Business Administration
No 2020:1
April, 2020

Private Sector Data for Understanding Public Behaviors in Crisis – The Case of COVID-19 in Sweden

Version 1.1
April 9, 2020

Erik Wetter[1,2], Sara Rosengren[1], Fredrik Törn[3]
[1]Stockholm School of Economics, Sweden [2]Swedish Defence University, Sweden [3]Coop Sverige, Sweden

Abstract

The novel Coronavirus (SARS-CoV-2) and the associated Coronavirus disease (COVID-19) has in early 2020 rapidly spread to become one of the biggest global public health crises in a century, with global economic impacts and supply chain shocks never before seen in modern history. Most countries have responded with drastic measures, and at the time of writing this 3.9 billion people – half the world´s population – are under lockdown or government-imposed mobility restrictions.

Sweden can be seen as a case of special interest as unlike most other EU countries, Sweden has not ordered 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.

Here we describe a new project in which we combine data and insights from private sector partners in retail and telecom to provide new insights in public behavioral dynamics with a specific focus on mobility, consumption, and hoarding behaviors e.g. bulk buying. 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 crisis. 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 such age groups designated as high-risk and thus considered more vulnerable.

This working paper will be continuously updated as new insights are produced in order to provide relevant insights that can hopefully assist in supporting more facts-based decision making for the public good.

Acknowledgements: We would like to thank Kristofer Ågren, Marisa Leysen-Jestin, Patryk Larek, and Linus Brännström for their active support for this project.

Introduction

Private sector data sources are growing exponentially with regards to collection and size and has radically transformed how economic behavior can be measured. When combined with other private and public data sources such as online search behavior, media analytics, or official statistics such data has been shown to offer high-resolution and close to real time measurement of economic activity (Cavallo & Rigobon, 2016; Einav & Levin, 2014).

Company data on consumption and mobility, for example from mobile operators, has been proven to provide rapid and unique insights to support crisis response after natural disasters such as the earthquakes in Haiti 2010 and Nepal in 2015 (Bengtsson et al., 2011; Wilson et al, 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 modelling and predicting the spread (Wesolowski et al. 2014; Bengtsson et al. 2015) as well as understanding the effectiveness of and population adherence to mobility restrictions (Peak et al., 2018).

The industry response to the COVID-19 crisis has been overwhelming with numerous initiatives, and over 150 public-private data collaboratives worldwide to improve data collection and private sector insight sharing to support the response[1]. Large tech companies like Google has with their Community Mobility Reports[2] provided 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.

The focus of this project is to create a better understanding of public behaviors in different segments of the population during a crisis, by combining and analyzing multiple private and publicly available datasets on media, consumption, and mobility. The project will evolve as we continue to combine datasets, timeframes, and insights. While numerous open data initiatives already exist, and some of our project partners already conduct and disseminate their own analytics, we see this project providing additional value in the following ways:

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

The value of high-resolution company data is intuitive; however, it is evident that the real insights come from 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.

While most people easily grasp the concept of analyzing consumption and mobility patterns from retail and mobile data, there exist 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.

[1] Data4Covid; Data Collaboratives for COVID-19 Response, NYU GovLab Living Library accessed April 8

[2] https://www.google.com/covid19/mobility/

Timeline of COVID-19 in Sweden

On January 16, the Public Health Agency of Sweden (PHAS) puts out their first communication about the novel Coronavirus outbreak in China. At the time, the virus was considered low risk for Sweden[3]. On January 30 the World Health Organization (WHO) declares 2019-nCoV as a Public Health Emergency of International Concern (PHEIC)[4]. On January 31st, the first COVID-19 case, a person who had returned from Wuhan, China, is confirmed in Sweden[5], and the following day on February 1st, the PHAS classifies COVID-19 as a public health concern[6].

Looking at the information conveyed to the public about the corona virus the first wave of news from the PHAS is released January 30 and 31 when it is suggested that the new Corona virus is classified as a public health concern. The second wave is released February 24-26 and focuses on the international spread of the virus and the risk that it reaches Sweden through people who have travelled in affected regions. From February 27-March 9 there are continuous updates mostly related to the risk of travel and international spread of the virus.

On March 10, however, PHAS announces that there are now signs of domestic spread of the new Coronavirus in Sweden, leading to the first domestic restrictions: prohibition of public gatherings with more than 500 people (March 11), changes in the process for testing and recommendation for anyone with symptoms to stay home (March 13), recommendations for people over 70 to avoid social contacts (March 16), distance teaching for students over 16 years and all higher education (March 17), recommendations to avoid travel (March 19), and new rules for restaurants and pubs (March 24). Many of the actions called upon in these announcements have been mentioned at earlier times. On April 1, PHAS decides to provide a more comprehensive overview based on social distancing and personal accountability.

Methods and data

This section describes the three data sources currently in use in the project; media analytics, retail data, and mobile data.

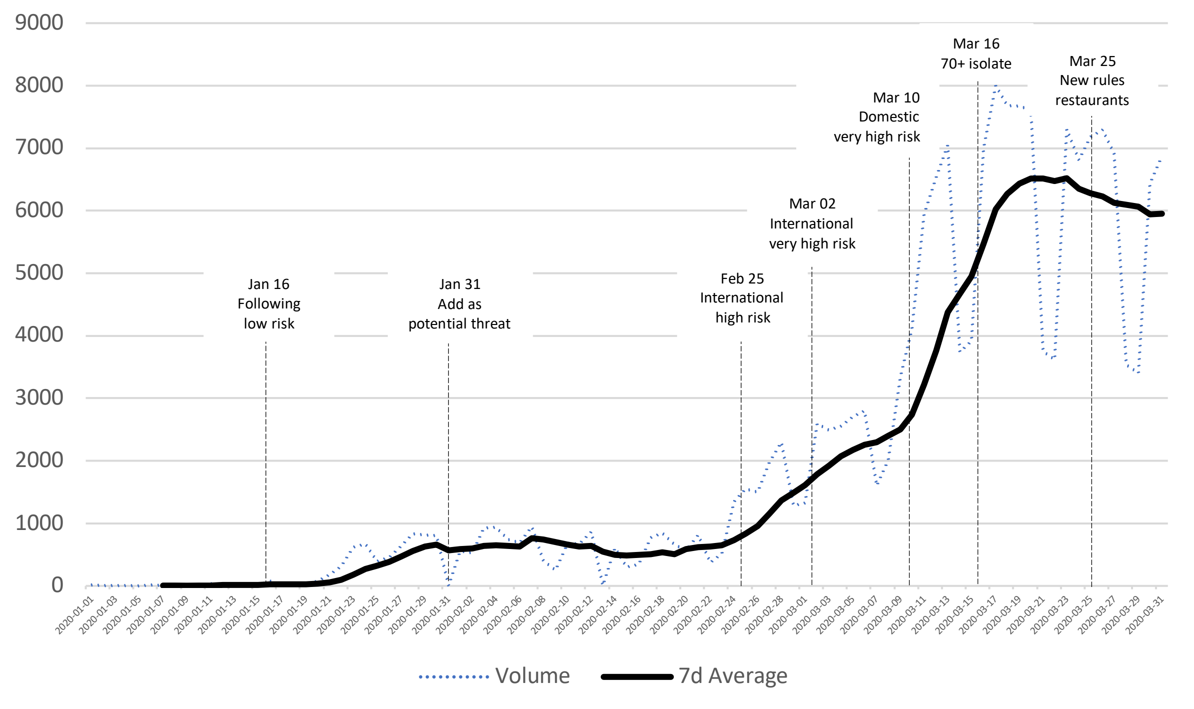
Media data from Meltwater

Editorial media is one of the main information sources of the COVID-19 crisis for the public; both through communicating government guidelines and recommendations, and also through own editorial coverage and interviews.

Meltwater is a Norwegian media monitoring and business intelligence service provider that was founded in 2011 and today has 55 offices worldwide. One of their key offerings is collecting global online media in a structured database that monitors over 300,000 news outlets globally and that allows for monitoring and searches on topics and events.

In collaboration with Meltwater the project will analyze the patterns of media communication around the COVID-19 crisis and how this affects population behaviors. An illustrative graph is found below in Figure 2. This is a simple volume graph that indicates how media interest and exposure has evolved during Q1 2020 and is based on Boolean search query of all Swedish editorial articles containing the term ‘Corona-’ or ‘Covid-’.

Figure 1. Sweden COVID-19 media mentions Q1 2020



From the media data we can see indications that editorial media interest in COVID-19 was initially quite low, and then increased slightly with the announcement that COVID-19 could be a potential risk on January 31st. Once it is flagged as an international high risk on February 25 media interest increases significantly and a sharp increase is noted after March 10 when the Public Health Agency announced a very high risk of community spread in Sweden. The announcement that people above 70 years of age should follow social distancing rules also received media coverage, and after that the media volume appeared to have peaked at a high level.

Retail data from Coop Sweden

Retail companies have a long history of working systematically with data to guide decision making (Bradlow et al 2017, Germann et al 2014; Wedel and Kannan 2016). Loyalty programs are the most common way for tracking customer behaviors in this industry. Data is typically stored in customer relationship management-systems (CRM-systems) comprising various data such as demographic data coupled with purchase behaviors.

In the current project we received data from Swedish grocery retailer Coop Sweden on food sales during Jan-March 2020 and Jan-March 2019. Coop Sweden is a retail chain with 650 store locations throughout Sweden that are owned by 3.5 million members through KF (Swedish: Kooperativa Förbundet, "Swedish Co-operative Union"), a federation of consumer co-operatives that was founded in 1899.

The data available by most contemporary retailers comprises rich information in a five-dimensional space: across customers, products, time, geo-spatial location, and channel (Bradlow et al 2017). Whereas there is a growing literature on how combining data across these five dimensions might benefit retailers, this data is, to the best of the authors knowledge, yet to be used in a public health crisis context. Below follows a brief description of the five dimensions and how they can be used in this context:

- *Customers*: CRM-systems allow retailers to move from aggregate data analyses (overall sales) to individual level data analysis (buying behaviors of specific individuals). In retailing the ability to track customers and link transactions over time is key. In a public health crisis, this granular level of data allows for analyses of overall responses as well as responses in specific groups. Below we move from an overall analysis of aggregated sales to a comparison of sales between risk groups (age 66+) and other groups (18-65).

- *Products*: Information about product are typically stored in several different systems based on a SKU-identifier (Stock Keeping Unit). Most retailers have access to product information in two-dimensions. First, information about the actual products available in the store. Second, information about the properties of each product that is available. In a public health crisis, this data would allow for analyses of overall responses in terms of increasing/decreasing demand of certain product categories, but potentially also provide opportunities to link content of products with certain outcomes.

- *Time*: Any transaction (either on the customer or product level) comes with a time stamp of date and time. This means that retail data allows for a continuous measurement of, for example, customer behaviors and product stock-outs. In a public health crisis, this data would allow for analyses of public behaviors during different days and time of the day, as well as help understand product availability over time (such as a week or month).

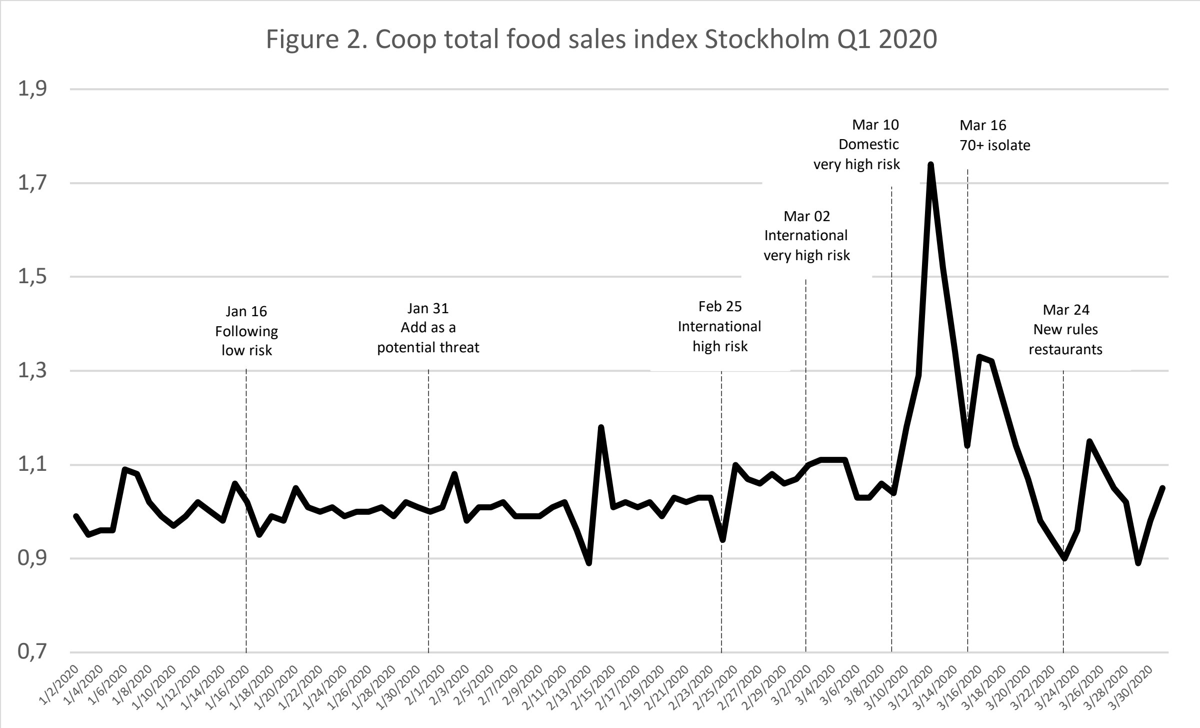
- *Location*: Retail data provides spatial information of customers in terms of the location of physical stores where transactions occur. Coupling this information with customer, product, and time, retailers have the potential to hyper-target their offers. In a public health crisis, it allows for analyses of public behavior in terms of mobility and changing habits. Potentially it could also help identify opportunities in which information would be more likely to have the desired effect.

- *Channel*: Retail data also provides opportunities to link online and offline behaviors. For retailers this helps understanding how customer move between channels and how this movement impact retailer profitability. In a public health crisis, this data allows for analyses of public behavior in terms of mobility and changing habits.

In order to track any changes in food sales connected to the COVID-19 crisis without divulging any commercially sensitive information, we developed a time series index that compares sales numbers per day in Q1 2020 compared with sales numbers for the corresponding weekday in Q1 2019.

Figure 2. below illustrates the volume of total sales for Stockholm in Q1 2020. We can see that food sales follow an expected pattern within normal variations until the March 10 announcement of very high risk of domestic community spread in Sweden, which caused sales to peak at +74% over expected index in the following days, indicating bulk buying behaviors. After a peak day, the sales volumes drop sharply almost back to the index level, but are then boosted again to 32% above expected index level with a high likelihood driven by the announcement that people above age 70 and in the high risk group from serious COVID-19 complications are instructed to isolate and apply strict social distancing.

After this sales volumes drop to -10% below the index, until they rise 20% in what could be a payday effect and also in response to the announcement that restaurants are to apply social distancing rules for patrons. However these variations are also within the normal range of deviations, so causality needs to be further tested before any conclusions can be drawn.



Mobile data from Telia Sweden

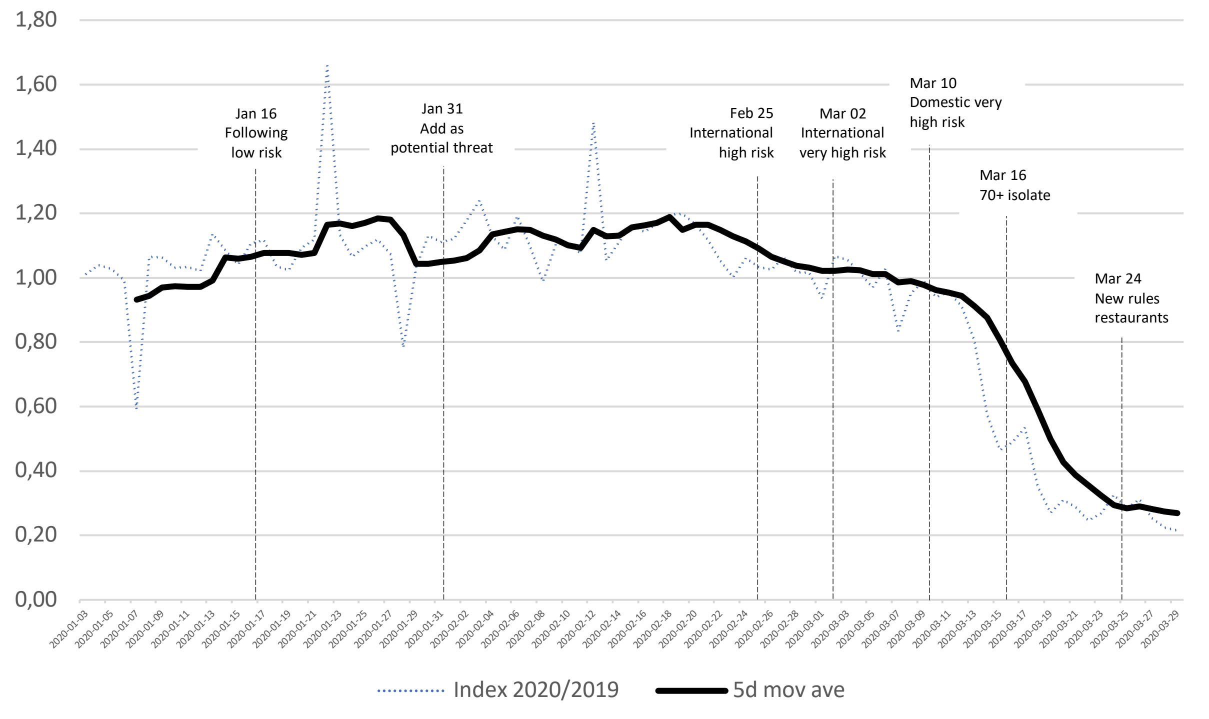
All mobile operators routinely collect and store global industry standard various network data that are used for billing and network optimization purposes. These data contain geographical information in that it stores the position of nearest cell tower to which each mobile handset is connected to at any given time. While this is a rough geospatial indicator, the mobile phone or SIM-card can be anywhere within the nearest cell tower coverage area – this geographical resolution has been shown to be enough for modelling population mobility for a number of commercial and research applications.

Telia is the largest mobile operator in Sweden. One of their commercial data offerings is called Telia Crowd Insights. Telia Crowd Insights is a service that analyzes anonymized and aggregated mobile network data from the Telia network. This provides a way to understand grouped movement behavior in society, such as travel patterns. Only grouped movement patterns are used, and the data is irreversibly anonymized. This means individuals cannot be identified as all personal information is removed and it is fully GDPR compliant.

One of the analytical products is an Activity Report that captures how many subscribers that have been spending a defined amount of time in a defined geographical area. These scalable geographical areas are aligned with the national statistics grids developed by Statistics Sweden, and allows for a grid resolution of 500x500m.

For an indicative graph, four grids that cover Stockholm Central Station and surrounding area were chosen as an indicator of Stockholm mobility. This location covers commuting through Stockholm Central (subways and trains), and the surroundings mostly contain offices, retail locations and hotels, but relatively few residential units.

Figure 3. Stockholm mobility index Q1 2020



From Figure 3. we can see that our Stockholm mobility index is hovering slightly above the expected range, which could be an expected development from the previous year given growth of population and thus the pool of potential mobile subscriptions. In the weeks after the February 25 announcement (International high risk) and March 02 announcement (International very high risk) we can see a slight downwards trend in mobility.

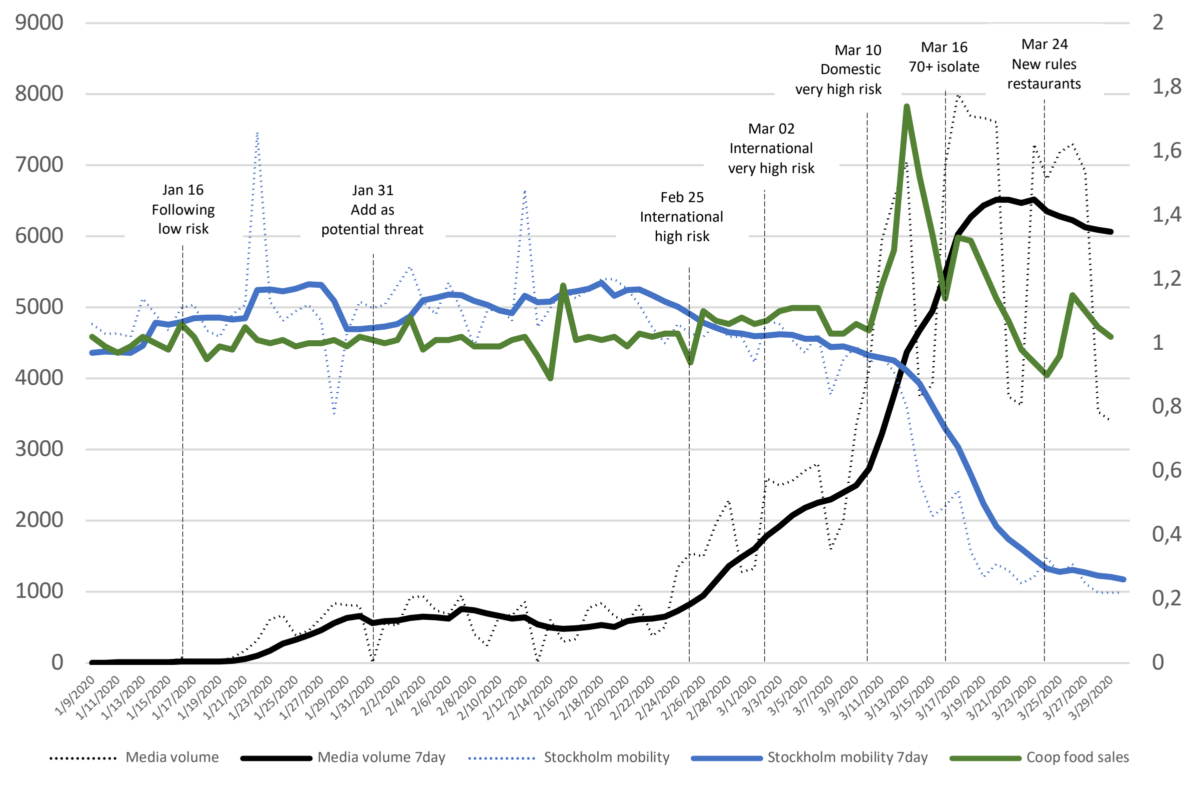
In the week after the March 10 announcement (Domestic very high risk) we can see that mobility decreases by -20% from the expected index, and in the week after the March 16 announcement, mobility decreases by another -40% below the expected index. After the March 24 announcement (new rules for restaurants), Stockholm mobility seems to have stabilized at ca -75% below the expected index, but the data time series is too short to make any assessment at this time.

A first look at a combined overview

Combining the above data sources with food sales data for Stockholm from Coop Sweden, Figure 4. below gives an indicative overview of how COVID-19 media mentions, and Stockholm mobility and food sales behaved during Q1 2020. Though this figure is descriptive only and no causal relationships have been quantitatively established or tested, there seems to be some indicative patterns and dynamics between the population behaviors as indicated by the data sources.

It almost appears as if there is an inverse relationship media mentions and mobility, although the direction of any potential causality is untested, and also whether it could be the quantity or content of media that would be driving population behaviors, if a causal link was established. It also seems like the bulk buying was a temporary activity.

Figure 4. COVID-19 media volume and population behaviors (Stockholm)



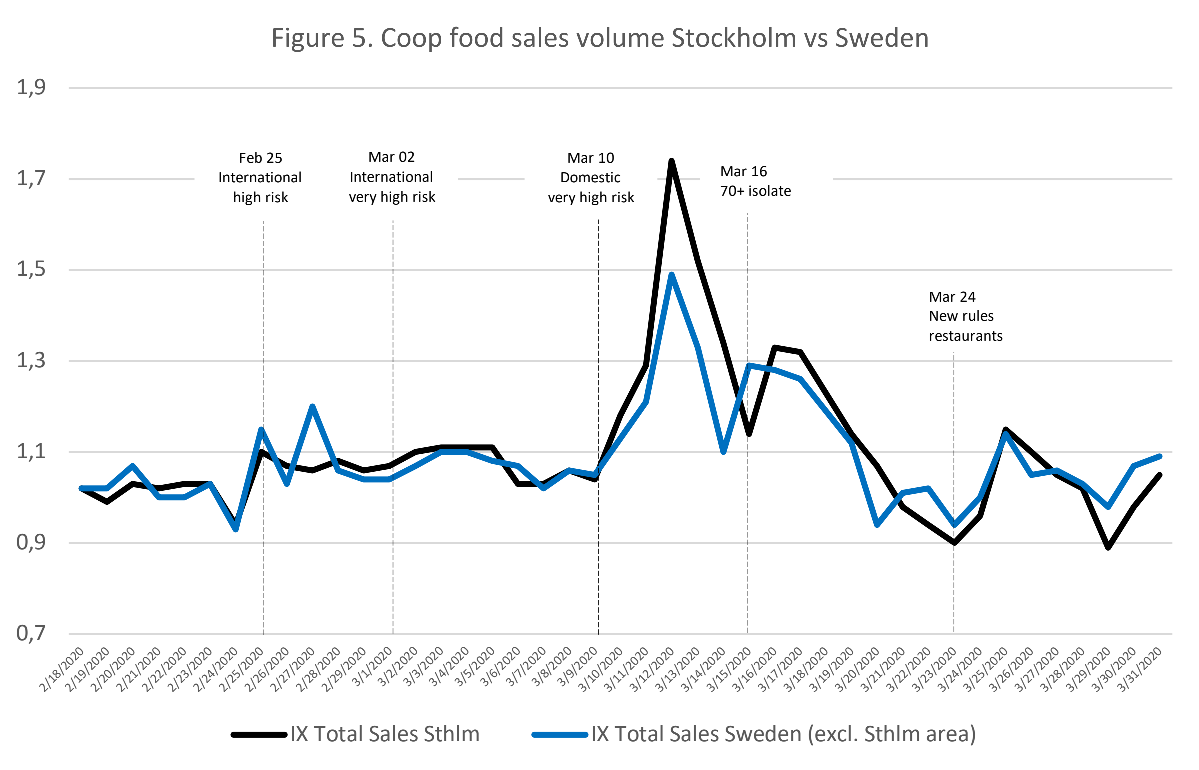
Breakdown of Stockholm grocery shopping behaviors

Figure 5. plots the food sales of Coop in Stockholm (black) and Sweden excluding the greater Stockholm area (blue) for January-March 2002. Sales are indexed using same week, same weekday sales from 2019 as a benchmark, meaning that a value over 1.0 is an increase and a value under 1.0 a decrease versus the day of the previous year. During January and February, we see some fluctuations in sales (mostly within the range of -10% to + 10%, with bigger fluctuations around the 23-25th every month as 25th is payday in Sweden, but not if on a weekend).

On March 10, PHAS raises the risk for domestic spread to very high and this sets off a big increase in total sales. On Thursday March 12, sales in Stockholm are up 74% and in Sweden outside Stockholm they are up 49%. In Sweden (non-Stockholm) there is another spike on Sunday March 15 (+29%) and in Stockholm there are two on March 16 (+33%) and 17 (+32%) respectively.

The data shows that consumers buy more food than expected during the period of March 10-March 19, and the pattern is stronger in Stockholm than in the rest of Sweden. From March 20 the situation stabilizes, although there is a 14% peak in both Stockholm and Sweden (non-Stockholm) on March 25 (payday). Also, note how stockouts on March 16 (Sweden) and March 18 (Stockholm) reduce consumer demand following bulk buying the days before.

On an aggregate level, the results do not show any clear evidence of bulk buying until March 10. Bulk buying then continues until March 19. Consumers in Stockholm engage in more extreme bulk buying than people outside of Stockholm. Based on the currently available data, Thursday March 12, two days after the PHAS announcement of domestic spread, is peak bulk buying.



Purchase patterns as a measure of social distancing

Although the total food sales provide an overview of how people behave on the aggregate, retailers typically break up aggregate sales into the number of visits (the amount of transactions made) and the average basket value (the average amount spend on each transaction). Doing so allows a better understanding of how these sales come about and thus allows more in-depth insights into consumer behaviors.

Figure 6. plots the total visits at Coop in Stockholm (black) and Sweden excluding greater Stockholm (blue) for January-March 2020. The number of visits is indexed using same week, same weekday sales from 2019 as a benchmark, meaning that a value over 1 is an increase and a value under 1 a decrease versus previous year.

For the most part of the time period we see some fluctuations in the number of visits (mostly within the range of -10% to + 10%), with a few spikes in mid-February. More notably we see a big drop in the number of visits from March 16 and onward. This drop is bigger for Stockholm than for the rest of Sweden. On March 23 the number of visits in Stockholm is down by 29%, but this is most likely a payday effect.

It is interesting to compare the use of store visits as a population mobility indicator in comparison with the mobile data; both can be used to indicate the number of people that are moving around the city, but where the mobile data gives a broader overview and can span large geographical areas and also capture commuting patterns, in the COVID-19 scenario the store visit data also contains a risk indicator; as a large number of people in a grocery store could indicate higher risk of viral transmission.

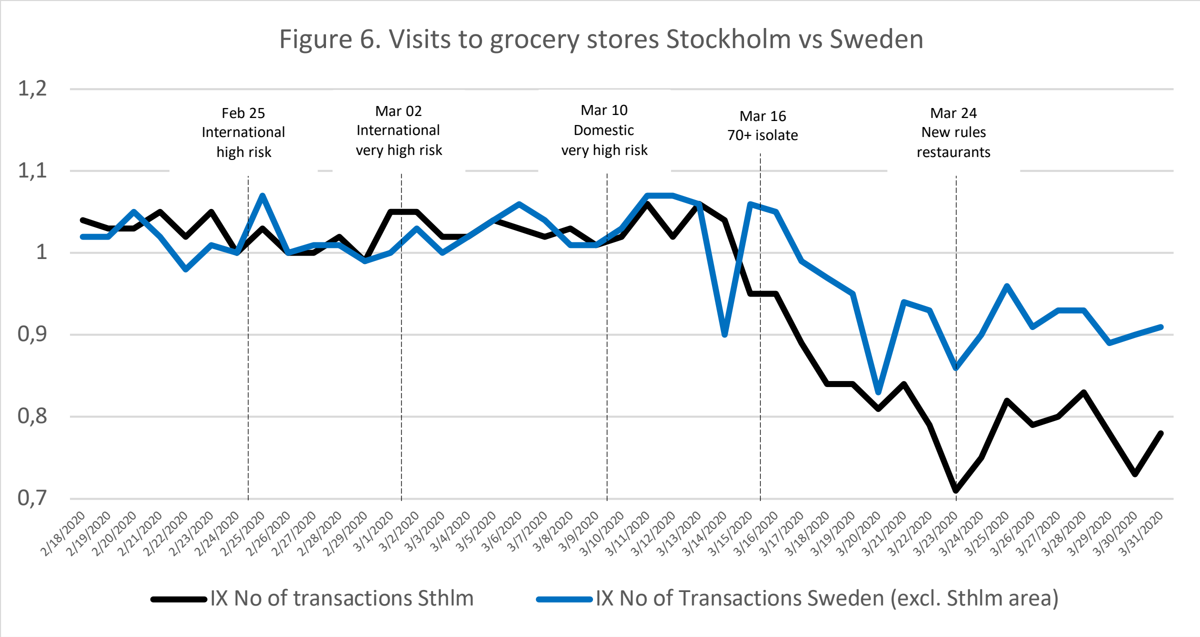
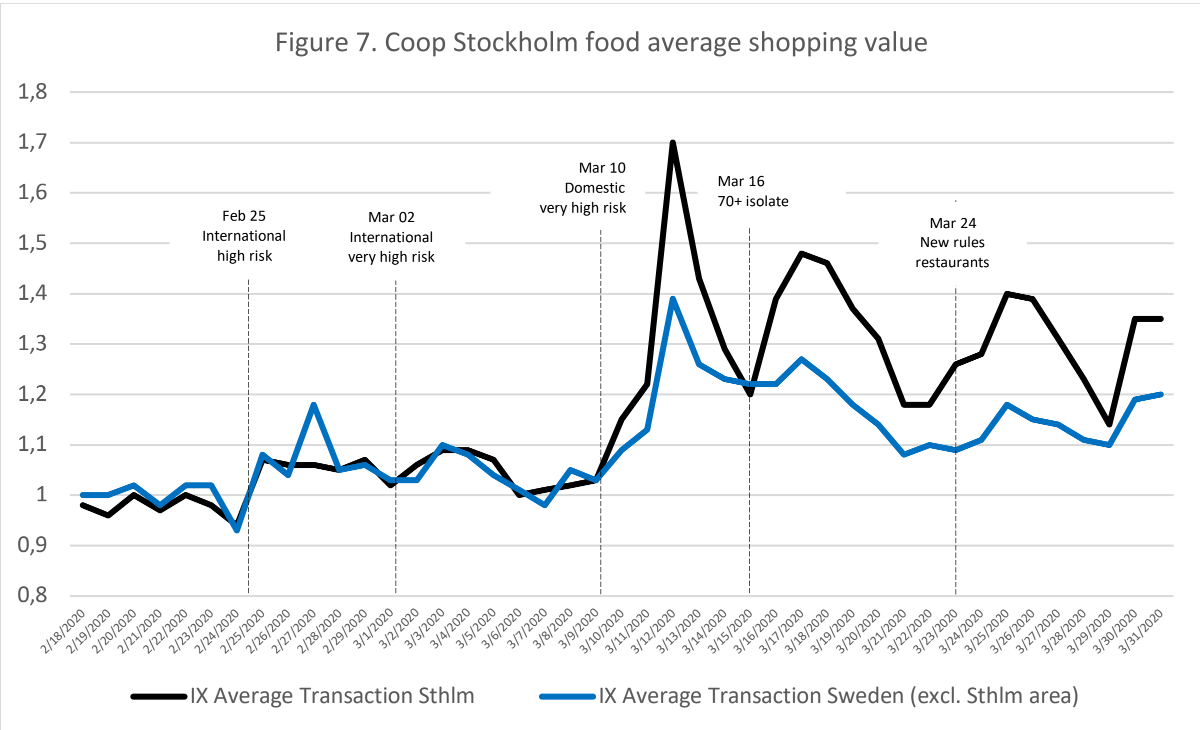


Figure 7 plots the average basket value of each visit at Coop in Stockholm (black) and Sweden outside Stockholm (blue) for January-March 2020. The average basked value is indexed using same week, same weekday sales from 2019 as a benchmark, meaning that a value over 1 is an increase and a value under 1 a decrease versus previous year. For the most part, we see some fluctuations in the average basket value (mostly within the range of -10% to + 10%). More notably we see a big increase in the average basked value staring on March 11. Peak is on March 12, when the average basket in Stockholm is 70% larger than in 2019. For Sweden outside Stockholm the increase is not as large, but still significant (+39%).

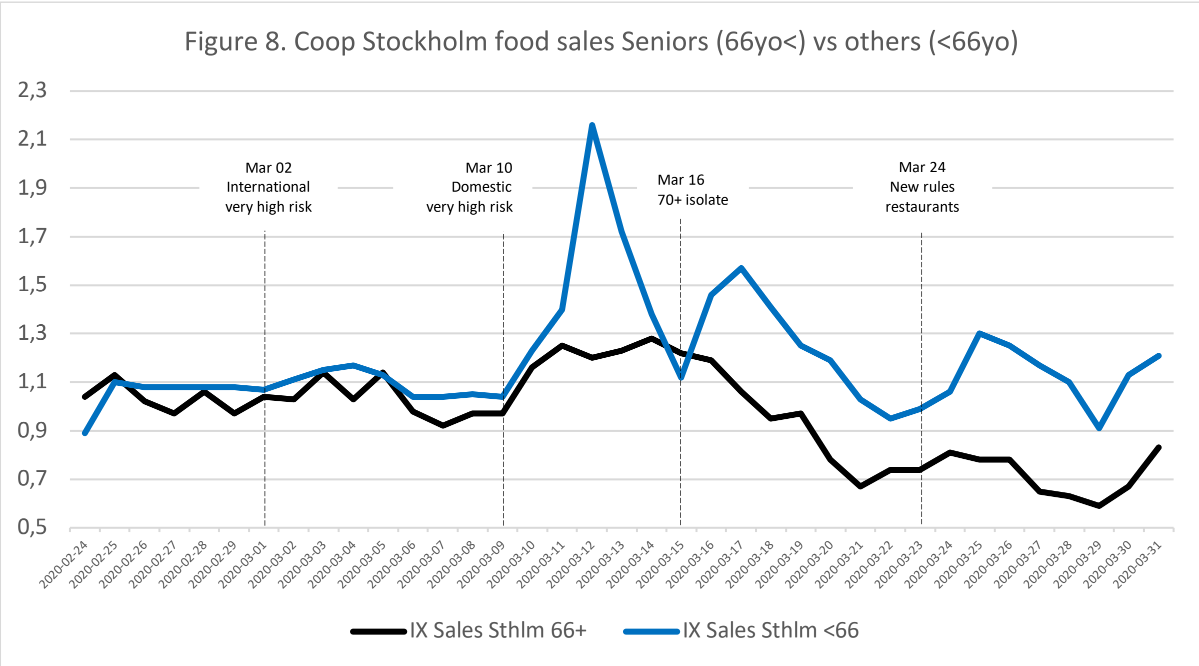
Whereas the bulk buying behaviors in our aggregate level analysis levelled off we see no such difference in the average receipts suggesting that the change in behavior is more permanent. From March 16, Swedes buy groceries less frequently but when they buy, they buy more.



Purchase patterns among risk groups in Stockholm

The retail data allows us to break down the analyses based on different individual traits. As recommendations related to the new Coronavirus are based on age, we follow this procedure and make a disaggregate analysis between older (risk group) and younger (non-risk group) consumers.

Based on the analysis above we notice changed behaviors from March 10. In the following analysis, however, we include data from Monday February 24 as this was the week when the risk for international spread of the new Coronavirus was initially raised. In the following analysis we focus on Stockholm and compare the behaviors of risk groups (age 66 and older) versus others (age up to 65). This is by no mean a perfect assessment of risk groups, but it gives an initial understanding of whether the change in behavior found on the aggregate level are primarily based on changes in the behaviors of the general population or the risk groups.



We can see that it is primarily non-risk groups that start to purchase more in early March. On March 12 the non-risk group in Stockholm increase their purchases by more than double (+116%). On March 17 there is another spike (+57%). For the risk group we see some, but less intense, bulk buying during the period of March 10-March 19. On March 14, the risk group shows its largest spike with +28%.

Could spikes in purchasing patterns post March 16 be due to buying for other people in the risk group? We are currently surveying Coop customers to better understand this. But initial evidence indicates that it might, at least partially, be the case is given by breaking down total sales into number of visits (Figure 9) and average basket value (Figure 10).

When it comes to number of visits, we see a big drop in purchases from the risk group after March 16. From March 20 the number of visits made by this group are significantly reduced by as much as -50% (March 28-30). We also see a shift in the number of visits from the non-risk groups but here the changes are smaller (up to -20%). In terms of average basket value, we see that it increases for both risk and non-risk groups. Initially this effect is larger for the non-risk group.

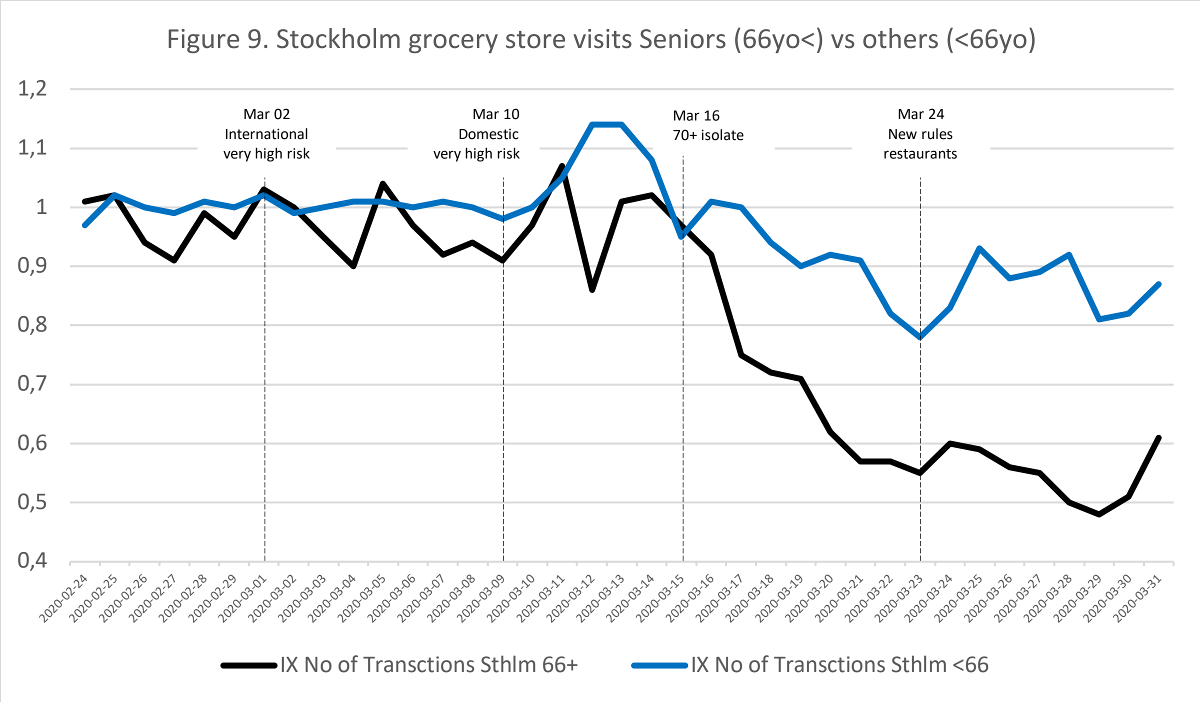


Figure 9. Stockholm grocery store visits Seniors (66yo<) vs others (<66yo)

In Figure 10. Below we can see the average basked value in grocery Stores in Stockholm (risk groups: age 66 and older vs. non-risk groups: up to 65 years of age).

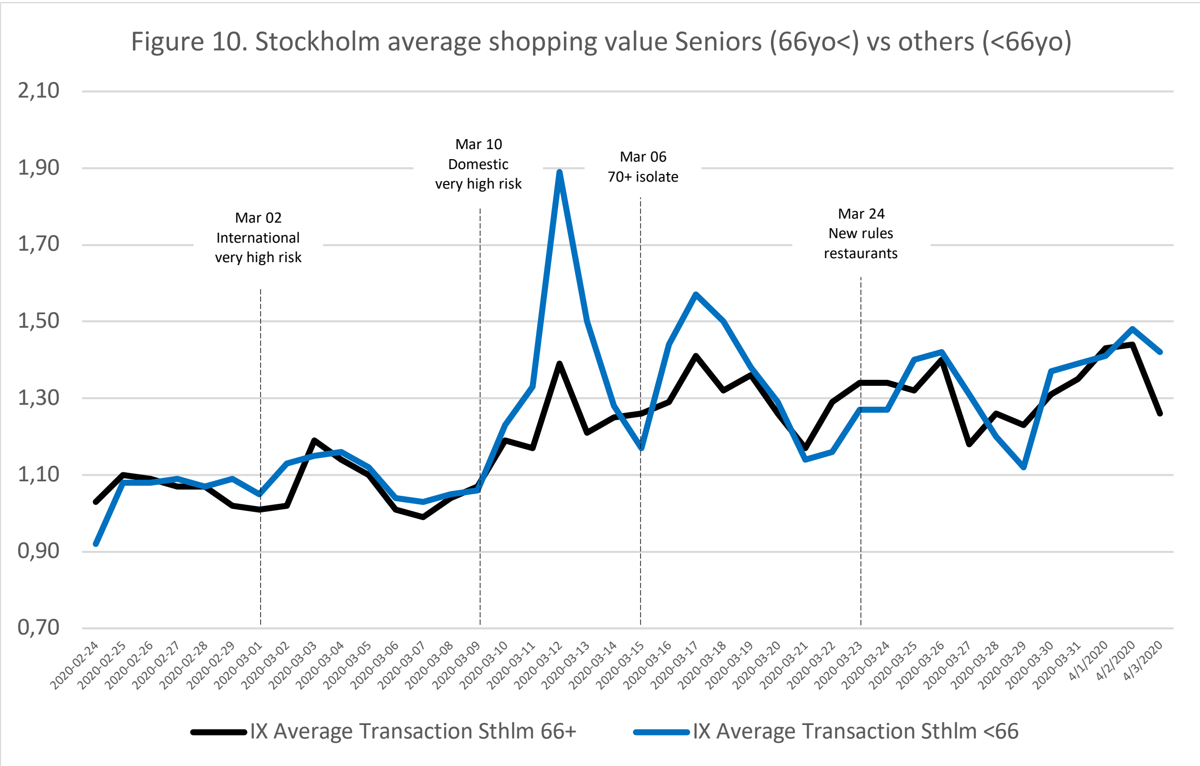


Figure 10. Stockholm average shopping value Seniors (66yo<) vs others (<66yo)

The analysis indicates that the bulk buying, and consequent stock-outs, are mainly due to non-risk groups, rather than risk groups, changing their behaviors. This suggest that that reactions are caused by other things than PHAS guidelines and recommendations. We also see a lag in reactions when it comes to the risk-group consumers. From March 16, however, these group seems to adjust their behavior more to the guidelines.

Further research

An interesting avenue of research would be to dig deeper into the quantity and quality of media coverage and assess how that impacts population behavior. For example, sentiment analysis of editorial and potentially social media could be compared with survey data to examine and quantify how media narratives potentially affect attitudes as a driver for public behaviors.

The retail data available would allow for further analyses based on channel (online vs. offline, big box store vs. convenience) as well as product (bulk buying categories such as toilet paper, pasta, and canned food. It is also possible to compare the behaviors across different regions in Sweden where the outbreak of the new Coronavirus has been higher or lower than in Stockholm. More sophisticated analytical procedures could also be used to model how and when different information and guidelines lead to changes in behavior for different groups (e.g. high risk age groups vs. other age groups), and also examine to what extent risk groups have been able to conform to social distancing and minimize their risk exposure.

Finally, it could be interesting to do more fine-grained comparisons of retail store visit data in combination with mobile data to understand how they together can combined better understanding of population mobility, especially with regards to high risk groups. Additional data sources that would provide commuting and mobility information could be added to assess this question.

Contributions

This paper describes the early stages of an ongoing research project in which we combine data and insights from private sector partners to provide new insights in public behavioral dynamics related to a public health crisis. The analyses above are purely descriptive and can be seen as a first glance in terms of what is possible from a collaboration like this. As such, they illustrate the value of private data for understanding public behaviors and dynamics in a crisis setting. It also highlights the importance of academia as a trusted partner in crafting a facts-based view of what is going on in society and the role of academia as a patterner to pool different private data sources to provide it.

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. The importance of grocery retailers in society has been highlighted in the current crisis due to their important supply chain function. However, as suggested in this working paper another important function that grocery retailers can play is that of an information resource to understand how people are reacting to information and guidelines from authorities. This kind of information should also be useful for understanding public reactions through governmental guidelines. Retail data has several advantages as it is typically large in volume and variety, but also has high velocity. Sales data are continuously updated and tracked by retailers and as such it could be valuable source of information for understanding public reactions in a time of crisis.

Our aim is to continue to add behavioral and potentially attitudinal data from additional sources for the duration of this crisis last. Doing so will enable us to provide insights to governments on how the public is behaving and what kind of information seems to be working as well as to the project partners who through participation will be able to gain a more complete picture of what is going on compared to when looking at their own data only.

References

Bengtsson, L., Lu, X., Thorson, A., Garfield, R., & von Schreeb, J. (2011) Improved Response to Disasters and Outbreaks by Tracking Population Movements with Mobile Phone Network Data: A Post-Earthquake Geospatial Study in Haiti. *PLoS Medicine* 8(8): e1001083.

Bengtsson, L., Gaudart, J., Lu, X., Moore, S., Wetter, E., Sallah, K., Rebaudet, S. & Piarroux, R. Using Mobile Phone Data to Predict the Spatial Spread of Cholera, *Scientific Reports,* Vol 5 (8923) https://doi.org/10.1038/srep08923

Bradlow, E., Gangwar, M., Kopalle, P., and Voleti, S. (2017) The Role of Big Data and Predictive Analytics in Retailing, *Journal of Retailing*, Vol. 93 (1), 79–95.

Cavallo, A. & Rigobon, R. (2016) The Billion Prices Project, *Journal of Economic Perspectives*, Vol. 30 (2), 151-178

Einav, L. & Levin, J. (2014) Economics in the age of big data, *Science*, Vol. 346 (6210) DOI: 10.1126/science.1243089

Germann, F., Lilien, G., Fiedler, L., & Kraus, M. (2014), Do Retailers Benefit from Deploying Customer Analytics? *Journal of Retailing*, Vol. 90(4), 587–93.

Lu, X., Wetter, E., Bharti, N., Tatem, A. & Bengtsson, L. (2013) Approaching the Limit of Predictability in Human Mobility. *Scientific Reports*, Vol 3(2923). https://doi.org/10.1038/srep02923

Oliver, N. et al. (2020) Mobile phone data and COVID-19: Missing an opportunity? Working paper, *arXiv*:2003.12347

Peak, C., Wesolowski, A., zu Erbach-Schoenberg, E., Tatem, A., Wetter, E., Lu, X., Power, D., Weidman-Grunewald, E., Ramos, S., Moritz, S., Buckee, C., Bengtsson, L. (2018) Population mobility reductions associated with travel restrictions during the Ebola epidemic in Sierra Leone: use of mobile phone data, *International Journal of Epidemiology*, Vol 47(5), 1562–1570.

Sallah, K., Giorgi, R., Bengtsson, L., Lu, X., Wetter, E., Adrien, P., Rebaudet, S., Piarroux, R. & Gaudart, J. (2017) Mathematical models for predicting human mobility in the context of infectious disease spread: introducing the impedance model *International Journal of Health Geographics* Vol. 16(42) https://doi.org/10.1186/s12942-017-0115-7

Wedel, M. and Kannan, P.K. (2016) Marketing Analytics for Data-Rich Environments, *Journal of Marketing*, Vol. 80 (November), 97-121.

Wesolowski, A., Buckee C., Bengtsson, L., Wetter, E., Lu, X., and Tatem A. (2014) Commentary: Containing the Ebola Outbreak – the Potential and Challenge of Mobile Network Data, *PLoS Currents*, Edition 1. doi: 10.1371/currents.outbreaks.0177e7fcf522b8b634376e2f3efc5e.

Wilson, R., Erbach-Schoenberg, E., Albert, M., Power, D., Tudge S., Gonzalez M., Guthrie, S., Chamberlain, H., Brooks, C., Hughes, C., Pitonakova, L., Buckee, C., Lu, X., Wetter, E., Tatem, A. & Bengtson, L. (2016) Rapid and Near Real-time Assessments of Population Displacement Using Mobile Phone Data Following Disasters: The 2015 Nepal Earthquake. *PLoS Currents*, Vol. 24(8). February.