Forthcoming, *Canadian Public Policy*

Estimating the Effects of Non-Pharmaceutical Interventions (NPIs) and Population Mobility on Daily COVID-19 Cases: Evidence from Ontario[1]

Nathaniel T. Stevens
Department of Statistics and Actuarial Science
University of Waterloo

Anindya Sen
Department of Economics
University of Waterloo

Francis Kiwon
Department of Statistics and Actuarial Science
University of Waterloo

Plinio P. Morita
School of Public Health and Health Systems
University of Waterloo

Stefan H. Steiner
Department of Statistics and Actuarial Science
University of Waterloo

Qihuang Zhang
Department of Biostatistics
Perelman School of Medicine
University of Pennsylvania

[1] The authors confirm that they do not have any conflict of interest. We are grateful to two anonymous referees an especially to Mike Veall for some very useful advice. All errors are the responsibility of the authors.

Abstract

This study employs COVID-19 case counts and Google mobility data for twelve of Ontario's largest Public Health Units from Spring 2020 until the end of January 2021 to evaluate the effects of Non-Pharmaceutical Interventions (NPIs: policy restrictions on business operations and social gatherings) and population mobility on daily cases. Instrumental Variables (IV) estimation is used to account for potential simultaneity bias, as both daily COVID-19 cases and NPIs are dependent on lagged case numbers. IV estimates based on differences in lag lengths to infer causal estimates, imply that the implementation of stricter NPIs and indoor mask mandates are associated with COVID-19 case reductions. Further, estimates based on Google mobility data suggest that increases in workplace attendance are correlated with higher case counts. Finally, from October 2020 to January 2021, daily Ontario forecasts from Box-Jenkins time-series models are more accurate than official forecasts and forecasts from a Susceptible-Infected-Removed (SIR) epidemiology model.

Keywords: COVID-19; Population Mobility; Google Data; Time-Series Modelling; Forecasts; Ontario; SIR

I. Introduction

With the enactment of stringent restrictions on public mobility and rising vaccination rates, all Canadian provinces began to experience a downward trend in daily COVID-19 cases from June 2021 onwards. There are a limited number of studies that have used econometric modelling to evaluate the effects of Non-Pharmaceutical Interventions (NPIs) on daily cases across Canadian provinces and/or sub-provincial jurisdictions.[2] Prior to vaccines, public health officials maintained that reduced social contact, mobility, and access to businesses, would be the best way to reduce COVID-19 cases. This paper provides some evidence about the possible magnitude of these effects.

This study employs a policy stringency index developed by Karaivanov et al. (2021) to evaluate the effects of NPIs and population mobility on daily COVID-19 cases from April 2nd - September 30th 2020 and across the twelve largest Public Health Units (PHUs) in Ontario.[3] Employing PHU level data enables an evaluation of the effects of business closures and restrictions on public gatherings, while controlling for potentially confounding unobserved jurisdiction-specific and time-invariant characteristics. Publicly available Google data are employed to estimate the effects of population mobility on daily new cases. While NPIs reduce the spread of infections through decreases in population mobility, it is important to study the impacts of overall social mobility on daily case counts, as the effects of stricter policies on population movements may diminish over time with lower public compliance. Further, the use of

[2] According to the Centers for Disease Control and Prevention (CDC], 'Nonpharmaceutical Interventions (NPIs) are actions, apart from getting vaccinated and taking medicine, that people and communities can take to help slow the spread of illnesses like pandemic influenza (flu).' Please see https://www.cdc.gov/nonpharmaceutical-interventions/index.html for further details, last accessed April 17th 2021.

[3] Public Health Units (PHUs) are administrative areas consisting of cities and adjoining suburbs that are charged with overseeing and managing public health according to policies and directives issued by provincial ministries of health. Being the largest province in terms of population, Ontario has the most health regions (36).

Google data enables an assessment of the effects of public mobility to destinations commonly frequented by individuals and households, such as grocery and retail stores and workplaces. Matching these mobility measures to corresponding trends in COVID-19 cases should be useful to policymakers in deciding specific types of economic and social lockdowns, as there is a paucity of knowledge on which types of population mobility are the most responsible for the spread of COVID-19 cases.

The challenge with identifying causal policy effects in this exercise, is that OLS estimates might be confounded and biased downward, as increases in daily cases are also likely to lead to more stringent policies. We attempt to identify a causal interpretation by using lagged cases as instruments, under the assumption that the policy implementation lag is likely longer than the disease transmission lag. Specifically, while current daily cases are impacted by recent daily trends, the impact of successive daily case counts on current case counts should diminish over time. On the other hand, there is a higher probability that the implementation of stricter restrictions on population mobility in response to surges in daily cases are not as immediate and take a longer time-period to occur. While we do not claim that such identification is unimpeachable, standard statistical tests of instrument strength and of overidentification yield statistics suggest that the approach may have some validity, and the resulting IV coefficients indeed suggest stronger policy effects than corresponding single equation estimates.

Some papers have used Google mobility data to understand the spread and propagation of COVID-19 cases in Canada.[4] The study most similar to ours is Karaivanov et al. (2021) who

[4] Most recent research has employed daily social mobility data from Facebook, Google, Apple, and cellular providers (Chan (2020a, 2020b), Goolsbee and Syverson (2020), Maloney and Taskin (2020), Armstrong II et al. (2020), Barrios et al. (2021), and Nguyen et al. (2020)) to study social distancing, rather than estimating the effects of social mobility on COVID-19 spread. However, Glaeser at al. (2020) study the effects of mobility data generated by Safegraph on COVID-19 cases for some U.S. cities, while Kuchler et al. (2021) use aggregated data from

employ data across Ontario PHUs and Canadian provinces to estimate the impact of mask mandates and other NPIs on COVID-19 case growth in Canada. Similar to our research, they attempt to account for behavioural responses by using Google mobility data. However, they average values across different Google mobility measures and focus on the effects on case growth rates as opposed to the incidence of daily cases. Chu and Qureshi (2020) study the relationship between COVID-19 confirmed cases and Google mobility patterns by province/state level in Canada and the U.S. They find evidence of a lagged relationship between Google mobility indicators and case counts. On the other hand, it is not clear which types of social mobility are the most responsible for variation in daily cases as they only consider an aggregate measure of mobility, rather than each individual Google mobility index. Further, both Karaivanov et al. (2021) and Chu and Qureshi (2020) do not employ instrumental variables to account for possible simultaneity bias. Sen (2020) focuses on the lagged effects of different Google social mobility indicators through individual time-series regressions for different Ontario health regions as opposed to pooling data across jurisdictions and over time.

Finally, this study also contributes to the evolving literature on forecasting daily COVID-19 cases by investigating the predictive power of Google population mobility indicators. While there are different research institutes offering long term forecasts based on epidemiological models, the amount of corresponding research on short term predictions is much more limited. Altieri et al. (2021), Bryant and Elofsson (2020), and Liu et al. (2020) are examples of research which have focused on constructing models generating one and two week-ahead forecasts. On the other hand, the methods used by these studies are computationally intensive, involving either

Facebook and demonstrate that the spread of COVID-19 between regions is correlated with increases in the number of Facebook relationships.

different types of linear and exponential predictors or Bayesian methods that are not easily replicable or interpretable. Chu and Qureshi (2020) find that time series models with a quartic trend function can generate comparable short-term out-of-sample forecasts for one to seven-day logarithmic case counts relative to the classic epidemiological SIR (Susceptible, Infected, and Recovered) approach.[5] Chen et al. (2021) employ smooth transition autoregressive (STAR) models, neural network (NN) models, and a susceptible-infected-removed (SIR) model to predict cumulative daily cases counts for Ontario, Alberta, British Columbia, and Quebec and find that NN models outperform other approaches in terms of prediction accuracy. However, these studies do not investigate the usefulness of mobility patterns in generating accurate forecasts. This research evaluates the efficacy of a wide range of Box-Jenkins models (Box et al., 2015). The models used here should be useful for policy purposes, as they are easily interpreted and implementable through standard statistical software packages such as R, STATA, SPSS and Excel.

In terms of primary findings, WLS (weighted least squares) estimates of the COVID-19 policy index are in most cases, statistically insignificant. However, the corresponding IV estimate is statistically significant. Further, increases in the policy index and the implementation of mandatory indoor mask mandates are correlated with reductions in social mobility. Hence, stricter policies may also have an indirect impact in lowering daily cases through decreasing population mobility. The coefficient estimate of mandatory mask mandates is also statistically significant in the IV regression. Both WLS and IV regressions reveal a robust and statistically significant association between increases in workplace mobility and daily COVID-19 cases.

[5] Holmdahl and Buckee (2020) and Liu et al. (2020) have good discussions of findings from recent epidemiological models. Ogden et al. (2020) and Tuite et al. (2020) are examples of Canadian studies that construct long term forecasts based on epidemiological models,

With respect to prediction, we construct one week ahead forecasts of daily COVID-19 cases starting from October 1st 2020 and up to January 31st 2021, updating our models and parameter estimates on a weekly basis. ARIMA models conditioned on weekly seasonality are able to predict daily COVID-19 cases in Ontario with good accuracy, as our daily forecasts differ on average from the actual daily case numbers by roughly 10%. In contrast, predictions generated by a Susceptible Infectious Removed (SIR) model have an average forecast error of roughly 39%.

The remainder of the paper is structured as follows. Sections II discusses the data and results are presented in Section III. Section IV concludes with a summary of key findings and policy implications.

II. Data

Google Mobility Indicators

The mobility data being employed in this research have been extracted from the location history associated with Google Maps app use. The information has been passively generated, collected, and now is being made available for use by researchers and policymakers through Google's 'COVID-19 Community Mobility Reports' that can be downloaded from https://www.google.com/covid19/mobility/.

The Google Mobility data captures total visits to the following specific destinations commonly frequented by individuals and households: (1) grocery and pharmacy stores, which include grocery markets, food warehouses, farmers markets, specialty food shops, drug stores, and pharmacies; (2) parks, which consist of local parks, national parks, public beaches, marinas, dog parks, plazas, and public gardens; (3) transit stations, comprising of subway, bus, and train stations; (4) retail stores & recreation outlets consisting of places like restaurants, cafes, shopping centers, theme parks, museums, libraries, and movie theaters; and (5) workplaces. With respect to places of residence, google social mobility represents duration of stay.

As detailed on its website, Google creates these aggregated and anonymized sets of data from users who have turned on the Location History setting of Google accounts on their phones and have agreed to share this information. Consequently, a limitation to acknowledge is that Google data on social mobility trends are based on a sample of users who own mobile devices and who chose to share their location history. These data therefore may not be representative of the population. Additionally, Google has not made public its precise methodology for calculating social mobility. Hence there is some ambiguity on the extent to which Google data are representative of population level trends. However, data on the number of people using Google Maps in the U.S. indicates that Google social mobility indicators might be fairly representative of overall population trends.[6]

Daily values are aggregations across individuals who have enabled their location history and are available for each province in Canada from February 15th onwards. These values are calculated relative to a baseline, which is defined as the median for the corresponding day of the week, during the 5-week period January 3rd – February 6th, 2020. Hence, each daily value is the percentage change in the social mobility category relative to its baseline, which shows how visits and length of stay at different destinations have changed since the onset of the pandemic. A visual inspection is useful to evaluate whether trends in social mobility correspond with intuition.

---

[6] Statistics Canada data reveals that 88% of Canadians (15yr+) have a smartphone (https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=2210011501). According to Statista, in the U.S., Google Maps had 154 million users in April 2018 (https://www.statista.com/statistics/865413/most-popular-us-mapping-apps-ranked-by-audience/). However, appropriate caution should be used in interpreting results based on Google data, given that Google has not revealed how it aggregates individual level to create a geographic specific index.

A caveat is that while data for all Google social mobility indicators are available at the province level, this is not the case for PHUs, with many missing values for park, transit, and residences.

Place Figure 1

Figure 1 contains trends in the different social mobility categories for the entire province from April 2nd 2020 – January 31st 2021. Grocery and pharmacy mobility increased over the sample period along with mobility at retail and recreational venues. However, a difference is the decline in retail social mobility from August onwards. Residential duration of stay was initially high and then fell during the latter part of the sample period. Unsurprisingly, movements in work and transit mobility are significantly correlated, with both indicators increasing over time. The top spikes in both these variables are mobility values during the weekend, which did not significantly decline relative to pre-pandemic observations. Finally, the sharp rise and fall of social mobility at parks reflects outdoor activities in the warmer months. Additionally, the time series visualized in Figure 1 all show a strong seasonal day-of-week effect, meaning that every seventh observation is highly correlated.

Trends in Daily COVID-19 Cases

Figure 2 visualizes trends in daily cases over the same period. These data have been made publicly available by the Ontario provincial government via its online data-sharing website.[7] Ontario is the only province in Canada and one of the few jurisdictions in the world, which publishes daily COVID-19 case data based on date of specimen collection. This is an important qualification to using social mobility data to capture trends in population movements in COVID-19 cases. Specifically, relying exclusively on daily case data constructed using date of

[7] The data are available from https://www.publichealthontario.ca/en/data-and-analysis/infectious-disease/covid-19-data-surveillance/covid-19-data-tool.

confirmation of test results, might lead to misleading estimates of the relationship between social mobility and daily cases if there are significant and inconsistent delays in the release of test confirmations.

Place Figure 2

As can be seen in Figure 2, the number of new cases each day steadily increased in Ontario until the second week of April when it began to decrease. This decline continued until mid August after which a sharp increase in cases began to occur – an increase that continued through the remainder of 2020. This sharp increase, however, was followed by an even sharper decline early in the new year. Increases in mobility seen throughout the Summer preceded the corresponding increase in cases we observe in the late Summer and into the Fall. Likewise, the decline in mobility in the Fall preceded the decline in cases observed early in the Winter. This suggests that past social mobility information may be useful in forecasting future COVID-19 case counts. From a modeling perspective we see some notable structure that, when accurately accounted for, may be exploited for purposes of predicting daily new COVID-19 case counts. For instance, the general pattern of increases and decreases just discussed represents a strong non-linear trend that should be accounted for. Additionally, just like the Google mobility data, we see a strong seasonal day-of-week effect. Accounting for the weekly seasonality exhibited by both this data and the Google mobility data will be very important. Note that we also observe an increase in volatility in daily case counts as times passes.

Our regression analyses exploit differences in daily cases across the twelve largest public health units in Ontario. In particular, we have data for the following twelve PHUs (with population size in parentheses): Durham (645,862); Hamilton (1,399,073); Halton (548,430); Middlesex-London (455,526); Ottawa (1,306,249); Niagara (447,888); Peel (1,381,744);

Simcoe-Muskoka (540,249); Waterloo (535,154); Windsor (398,953); Toronto (2,731,571); and York (1,109,909). Cumulatively, these health units account for more than 85% of the province's population. Other PHUs have much smaller populations and did not experience a significant number of COVID-19 cases. In terms of sample means of daily cases for April 2nd – September 30th 2020, the PHUs are ranked as follows (with the mean of daily cases in parentheses): Toronto (100.41); Peel (50.86); York (22.06); Ottawa (16.96); Windsor (13.344); Durham (11.273); Waterloo (8.9126); Halton (6.169); Hamilton (5.781); Niagara (5.24); Simcoe-Muskoka (4.754); and Middlesex-London (4.01).

Policy Variables

The effects of NPIs at the province level are measured through the Bank of Canada Policy Stringency Index created by Cheung et al. (2021). This index is based on the methodology of the Oxford COVID-19 Government Response Tracker (OxCGRT) developed by the University of Oxford's Blavatnik School of Government.[8] The index is comprehensive in capturing different policies aimed at restricting public mobility and include: school and university closures; workplace and office closures; public event cancellations and restrictions; restrictions on private gatherings; public transport closures; stay-at-home requirements; restrictions on intra-provincial travel (between cities or regions within a province); restrictions on interprovincial travel (between provinces); restrictions on international travel; enforcement mechanisms for individuals; enforcement mechanisms for firms; and public information

[8] For more details on methodology please see Hale et al. (2020).

campaigns.[9] The index ranges from 0 (no restrictions) to 1 (maximum restrictions). We employ this index in our forecasting of Ontario level daily cases.

To estimate the effects of NPIs at the PHU level, we employ the policy stringency index created by Karaivanov et al. (2021). Karaivanov et al. (2021) were able to compile restrictions on business operations, and compute values (from 0 to 1) capturing the intensity of restrictions for businesses and gatherings. Their index captures restrictions on: non-essential and retail business; personal services business; restaurants, bars and nightclubs; places of worship; events and gatherings; and recreation, gyms and parks. Zero denotes no restrictions; and 1 denotes the highest level of restrictions.

Place Figure 3

As noted, much of our estimation uses the period April 2nd to September 30th. Accordingly, Table 1 gives some summary statistics for Ontario aggregate data over that period and Figure 3 decomposes time-series variation in the PHU Index for the largest regions of Durham, Toronto, Peel, York, Ottawa, and Hamilton, as well as the corresponding values for the Bank of Canada Policy Index for Ontario. Figure 4 depicts movements in the PHU index for the smaller regions of Halton, Middlesex-London, Niagara, Waterloo, Simcoe-Muskoka, and Windsor. As can be seen in Figure 3, the PHU Index is identical for; Durham and Hamilton; and Peel, Toronto, and York. There is time-series variation for all PHUs as captured by the loosening of restrictions on mobility through the sample period. There is also variation across PHUs with relaxation in restrictions for Peel and Toronto in late June and July, following the lifting of restrictions on mobility in other health regions. The Bank of Canada Policy Stringency Index

[9] Please refer to Cheung et al. (2021) for further details. We are grateful to an anonymous referee for bringing this index to our attention.

follows a similar decline through time and is highly correlated with the PHU Index for Toronto (Pearson correlation coefficient of 0.95). Figure 4 shows a similar relaxation of mobility restrictions for smaller PHUs over the sample period. Halton and Niagara have identical trends, as do London, Waterloo, and Simcoe-Muskoka. For these health regions, lifting of policy restrictions occurs later in Windsor, relative to other health regions.

Place Figure 4

We also construct a dummy variable to represent the implementation of mask mandates in indoor settings. There is time-series variation in their enactment across PHUs. The dates are: Durham (July 10, 2020); Halton (July 22, 2020); Hamilton (July 20, 2020); Middlesex-London (July 18, 2020); Niagara (July 31, 2020); Ottawa (July 07, 2020); Peel (July 10, 2020); Waterloo (July 13, 2020); Simcoe Muskoka (July 13, 2020); Toronto (July 07, 2020); Windsor-Essex (June 26, 2020); and York (July 17, 2020).[10]

III. Results

The Effects of NPIs on Population Mobility

Table 2 reports results of basic WLS regressions with retail and recreational, grocery and pharmacies, and workplace Google mobility indicators as dependent variables.[11] The motivation is to explore the impacts of policy stringency on mobility. In this respect, the local COVID-19 Policy Stringency Index may share a different relationship with population mobility measures,

[10] This information was taken from Karaivanov et al. (2021).

[11] We follow Karaivanov et al. (2021) in using Weighted Least Squares (WLS), where observations are weighted by PHU population size. Parks, transit, and residential mobility are omitted given missing observations for some Public Health Units. The control variables are: one and two-day lags in the dependent variable; the COVID-19 Policy Stringency Index; the mask mandate dummy; average daily temperature for the PHU; PHU specific fixed effects; and day of week dummies as covariates.

relative to indoor mask mandates. Specifically, increases in the Stringency Index should be associated with reductions in social mobility, as it captures restrictions on public gatherings and access to businesses. The implementation of mask mandates in indoor settings might also be correlated with lower public mobility if individuals view such regulation to be indicative of a heightened risk of infection to the public. On the other hand, if individuals feel safer with mask mandates, it is possible that such regulation will result in increased public movements.[12]

The results reveal that controlling for other factors, both stricter policies on mobility and mask mandates are significantly correlated (at the 1% level) with reduced mobility as measured by all Google variables. In most cases, the lagged dependent variables are statistically significant. The coefficient estimate of the average temperature covariate is positive and also statistically significant, which reflects the association between warmer temperatures and higher social mobility. Although the underlying model is simple, the adjusted R^2 is above 0.7 with respect to retail and workplace mobility. Given that an objective of stricter COVID-19 policies is to reduce public mobility, these results suggest that these specific initiatives were successful.

The Effects of NPIs & Population Mobility with PHU Data

Table 3 contains Weighted Least Squares (WLS) regression results based on data pooled across twelve PHUs and over time. The dependent variable is the number of daily cases. Column (1) contains estimates of the local Policy Stringency Index conditioned on one and two-day lagged cases, PHU dummies, and day of week dummies. Column (2) adds the mandatory mask dummy, one-week lagged Google mobility and average temperature variables, while Column (3)

[12] This possibility is based on the existence of risk compensation as a part of rational decision-making. For example, some previous studies find that the implementation of mandatory seatbelt legislation can be associated with more injuries or accidents, as individuals respond to a feeling of enhanced safety by driving more aggressively. Please see Sen (2001) for a discussion of the literature.

adds three, four, five, six, and seven day lagged dependent variables to assess the fit of a more dynamic specification.[13] [14] Column (4) contains second-stage IV results where the Policy Stringency Index is instrumented by 12, 13, 14, and 15 day lags of daily cases.

The motivation for employing IV analysis is to account for the possibility that single equation estimates of the local Policy Stringency Index might be biased downwards by simultaneity bias. Single equation models assume that changes in policy exogenously impact daily cases. However, changes in historical daily case trends may also influence the enactment or easing of more stringent policies aimed at restricting public mobility. The 12-15 day lagged dependent variables we employ are far enough in the past, that it can be argued that they should not be strongly correlated with the current daily COVID-19 cases. We used multiple lagged values to lessen the possibility of a spurious correlation and to enable a test of overidentifying restrictions.[15]

The coefficient estimate of the local Policy Stringency Index is negative and statistically significant at the 1% level in Column (1) but becomes insignificant in Columns (2) and (3) with the addition of the mask mandate dummy and other control variables. One and two-day lags of the dependent variable are positive and statistically significant at the 1% or 5% levels in all columns. The mask mandate dummy is statistically insignificant in Columns (1) – (3). In Column

[13] We are grateful to an anonymous referee for recommending this sensitivity test.
[14] Seven day-lags of mobility variables are used as they remained significant in LASSO regressions after employing different combinations of lagged values.
[15] The implementation of mandatory mask regulation may also be endogenous to rising COVID-19 case counts and government advisories. Studies based on self reported mask use in Canada (Sheluchin et al. (2020) and Jehn et al. (2021)) report public increases in mask usage that are correlated with public health advice and, which occurred during the early part of the pandemic. However, an argument might be made that population mask use is likely more endogenous with respect to daily COVID-19 cases, relative to the enactment of mandatory mask regulation. This is because of the ease in which individuals may alter behaviour in response to perceived risk. In any case, our inability to conduct IV analysis of mask mandates is a shortcoming that we acknowledge. We were unable to identify plausible instruments that matched time-series variation in mask mandates across Public Health Units.

(3), the three, six, and seven-day lags are statistically significant at the 1% levels. Coefficient estimates of seven day lagged retail mobility are statistically significant at the 5% levels in Columns (2) and (3) and imply that a 10-percentage point mobility is correlated with on average, a roughly 2-3 daily case increase across PHUs. Coefficient estimates of groceries and pharmacies are significant in Columns (2)-(3) but have a counter-intuitive negative sign. Results in the same columns imply that a 10-percentage point rise in work mobility is associated with approximately a 1 daily case increase (statistically significant at the 1% level). Higher temperatures are significant at the 1% level and possess the expected negative signs, as an increase in temperatures should result in more outdoor and socially distanced mobility, and therefore, fewer cases.

Results in Column (4) confirm the possibility that WLS estimates of the Policy Index are likely biased downward by simultaneity bias, as the IV coefficient estimate of the Policy Index is negative and statistically significant (at the 5% level). The *p* value and *F*- statistic of the test of joint significance of the instruments (reported in the table) enables us to reject the null hypothesis that coefficient estimates of 12-15 day lags in the dependent variable from the first stage regression are zero. Further, the use of multiple instruments allows us to employ a Sargan test for overidentifying restrictions. As reported in Table 3, we could not reject the null hypothesis. Nonetheless, the statistical significance of the IV estimate should be treated with caution as it is based on a specific set of lagged dependent variables.

With respect to other findings from Column (4), the mask mandate dummy is negative and statistically significant at the 5% level. The lagged grocery and retail mobility variables are either statistically significant at the 1% or 5% levels. However, their signs are reversed relative to corresponding WLS estimates. Hence this paper does not provide evidence that retail mobility increases daily cases. In contrast, the coefficient estimate of workplace mobility remains positive and significant at the 1% level. [16] In summary, the IV results imply that more stringent policies – as measured through the local Policy Stringency Index and the mask mandate dummy - are correlated with lower daily cases. As demonstrated by the results in Table 2, stricter policies on social mobility are also associated with reduced population movements, which in turn, are correlated with lower daily case counts.

Forecasting

- *Forecasting at the Ontario Level*

Given that the time series nature of case counts in Ontario exhibits non-stationarity and strong weekly effects, we use Box-Jenkins models for their ability to flexibly model and forecast complex correlation structure. We specifically consider pure seasonal autoregressive integrated moving average (SARIMA) models (which model daily new COVID-19 case data as a function of historical daily cases only) as well as SARIMA models augmented with all Google mobility variables and the BOC policy index.[17] These latter models may also be thought of as regression with SARIMA errors. It should be noted that in the context of SARIMA models, a seasonal effect is one that recurs predictably with some fixed frequency, independent of the specific frequency, which could be weekly, monthly, or quarterly, Hence, SARIMA models are

[16] In contrast to our results, Karaivanov et al. (2021) do not find Google mobility indicators to be statistically significant. However, there are possible reasons for this difference in research findings. First, Karaivanov et al. (2021) focus on weekly case growth rates as opposed to the number of daily new cases. Second, they use an average across population mobility indicators, rather than the actual individual values. Third, they may be employing daily cases by reported date as opposed to by date of specimen collection.

[17] We employ all Google mobility variables when forecasting for the province as opposed to Public Health units, as there are no missing values at the aggregate province level.

appropriate to control for the day of week trends that we observe in our cases and Google mobility data.

We choose training and forecasting periods to avoid the confounding effects of public vaccination programs. Specifically, we partition the available data into training and testing sets where the training data (April 2nd – September 30th) is used to fit the model and the test data (October 1st – January 31st) is used to evaluate the accuracy of the model's forecasts. This partition has also been chosen for illustration because it showcases the model's ability to accurately forecast the pronounced increase in cases that began in September of 2020 as well as the sharp decline that followed in January of 2021. Further details of our SARIMA modeling are available in the Appendix.

Figure 5 visualizes the fit and forecasts of a $SARIMA(1,1,2)(1,1,2)[7]$ model without any exogenous information, over the training and testing time-periods.[18] The shaded regions represent 95% prediction intervals, and the vertical dashed line separates training from testing data. As evidenced by Figure 6, the model fits and forecasts the data very well. We quantify this good performance using the mean absolute error (MAE) which calculates, on average, the absolute difference between a forecast (blue line) and the true count (black line). For this model and these data, we have MAE = 172.8855, meaning that on a typical day, our forecasted daily case count is roughly 173 away from the truth. Quantified another way, the mean absolute percent error is MAPE = 9.74%.

Place Figure 6

[18] In $SARIMA(1,1,2)(1,1,2)[7]$, the first part of the notation (1,1,2) denotes the non-seasonal part of the model with the autoregressive part (p) = 1, differencing (d) = 1, and the MA component (q) of the process = 2. The seasonal part is given by the second bracket (1,1,2), and m = the seasonal period or the number of observations per time-period, which in this case is the number of days in a week, 7.

We may also quantify the efficacy of the methodology by considering the accuracy of the interval forecasts. In particular, we observe whether the prediction intervals (grey shaded area) contain the true daily case count (black line). An especially important time frame to consider is the first week of January. During this time Ontario saw a dramatic change from increasing to decreasing case numbers in a matter of one week. The blue line (which deviates more than usual from the black line in this time frame) indicates that the model's forecasts did not immediately predict the sudden downward trend; it was not until the second and third weeks of January that the forecasts re-aligned with actual case numbers. However, the 95% prediction intervals captured this sudden and dramatic change in trend, indicating the value of accurate interval estimates.[19]

Although not depicted visually, the predictive accuracy associated with SARIMA models that *do* include the Google mobility variables and the BOC stringency index perform similarly. The top section of Table 4 reports corrected AIC, MAE and MAPE values for four versions of SARIMA: the pure SARIMA model depicted in Figure 5; the SARIMA model that includes the BOC stringency index, but not the Google mobility variables; the SARIMA model that includes the Google mobility variables, but not the BOC stringency index; and the SARIMA model that includes the Google mobility variables as well as the BOC stringency index. The results indicate that the pure SARIMA model is (strictly speaking) superior but that including the exogenous variables does not drastically worsen performance. Though not included here, these conclusions generalize to different train-test partitions.

[19] Although not shown here, comprehensive residual diagnostics were performed confirming that the residuals were stationary (Augmented Dickey Fuller Test p-value = 0.01), uncorrelated (Ljung-Box Test p-value > 0.8 for lags in [1,10]), and homoscedastic (Levene Test p-value = 0.09), and hence that the necessary modeling assumptions are satisfied.

- *Forecasting at the Public Health Unit Level*

We also investigated using a SARIMA model like the one specified in the Appendix to forecast daily cases in each of the PHUs individually.[20] However, because of missing data associated with the parks, transit, and residential Google mobility variables at the PHU level, we omit them and focus on retail & recreation mobility, grocery & pharmacy mobility, and workplace mobility. Such models proved to be ineffective for small PHUs with relatively low case counts. As such, we present here the results only for the largest six PHUs: Durham, Toronto, Peel, York, Ottawa, and Hamilton. SARIMA models for the daily cases in each of these PHUs are visualized in Figure 6, and their prediction accuracy is quantified in the bottom six sections of Table 4. For illustration, the models depicted in Figure 6 are SARIMA models that include Google mobility variables but not the stringency index. Unfortunately, at the regional level there isn't one model specification that is uniformly superior to the others across all jurisdictions, but we see that including either the Google mobility variables or the stringency index is advisable. This is in contrast to forecasts with the Ontario level data, in which the policy variable does not seem to be important. The forecast errors with these models are on the order of 15%-30%, with values of 14%-15% for Toronto and Peel and roughly 20% for York, which also happen to be the worse hit PHUs in the province, with respect to daily COVID-19 case counts.

A valid question is how our forecasts compare against corresponding government projections. Through its COVID-19 Science Advisory Table (https://covid19-sciencetable.ca/), the province of Ontario collects information and data on COVID-19 health impacts as well as projections of daily cases that are compiled by different experts and researchers, which are also

[20] We still relied on the BOC Policy Stringency index in PHU level forecasting as we do not possess data for PHU level policy indices from December 2020 onwards.

released to the public. Unfortunately, these public briefs do not offer specific numerical daily forecasts but only time-trends through graphs. The forecasts are based on a fixed daily percentage increase in COVID-19 cases. Projections available from the Government of Ontario (September 30th 2020) specifically indicate a belief that daily case counts could reach more than 1,000 cases a day during mid-October.[21] Based on actual daily cases, this is consistent with a roughly 3.5% daily increase in reported cases from mid-September to mid-October. This daily increase results in an absolute forecast error of 16% with respect to daily case predictions. Over the same time period, our SARIMA models with Google mobility variables and the BOC Policy Index, produces forecasts over the same time-period with forecast errors of roughly 18 percent.

On the other hand, the SARIMA model with exogenous variables produces much more accurate daily forecasts between November 16th – December 15th 2020, compared to predictions generated by the Ontario COVID-19 Science Advisory Table for this time-period. In this report, specific daily case growth rates of 3% and 5% are assumed from mid-November to December 2nd 2020. When compared with actual daily cases, the results have MAPE values of roughly 22% and 69%, when assuming 3% and 5% growth rates in daily case counts. Our model produces daily forecasts with a much lower error of approximately 10% with Google mobility variables and the BOC Policy Index. Hence, for this time-period, the SARIMA forecasting model with exogenous variables offers more accurate predictions relative to available government projections.

[21] On page 8 of the slide deck, there is a statement: "This forecasting suggests Ontario could be around 1,000 cases per day in the first half of October." The report is available here https://files.ontario.ca/moh-fall-prep-modelling-deck-en-2020-09-30-v2.pdf.

- *Comparison with a Susceptible-Infected-Removed (SIR) Model at Ontario Level*

As a final sensitivity exercise, we evaluate the performance of the SARIMA model by constructing forecasts from a SIR model over the same period. Despite the widespread use of Box-Jenkins methods for forecasting, SIR models are the dominant methodology to model the spread of epidemics.[22] The SIR model uses a differential equations approach to model changes in the number of infections (I), by incorporating population size (N), the susceptibility of the population to the disease (S), and recovery rates (R). Infections (I) are calculated based on daily cases and recovery (R) is counted as the number of daily recovered and deceased individuals.[23] S is calculated as a function of β, which is the average number of contacts per infectious person per time unit. We specifically employ the approach detailed in Chen et al. (2021). As is the case with our SARIMA modeling, the parameters of the SIR model are updated weekly and the model is employed to construct week ahead daily predictions. Figure 7 visualizes SIR forecasts for Ontario from October 1st 2020-January 31 2021 against predictions generated from the SARIMA model with no exogenous variables and a "naïve model" in which the daily forecast is the average case count in the previous seven days. On average, the SIR model performs poorly against SARIMA, as the MAPE over the testing period is approximately 39%. In particular, we note that SARIMA is able to predict the downward trend in daily cases during January, while the SIR model forecasts a continuing rising trend. Even the MAPE from the naïve model at roughly 16%, is lower than the MAPE of SIR forecasts.

[22] Please see Tolles and Luong (2020) for further details.
[23] Specifically from https://data.ontario.ca/en/dataset/status-of-covid-19-cases-in-ontario.

IV. Conclusion

This paper studies the effects of Non-Pharmaceutical Interventions (NPIs) in the form of policy restrictions on businesses and public gatherings and population mobility on daily cases in the twelve largest Public Health Units (PHUs) in Ontario. These estimates are conditioned on the use of Google mobility data, which is intended at controlling for the magnitude of population level movements. Given declining daily case counts across the country and increases in vaccination rates, it is important to gain an understanding of the effects of government policies and public mobility on COVID-19 cases during a time-period in which cases were rising rapidly and vaccines were unavailable.

Results from IV regressions based on PHU level data demonstrate that stricter policies are correlated with reductions in daily COVID-19 case counts. Increases in the local Policy Stringency Index and the enactment of mask mandates is associated with reduced public mobility. We observe a statistically significant positive correlation between mobility at workplaces and daily cases. This is unsurprising given recent evidence that for some regions, surges in COVID-19 infections are associated with congested workplaces, prompting Toronto and Peel regions to ask all businesses with five or more employees to shut down for a ten-day period.[24] We note that the coefficient estimate of mask mandates is statistically significant in our IV regression. Caution should be used in interpreting this result given that we did not instrument the mask dummy variable. On the other hand, the importance of mask mandates cannot be dismissed given the robust correlation between such regulation and reduced case growth

[24] https://www.theglobeandmail.com/canada/article-ontarios-peel-region-sees-surge-in-workplace-covid-19-infections/, and https://toronto.ctvnews.ca/toronto-peel-require-all-businesses-with-5-or-more-work-acquired-covid-19-cases-to-shut-for-10-days-1.5394284.

obtained by Karaivanov et al. (2021). A conservative interpretation of the effects of indoor mask mandates can be obtained by contextualizing the IV coefficient of -12.25 against the mean of daily cases in Toronto and Peel Region, the two PHUs with the highest sample means of 100.41 and 50.86, respectively. These summary statistics imply that mask mandates are associated with roughly a 12% and 24% decline in daily cases in Toronto and Peel Region.

Another objective of this research was to develop time series models that are capable of forecasting daily new COVID-19 cases. In this respect, SARIMA models fit daily Ontario data very well and provide accurate forecasts over the four-month period October 1st 2020 – January 31st 2021 that are roughly 10% different from actual values. Google mobility variables and the BOC stringency index do not offer much help in improving 7-day forecast accuracy for aggregated Ontario data, but at the PHU level, and over longer forecasting time frames, these exogenous variables do indeed help to improve forecast accuracy.

When compared against available evidence, forecasts based on SARIMA models with exogenous variables are comparable to government projections from mid-September to mid-October, but superior to corresponding predictions between mid-November to mid-December. Finally, we benchmark the SARIMA forecasts against corresponding forecasts generated from a SIR model. On average, the MAPE in SIR forecasts is 39%, which is much higher than the MAPE in SARIMA predictions (10%).

References

Altieri, N. et al. (2021). "Curating a COVID-19 Data Repository and Forecasting County-Level Death Counts in the United States." *Harvard Data Science Review*. https://doi.org/10.1162/99608f92.1d4e0dae

Armstrong II, David A., Lebo, Matthew J., and Lucas, Jack (2020). "Do COVID-19 Policies Affect Mobility Behaviour? Evidence from 75 Canadian and American Cities." *Canadian Public Policy* 46(S2): S127-S144.

Barrios, J. M., Benmelech, E., Hochberg, Y. V., Sapienza, P., & Zingales, L. (2021). "Civic Capital and Social Distancing during the COVID-19 Pandemic." *Journal of Public Economics* 193, 104310. https://doi.org/10.1016/j.jpubeco.2020.104310.

Box, George E.P., Jenkins, Gwilym M., Reinsel, Gregory C., Ljung, Greta M. (2015). *Time Series Analysis: Forecasting and Control,* 5ed. John Wiley & Sons Inc.

Bryant, Patrick and Arne Elofsson (2020). "Estimating the Impact of Mobility Patterns on COVID-19 Infection Rates in 11 European Countries." *PeerJ* vol. 8 e9879. doi:10.7717/peerj.9879.

Cavanaugh, Joseph E. (1997). "Unifying the Derivations for the Akaike and Corrected Akaike Information Criteria." *Probability Letters* 33(2): 201-208.

Chan, Jeff (2020a). "Using Google Data to Understand Canadian Movement Reductions During the COVID-19 Pandemic." Available at SSRN: https://ssrn.com/abstract=3599227 or http://dx.doi.org/10.2139/ssrn.3599227.

Chan, Jeff (2020b). "The Geography of Social Distancing in Canada: Evidence from Facebook." *Canadian Public Policy* 46(S1): S19-S28.

Chen, L.P., Zhang, Q., Yi, G.Y., and W. He (2021). "Model-Based Forecasting for Canadian COVID-19 Data". *PLoS One* 16(1):e0244536. doi: 10.1371/journal.pone.0244536. PMID: 33465142; PMCID: PMC7815137.

Cheung, Calista, Lyons, Jerome. Madsen, Bethany, Miller, Sarah, Sheikh, Saarah. (20201). "The Bank of Canada COVID-19 Stringency Index: Measuring Policy Response across Provinces." Staff Analytical Note 2021-1, available at https://www.bankofcanada.ca/2021/02/staff-analytical-note-2021-1/.

Chu, Ba and Shafiullah Qureshi (2020). "Predicting the COVID-19 pandemic in Canada and the US." *Economics Bulletin* 40(3): 2565-2585.

Glaeser, Edward L., Gorback, Caitlin S., and Redding, Stephen J. (2020). "How Much does COVID-19 Increase with Mobility? Evidence from New York and Four Other U.S. Cities." *Journal of Urban Economics*. doi: 10.1016/j.jue.2020.103292.

Goolsbee, Austan and Chad Syverson (2020). "Fear, Lockdown, and Diversion: Comparing Drivers of Pandemic Economic Decline 2020*." Journal of Public Economics* 193. doi: 10.1016/j.jpubeco.2020.104311.

Government of Ontario (September 2020). "COVID-19: Modelling Update", available at https://files.ontario.ca/moh-fall-prep-modelling-deck-en-2020-09-30-v2.pdf.

Hale, T., Angrist, N., Cameron-Blake, E., Hallas, L., Kira, B., Majumdar, S., Petherick, A., Phillips, T., Tatlow. H., and S. Webster (2020). "Variation in Government Responses to COVID-19." Blavatnik School of Government, University of Oxford, Working Paper No. 2020/032, available https://www.bsg.ox.ac.uk/research/publications/variation-government-responses-covid-19.

Holmdahl, Inga and Caroline Buckee (2020). "Wrong but useful — what COVID-19 epidemiologic models can and cannot tell us." *New England Journal of Medicine*: 383: 303-305 DOI: 10.1056/NEJMp2016822.

Jehn, Anthony, Stackhouse, Matthew, and Zajacova, Anna (2021). "COVID-19 Health Precautions: Identifying Demographic and Socio-Economic Disparities and Changes over Time." *Canadian Public Policy* 47(2): 252-264. https://doi.org/10.3138/cpp.2020-138.

Karaivanov, Alexander et al., Lu, Shih En, Shigeoka, Hitoshi, Chen, Cong & Stephanie Pamplona (2020). "Face Masks, Public Policies and Slowing the Spread of COVID-19: Evidence from Canada." *Journal of Health Economics* 78: 102475.

Kuchler, Theresa, Russel, Dominic & Johannes Stroebel (2020). "The Geographic Spread of COVID-19 Correlates with the Structure of Social Networks as Measured by Facebook." Forthcoming, *Journal of Urban Economics*.

Liu, Laura, Moon, Hyungsik Roger, and Frank Schorfheide (2020). "Panel forecasts of country-level COVID-19 infections." Available at https://laurayuliu.com/covid19-panel-forecast/ and forthcoming, *Journal of Econometrics*.

Maloney, William and Temel Maskin (2020). "Determinants of Social Distancing and Economic Activity during COVID-19: A Global View." Equitable Growth, Finance and Institutions Practice Group Office of the Chief Economist, World Bank Group. Available at http://documents.worldbank.org/curated/en/325021589288466494/pdf/Determinants-of-Social-Distancing-and-Economic-Activity-during-COVID-19-A-Global-View.pdf.

Nguyen, Thuy D., Gupta, Sumedha, Andersen, Martin, Bento, Ana, Simon, Kosali I. & Coady Wing (2020). 'Impacts of state reopening policy on human mobility,' *NBER Working Paper* 27235, DOI 10.3386/w27235.

Ogden, Nick H., et al. (2020). Modelling Scenarios of the Epidemic of COVID-19 in Canada. *Artificial Intelligence in Public Health*, 46–6.

Ontario COVID-19 Science Advisory Table (November 12th 2020). "Update on COVID-19 Projections', available eat https://covid19-sciencetable.ca/wp-content/uploads/2020/11/Update-on-COVID-19-Projections.pdf.

Sen, Anindya (2001). “An Empirical Test of the Offset Hypothesis.", *Journal of Law and Economics* 44(2): Article 6.

Sen, Anindya (2021). “Using Google Data to Estimate the Effects of Regional Mobility on Daily COVID-19 Cases: Evidence from Ontario Public Health Units.”, forthcoming, *Canadian Journal of Regional Science*.

Sheluchin, Anwar, Johnston, Regan M., and van der Linden, Clifton (2020). “Public Responses to Policy Reversals: The Case of Mask Usage in Canada during COVID-19.” *Canadian Public Policy* 46(S2): S119–S126. doi.org/10.3138/cpp.2020-089.

Tolles J, and Luong T. (2020). “Modeling Epidemics with Compartmental Models.” *JAMA* 323(24): 2515–2516. doi:10.1001/jama.2020.8420.

Tuite, Ashleigh R., Fisman David N. and Greer, Amy L. (2020). Mathematical Modelling of COVID-19 Transmission and Mitigation Strategies in the Population of Ontario, Canada. *CMAJ* 192 (19): E497-E505; DOI: https://doi.org/10.1503/cmaj.200476.

Tibshirani, Robert. (1996). “Regression Shrinkage and Selection Via the Lasso.” *Journal of the Royal Statistical Society. Series B (Methodological)* 58(1): 267-288.

Figure 1: Google Social Mobility Data

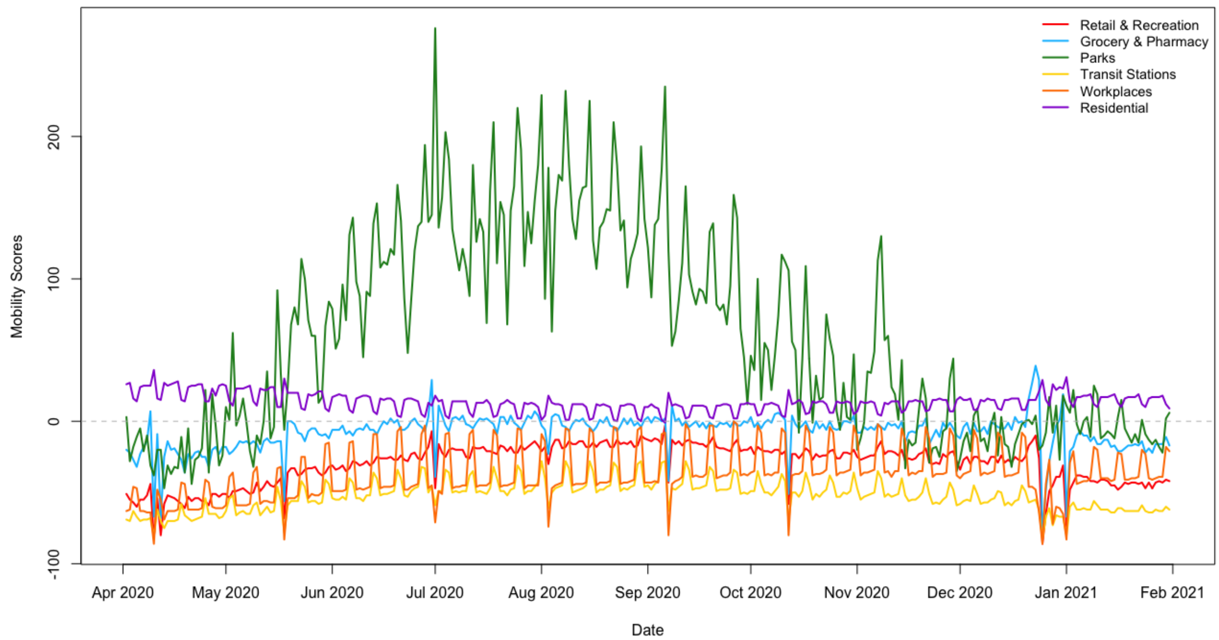

Figure 2: Ontario New Daily COVID-19 Case Counts

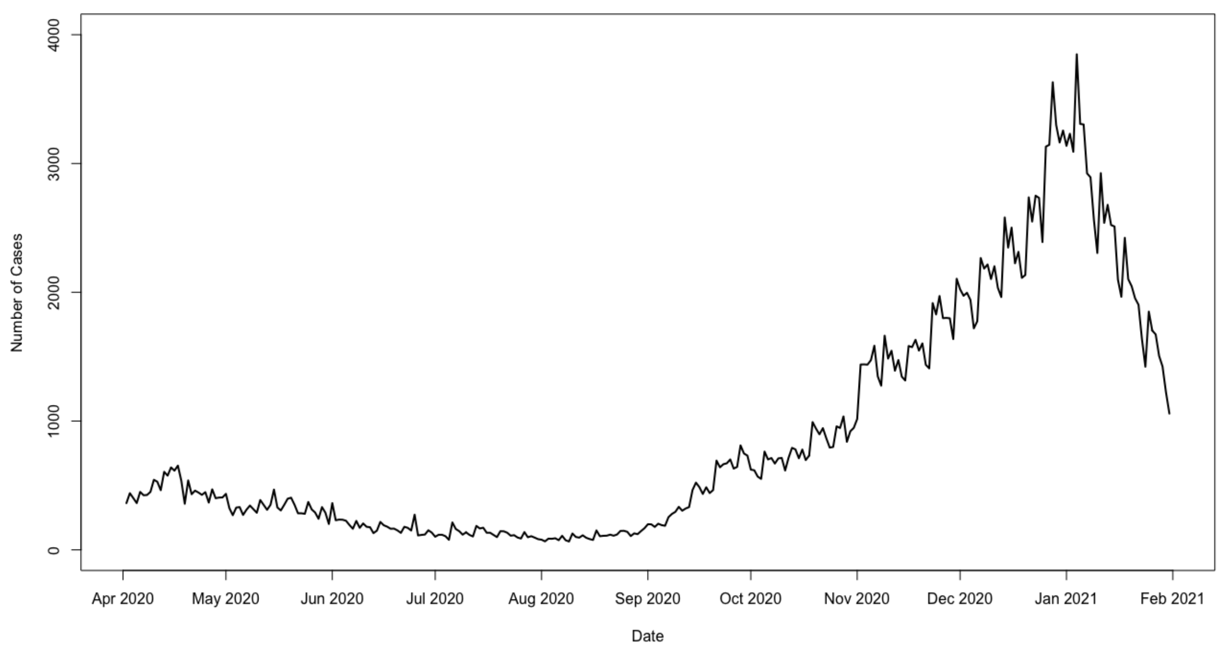

Figure 3. Large Public Health Units & Bank of Canada Policy Stringency Index

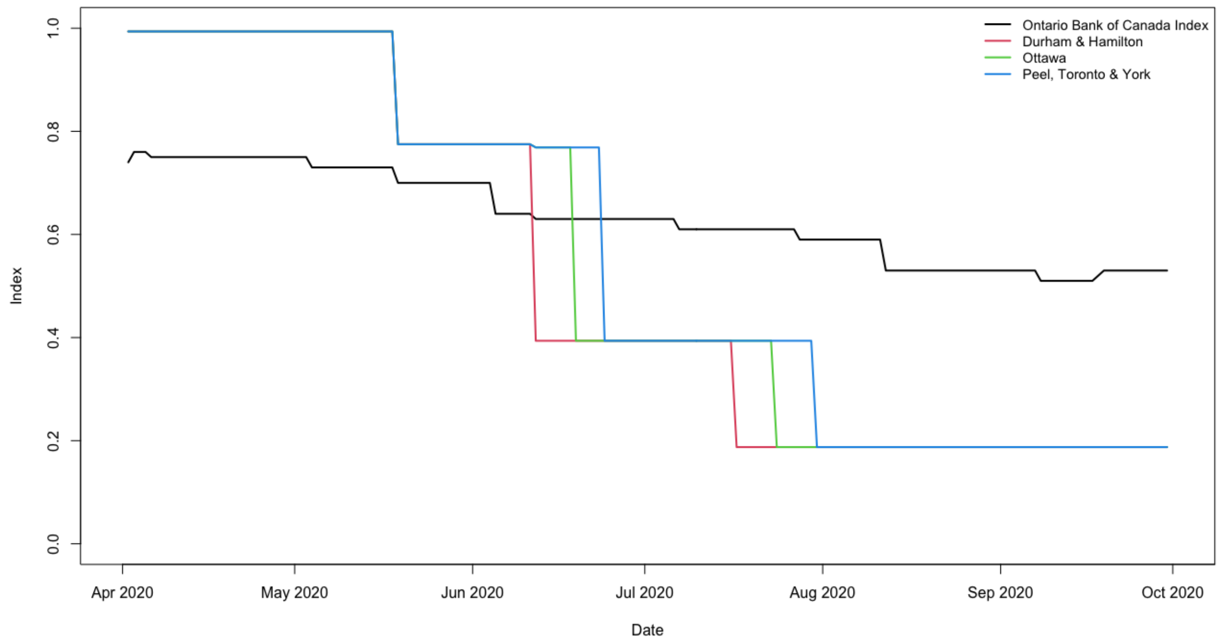

Figure 4. Stringency Index for Smaller Public Health Units

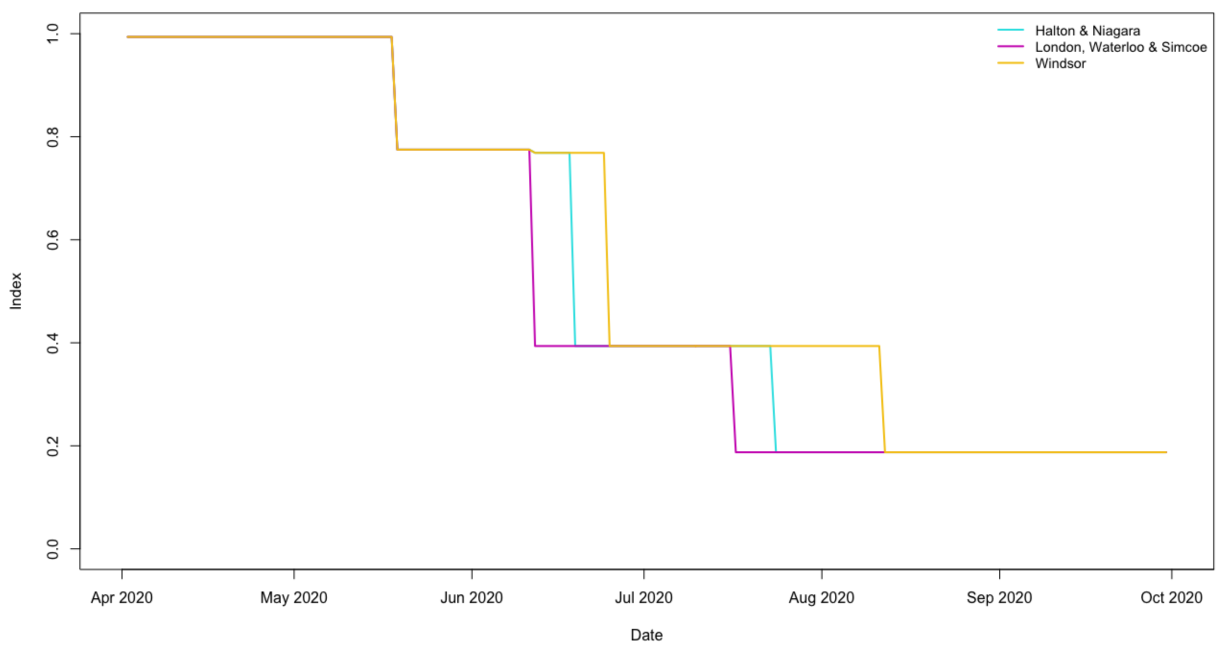

Figure 5: Observed and Forecasted Daily New COVID-19 Case Counts in Ontario

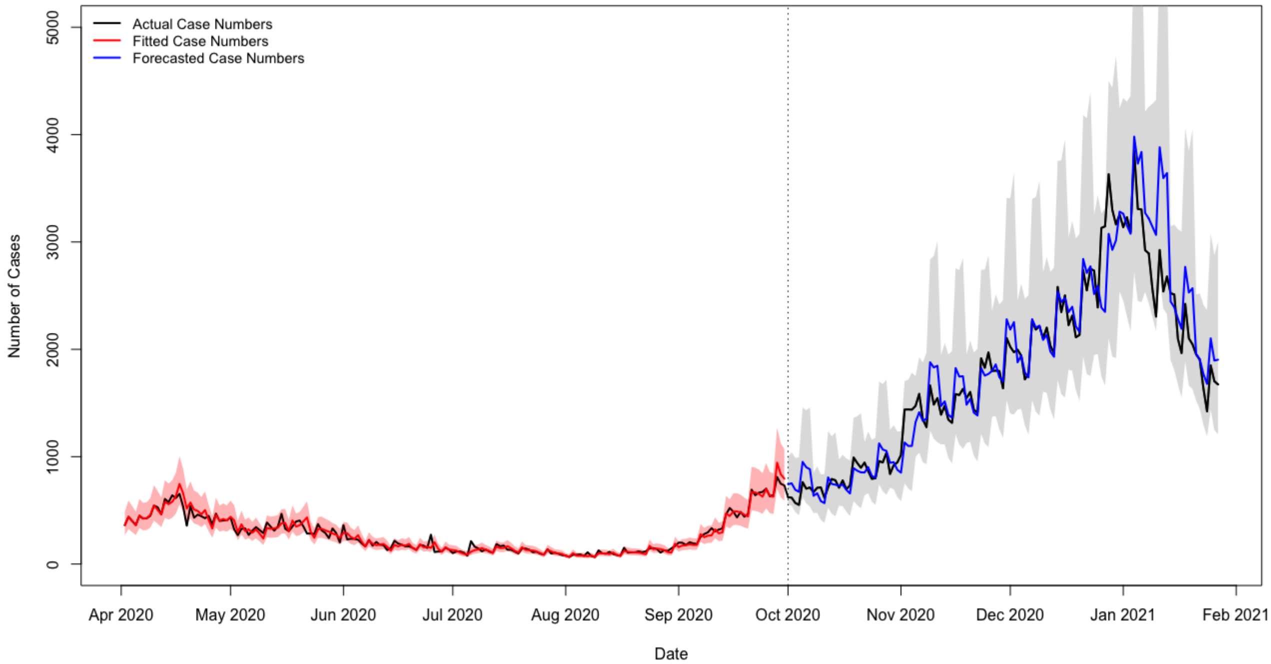

Figure 6. Observed and Forecasted Daily New COVID-19 Case Counts in PHUs

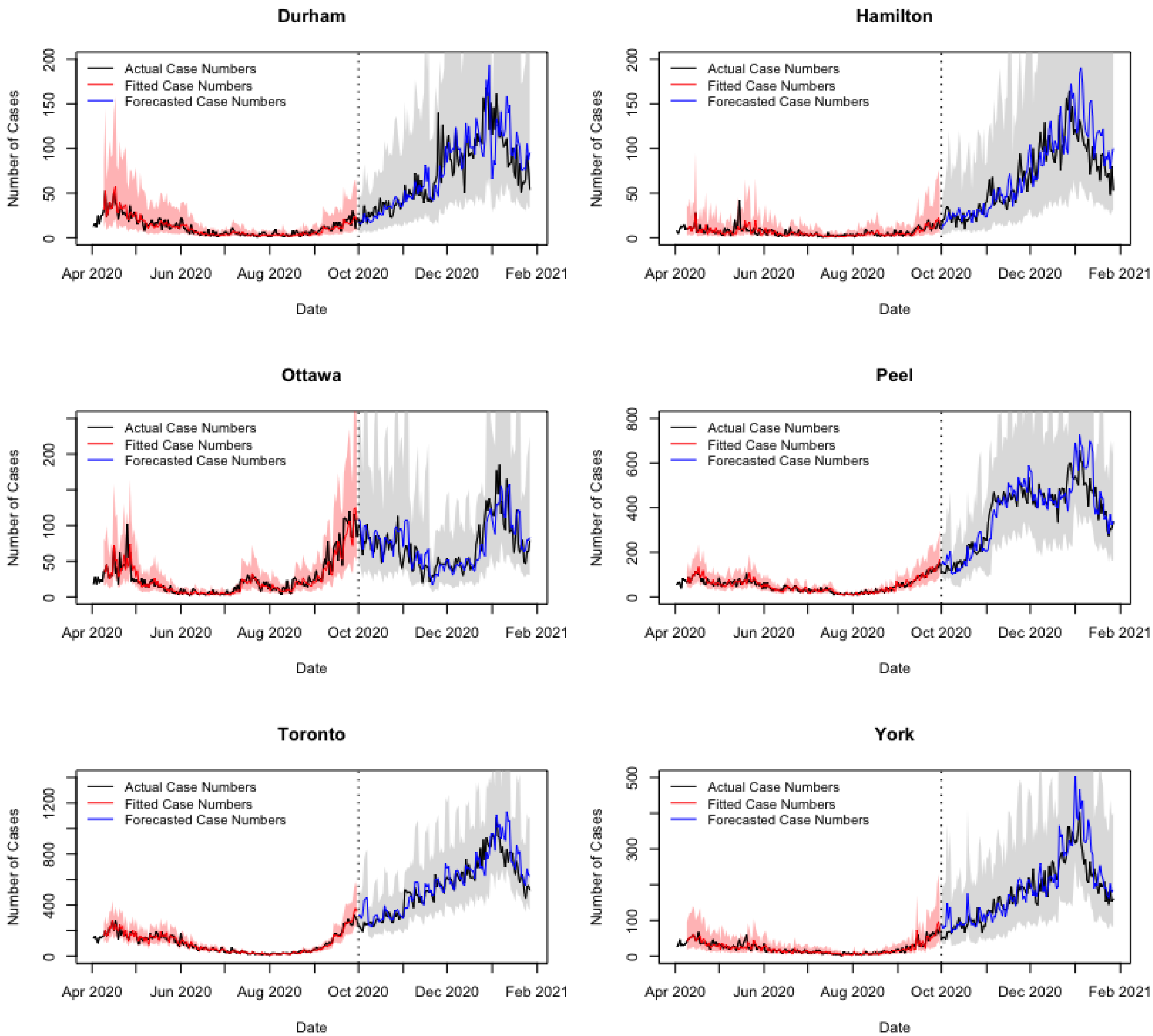

Figure 7. Forecasted Daily New COVID-19 Case Counts by Model (Ontario)

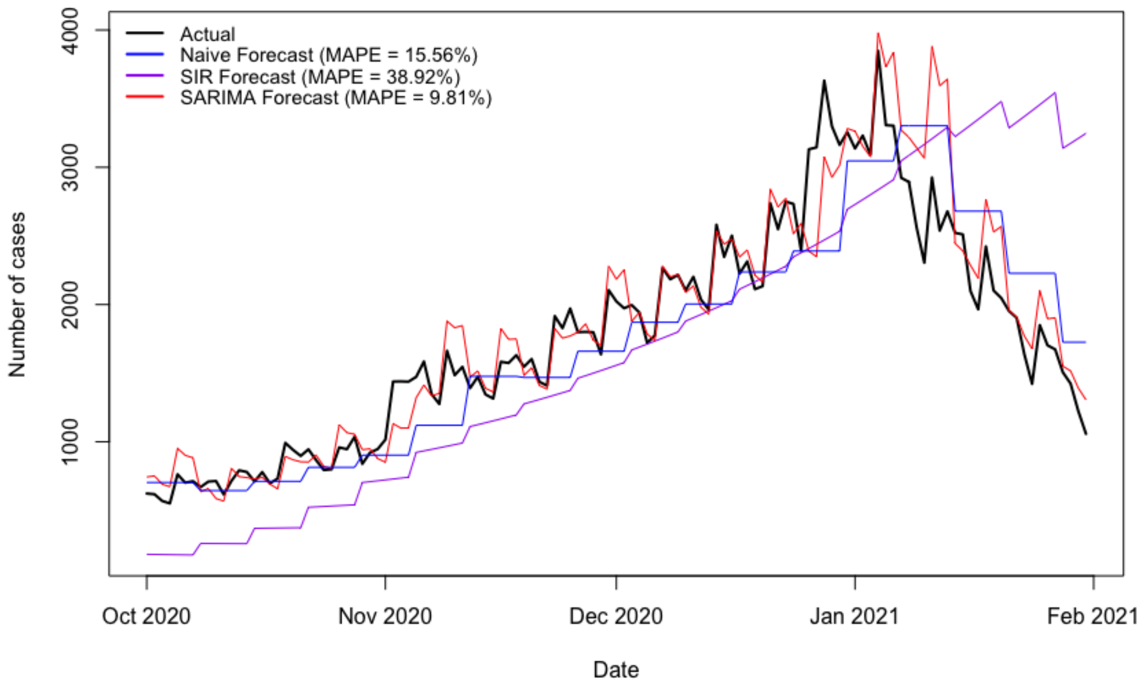

TABLE 1. Summary Statistics April 2nd – Sept 30th 2020

NAME	MEAN	ST. DEV	VARIANCE	MIN	MAX
Total Daily Cases	21.479	37.660	1418.3	0.10000E-08	324.00
7 Day Lag Mask Mandate Dummy	0.39526	0.48902	0.239	0.0000	1.0000
7 Day Lag Policy Stringency Index	0.566	0.33985	0.1155	0.18750	0.99375
7 Day Lag Temperature	16.017	7.6752	58.908	-3.9000	29.500
Tuesday Dummy	0.14208	0.34921	0.122	0.0000	1.0000
Wednesday Dummy	0.1475	0.35473	0.12583	0.0000	1.0000
Thursday Dummy	0.14208	0.34921	0.12195	0.0000	1.0000
Friday Dummy	0.14208	0.34921	0.12195	0.0000	1.0000
Weekend Dummy	0.31694	0.4654	0.2166	0.0000	1.0000
Google Mobility Indicators					
7 Day Lag Retail & Recreation	-30.901	18.834	354.72	-86.000	33.000
7 Day Lag Grocery & Pharmacy	-7.9167	14.985	224.55	-83.000	48.000
7 Day Lag Work	-42.099	19.657	386.40	-89.000	5.0000

TABLE 2 Estimates of the Effects of Non-Pharmaceutical Interventions (NPIs) on Daily Google Mobility across Ontario Public Health Units (PHUs)

	(1) Retail Mobility	(2) Groceries & Pharmacies Mobility	(3) Workplace Mobility
One Day Lagged Dependent Variable	0.191 (0.019)[a]	-0.003 (0.0195)	0.175 (0.020)[a]
Two Day Lagged Dependent Variable	0.346 (0.019)[a]	0.275 (0.019)[a]	-0.572 (0.026)[a]
Local COVID-19 Policy Stringency Index	-20.658 (1.223)[a]	-18.977 (1.457)[a]	-36.646 (1.813)[a]
Mask Mandate Dummy	-2.372 (0.498)[a]	-4.412 (0.741)[a]	-3.705 (0.940) [a]
Average Daily Temperature	0.139 (0.026)[a]	0.237 (0.038)[a]	0.356 (0.046)[a]
PHU Dummies	Yes	Yes	Yes
Day of Week Dummies	Yes	Yes	Yes
Adjusted R Square	0.8994	0.6418	0.7185

Notes: The estimates in this table are based on data from 12 Public Health Units (PHUs) between April 2nd– September 30th 2020. The dependent variables are different Google mobility variables. Regression estimates are obtained from Weighted Least Squares (WLS) regression where observations are weighted by PHU specific population. Standard errors are in parentheses below coefficient estimates. a, b, and c denote statistical significance at the 1%, 5%, and 10% levels.

TABLE 3 Estimates of the Effects of Non-Pharmaceutical Interventions (NPIs) on Daily COVID-19 Cases & Google Mobility across Ontario Public Health Units (PHUs)

	(1) WLS	(2) WLS	(3) WLS	(4) IV
One Day Lagged COVID-19 Policy				
Seven Day Lagged Local COVID-19 Policy Stringency Index	-2.313 (0.880)[a]	0.178 (3.084)	-1.746 (3.030)	-53.427 (21.77)[b]
Seven Day Lagged Mask Mandate Dummy Variable		-0.664 (1.177)	0.037 (1.154)	-12.255 (5.004)[b]
One Day Lagged Cases	0.629 (0.020)[a]	0.611 (0.020)[a]	0.549 (0.022)[a]	0.554 (0.023)[a]
Two Day Lagged Cases	0.363 (0.021)[a]	0.355 (0.021)[a]	0.234 (0.025)[a]	0.400 (0.023)[b]
Three Day Lagged Cases			0.177 (0.026)[a]	
Four Day Lagged Cases			-0.027 (0.026)	
Five Day Lagged Cases			0.042 (0.026)	
Six Day Lagged Cases			-0.082 (0.025)[a]	
Seven Day Lagged Cases			0.102 (0.022)[a]	
Seven Day Lagged Retail Mobility		0.281 (0.087)[a]	0.206 (0.086)[b]	-0.801 (0.378)[b]
Seven Day Lagged Grocery Mobility		-0.26 (0.063)[a]	-0.202 (0.062)[a]	0.384 (0.203)[c]
Seven Day Lagged Work Mobility		0.102 (0.027)[a]	0.122 (0.027)[a]	0.0598 (0.022)[a]
Average Daily Temperature		-0.396 (0.068)[a]	-0.362 (0.066)[a]	-0.534 (0.140) [a]
F Statistic (*P* Value) of Joint Significance of Instruments (12, 13, 14, and 15 day lags of daily COVID-19 cases)				13.763 (p value = 0.000)
Sargan Test for Overidentifying Restrictions				1.142 (p value = 0.331)
PHU Dummies	Yes	Yes	Yes	Yes
Day of Week Dummies	Yes	Yes	Yes	Yes
Adjusted R Square	0.9451	0.9466	0.9489	0.9151

Notes: The regressions in this table are based on data from 12 Public Health Units (PHUs) between April 2nd– September 30th 2020. The dependent variable in is the total number of daily cases. Regression estimates in columns (1) - (3) are obtained from Weighted Least Squares (WLS) regression where observations are weighted by PHU specific population while columns (4) contains Instrumental Variables (IV) estimates where the Seven Day Lagged Local COVID-19 Policy Stringency Index is instrumented by 12, 13, 14, and 15 day lags of daily COVID-19 cases. Standard errors are in parentheses below coefficient estimates. a, b, and c denote statistical significance at the 1%, 5%, and 10% levels.

TABLE 4 Comparison and Evaluation of PHU Forecasting Models (October 1st 2020-January 31st 2021)

Jurisdiction	Model	AICC	MAE	MAPE
Ontario (aggregate)	SARIMA	-145.45	172.89	0.0974
	SARIMA + STR	-130.74	193.66	0.1068
	SARIMA + GM	-129.81	187.17	0.1029
	SARIMA + GM +STR	-127.55	198.43	0.1091
Durham	SARIMA	282.77	14.93	0.2105
	SARIMA + STR	273.32	14.22	0.2054
	SARIMA + GM	272.00	16.24	0.2342
	SARIMA + GM +STR	274.14	15.65	0.2231
Hamilton	SARIMA	372.65	20.17	0.3167
	SARIMA + STR	300.80	16.88	0.2688
	SARIMA + GM	334.04	17.70	0.2852
	SARIMA + GM +STR	334.91	17.00	0.2727
Ottawa	SARIMA	269.75	17.41	0.2510
	SARIMA + STR	206.14	18.65	0.2857
	SARIMA + GM	205.74	15.83	0.2418
	SARIMA + GM +STR	208.04	16.28	0.2503
Peel	SARIMA	82.41	48.55	0.1468
	SARIMA + STR	74.23	51.32	0.1566
	SARIMA + GM	74.87	49.14	0.1451
	SARIMA + GM +STR	75.27	49.19	0.1485
Toronto	SARIMA	-29.7	71.11	0.1415
	SARIMA + STR	-14.14	69.19	0.1367
	SARIMA + GM	-14.02	66.23	0.1347
	SARIMA + GM +STR	-11.88	67.77	0.1379
York	SARIMA	176.89	28.01	0.1622
	SARIMA + STR	174.05	25.89	0.1559
	SARIMA + GM	209.5	32.3	0.2046
	SARIMA + GM +STR	210.55	31.7	0.1970

Notes: STR = Bank of Canada COVID-19 Policy Stringency Index, GM = Google Mobility

Appendix

The most general SARIMA model we consider is

$$\log(y_t) = \beta_1 \text{Retail}_{t-7} + \beta_2 \text{Grocery}_{t-7} + \beta_3 \text{Parks}_{t-7} + \beta_4 \text{Workplace}_{t-7}$$
$$+ \beta_5 \text{Transit}_{t-7} + \beta_6 \text{Residential}_{t-7}$$
$$+ \text{BOC Policy Stringency Index}_{t-7} + \eta_t$$

where $\eta_t \sim SARIMA(p, d, q)(P, D, Q)[7]$. Therefore $\log(y_t)$ (where y_t is the number of new COVID-19 cases on day t) is modeled by a SARIMA model with a given specification of p, d, q, P, D, Q and as a function of the seven-day lags of the six Google mobility variables as well as the BOC stringency index. Different values of the non-seasonal and seasonal orders p, d, q, P, D, Q give rise to different configurations of the model, accounting for different forms of correlation structure in daily case numbers. Note that we specify the seasonal component with a 7-day period (reflecting the weekly seasonality observed in Figure 2), and the values of p, d, q, P, D, Q are chosen to minimize the corrected Akaike Information Criterion (AIC) to ensure the model fits the observed data well (Cavanaugh, 1997). $\log(y_t)$ is taken to be the dependent variable in this forecasting model because the natural-log transformation takes into account the heteroscedasticity observed in daily COVID-19 case counts during the forecasting period.

We also consider sub-models that: exclude all exogenous information; include only the Google mobility variables; include only the BOC stringency index. We compare all four specifications in terms of their predictive accuracy which we evaluate using cross validation. We calculate 7-day forecasts and re-estimate the model and update the orders p, d, q, P, D, Q (if necessary) before forecasting the subsequent 7 days. Updating the parameter estimates and model orders serves to dynamically adapt the model as new data become available; it acknowledges the progression of the disease may change over time and so we would not expect the model derived during the training period to be relevant indefinitely. Empirical investigations indicate that updating the model less frequently (i.e., every 4 weeks) is not often enough to adequately react to rapid changes in the spread of COVID-19. On the other hand, updating too frequently risks needlessly reacting to noise. Updating weekly appears to balance these concerns and yields strong predictive performance.