Monitoring the COVID-19 epidemic with nationwide telecommunication data

Joel Persson[a,1], Jurriaan F. Parie[a], and Stefan Feuerriegel[a]

[a]Department of Management, Technology, and Economics, ETH Zurich (Swiss Federal Institute of Technology), 8092 Zurich, Switzerland

Edited by Amy Wesolowski, Johns Hopkins University, Baltimore, MD, and accepted by Editorial Board Member Adrian E. Raftery May 7, 2021 (received for review January 13, 2021)

In response to the novel coronavirus disease (COVID-19), governments have introduced severe policy measures with substantial effects on human behavior. Here, we perform a large-scale, spatiotemporal analysis of human mobility during the COVID-19 epidemic. We derive human mobility from anonymized, aggregated telecommunication data in a nationwide setting (Switzerland; 10 February to 26 April 2020), consisting of ∼1.5 billion trips. In comparison to the same time period from 2019, human movement in Switzerland dropped by 49.1%. The strongest reduction is linked to bans on gatherings of more than five people, which are estimated to have decreased mobility by 24.9%, followed by venue closures (stores, restaurants, and bars) and school closures. As such, human mobility at a given day predicts reported cases 7 to 13 d ahead. A 1% reduction in human mobility predicts a 0.88 to 1.11% reduction in daily reported COVID-19 cases. When managing epidemics, monitoring human mobility via telecommunication data can support public decision makers in two ways. First, it helps in assessing policy impact; second, it provides a scalable tool for near real-time epidemic surveillance, thereby enabling evidence-based policies.

COVID-19 | epidemiology | human mobility | telecommunication data | Bayesian modeling

The novel coronavirus disease (COVID-19) has evolved into a global pandemic, which, as of 15 December 2020, has been responsible for more than 70 million reported cases (1). In response, governments around the world have put policy measures into effect with the aim of reducing transmission rates (2–7). Examples of policy measures are border closures, school closures, venue closures, and bans on gatherings.

Prior literature has suggested the use of human mobility data to model the COVID-19 epidemic (8). Mobility patterns have been inferred from point-of-interest (POI) check-ins (9–13) and from location logs of smartphone apps (14–26). Other works have used telecommunication data to model spreading patterns (27, 28), for exploratory analysis of mobility patterns (29–31), for network analysis of structural changes in mobility (32), and for modeling the spatiotemporal distribution of COVID-19 (28), but none have yet empirically explored the link between telecommunication data and policy measures. Establishing such a link would provide a scalable tool for near real-time disease surveillance under policy measures and, in particular, enable evidence-based policies. Previously, the value of telecommunication data for disease surveillance has been studied in the context of malaria (33, 34), influenza (35), and other infectious diseases (36–38), where the objective was to make spatiotemporal forecasts. In contrast, this paper demonstrates the utility of telecommunication data for near real-time assessments of COVID-19 policies. In fact, nationwide data from mobile telecommunication networks have been used by governments during the first wave of COVID-19 (39). However, to the best of our knowledge, empirical evidence regarding the effectiveness of telecommunication data for epidemic surveillance in the context of COVID-19 is absent.

In this paper, we analyze human mobility during the COVID-19 epidemic. Our analysis is based on large-scale, granular data of human movements (anonymized and aggregated) consisting of ∼1.5 billion trips in Switzerland during the first COVID-19 wave (10 February to 26 April 2020) derived from telecommunication data. Using regression models, we estimate 1) the impact of policy measures on human mobility and 2) how mobility predicts the growth in reported COVID-19 cases. By establishing that policy measures reduce mobility and that mobility predicts reported cases, mobility insights can be used to inform when to implement policy measures. The findings are therefore of direct value to public decision makers: Monitoring human mobility through telecommunication data provides an effective and scalable tool for near real-time epidemiology and, thus, management of the COVID-19 epidemic.

To establish the ability of telecommunication data for near real-time monitoring of the COVID-19 epidemic, we follow a two-stage approach (*Materials and Methods*). We first study the reduction in mobility due to five different policy measures (bans on gatherings of more than 100 people, bans on gatherings of more than 5 people, school closures, venue closures, and border closures). We then estimate the extent to which reduction in mobility predicts decreases in reported case growth. Here, we compare the predictive ability over a forecast window from 7 to 13 d. Taken together, the results confirm the effectiveness of policy measures for reducing human mobility and, in turn, human mobility as a predictor of reported cases by a lead time of approximately (approx.) 7 to 13 d. The two-stage approach is repeated for total trips, three different modes of mobility (train, road, highway), and two different purposes for mobility (commuters vs. noncommuters). In an extended analysis, we further perform a mediation analysis. Here, we decompose the reduction in new reported cases due to the policy measures into 1) the

Significance

To manage the current epidemic, policymakers need tools that help them in evidence-based decision making. In particular, decision support is needed to assess policy measures by their ability to enforce social distancing. A solution is offered by our work: We use mobility data derived from telecommunication metadata as a proxy for social distancing, and, based on this, we demonstrate how the effect of policy measures can be monitored in a nationwide setting. Compared to the status quo, this provides a clear benefit: Monitoring policy measures through case counts has a substantial time lag, whereas our approach allows for monitoring in near real time.

Author contributions: J.P., J.F.P., and S.F. designed research; J.P., J.F.P., and S.F. performed research; J.P. and S.F. analyzed data (modeling); J.F.P. analyzed data (exploratory); and J.P., J.F.P., and S.F. wrote the paper.

This article is a PNAS Direct Submission. A.W. is a guest editor invited by the Editorial Board.

[1]To whom correspondence may be addressed. Email: jpersson@ethz.ch.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2100664118/-/DCSupplemental.

Published June 23, 2021.

part that is explained only by reductions in mobility and 2) the part that is explained by other behavioral adaptations.

Results

We analyze large-scale data on human mobility during 10 February to 26 April 2020 from Switzerland. For this, human movements derived from telecommunication data were obtained from a major telecommunications provider in Switzerland (*Materials and Methods*). Telecommunication data provide more reliable and extensive information on mobility compared to alternative data sources (check-ins or location logs from smartphone apps) (8, 40, 41). In particular, our telecommunication data represent routine signal exchanges ("pings") exchanged between mobile devices and network antennas. These were recorded for all mobile devices in Switzerland regardless of the mobile service provider. Based on the telecommunication data, granular locations (longitude, latitude) of individuals carrying a mobile device were inferred. This yields data on microlevel movements from all mobile devices in a nationwide setting. Altogether, the nationwide mobility for a population of ∼8.6 million people was estimated.

For the analysis, the telecommunication data were then processed to count the number of trips between postcodes per day and canton. All trips were further classified according to the mode of mobility (train, highway, and other road movements) and purpose (commuting vs noncommuting). For the time period of this study (10 February to 26 April 2020), our data include a total of ∼1.5 billion trips. Details are reported in *Materials and Methods*.

We further collected data on the use of policy measures in Switzerland. Switzerland comprises 26 member states at the subnational level (called "cantons"), each with a large degree of sovereignty. As a result, the use of policy measures varies across cantons with respect to their order and timing in a way that is similar to the variation among other European countries. Some cantons (e.g., Ticino at the border to Italy and Geneva at the border to France) showed epidemiological dynamics with large numbers of reported cases and responded with comparatively stringent policy measures during the first wave of COVID-19. Other cantons had lower case numbers and put policy measures into effect during a later phase of the epidemic. We followed a systematic procedure (*SI Appendix*, section 1) based on which we encoded policy measures according to five categories: bans on gatherings (>100 people), bans on gatherings (>5 people), school closures, venue closures (stores, restaurants, bars), and border closures. The implementation dates of the chosen policy measures varied greatly across cantons (but then remained in effect for the complete study period, i.e., until 10 April).

The spatiotemporal patterns of human movements in our data are as follows. Overall, ∼95 million trips were recorded in the first week (24 February to 1 March 2020), during which all five policy measures were in effect in all 26 cantons. In comparison, the same time period in 2019 (as a reference period) registered ∼186 million trips. This amounts to a reduction of 49.1%. The reduction occurred in all cantons (Fig. 1 *A* and *B*). The highest decline was observed in Ticino and Geneva, which are located at the borders with Italy and France, respectively. Both cantons also reported the highest number of COVID-19 cases.

The largest mobility reduction compared to 2019 occurred on Sunday, 22 March 2020 (Fig. 1*C*). In comparison to Sunday, 24 March 2019, the reduction in trip counts ranged between 49.3 and 77.0% across the 26 cantons (mean: 61.6% reduction per canton). Overall, the reduction in mobility is of similar magnitude for both rural (e.g., canton of Valais) and urban regions (e.g., cantons of Basel-City and Zurich). Furthermore, movements declined for all modes of mobility (Fig. 1*D*) and for all purposes (Fig. 1*E*). After the implementation of the policy measures, trips by train remained low for the rest of the study period, while highway traffic was on an upward trend (Fig. 1*D*). Similarly, trips by commuters remained at a low level after the implementation of the policy measures, whereas trips not for commuting (i.e., personal purposes) started increasing in early April (Fig. 1*E*).

Estimating the Reduction in Human Mobility Due to Policy Measures. We estimate the reduction in mobility due to the policy measures with a regression model. The estimates are identified via a difference-in-difference analysis and may thus be given a causal interpretation under certain assumptions (*SI Appendix*, section 4A).

The most effective policies for reducing trip counts are as follows (Fig. 2*A*). Based on our model, bans on gatherings of more than five people reduced total trips by 24.9% (95% credible interval [CrI]: 22.1 to 27.6%), venue closures reduced total trips by 22.3% (95% CrI: 15.6 to 29.0%), and school closures reduced total trips by 21.6% (95% CrI: 17.9 to 25.0%). For a precise ranking, the width of the credible intervals must be considered. Here, the aforementioned policy measures appear more effective at reducing total trips than the other two policy measures (i.e., bans on gatherings of more than 100 people and border closures). In particular, bans on gatherings of more than 100 people are linked to a comparatively smaller change in total trips than bans on gatherings of more than 5 people (i.e., the 95% CrIs of the estimates are disjoint). For border closures, the credible interval includes zero. Overall, policy measures are important determinants of mobility reductions during the COVID-19 epidemic.

The estimated mobility reduction depends on the underlying mode (Fig. 2*B*). Across all policy measures, the mobility reduction is more pronounced for highways than for road movements. This observation is to be expected. Highways are often used for long-distance travel, which is more likely to be suspended during an epidemic, while roads also include movements within close proximity and are more likely to correspond to routine or essential activities (e. g., grocery shopping). For the ban on gatherings and school closures, the largest reduction is seen in trips by train, which can be explained by the widespread use of public transportation in Switzerland. Finally, we observe a wide credible interval for the estimated effect of venue closures on trips by train. A potential reason for this is that the use of trains (e.g., for visiting stores) varies across cantons, as some cantons (e.g., Zurich) have a high population density with extensive shopping infrastructure, while others (e.g., Appenzell Innerrhoden) have only a few stores due to their low population density, resulting in the need for travel to visit stores.

The estimated effect sizes are fairly similar for trips made for commuting versus noncommuting (Fig. 2*C*). This is interesting, considering that no policy measure in Switzerland directly prohibited movement to and from work. The efficacy of border closures is uncertain since the credible interval for its estimated effect includes zero. In contrast, a negative effect is observed for commuting. Here, one potential reason is that border closures have reduced the number of cross-border commuters. A validation analysis supports this explanation (see *SI Appendix*, section 2B for details).

The findings are robust to alternative model specifications (see the robustness checks in *SI Appendix*, section 6). Specifically, changing the specification of time-related control variables still gives parameter estimates for the policy measures that imply decreases in mobility.

Estimating the Relationship between Mobility and COVID-19 Cases. The epidemic dynamics during the first wave of COVID-19 in Switzerland are as follows. The initial exponential growth rates

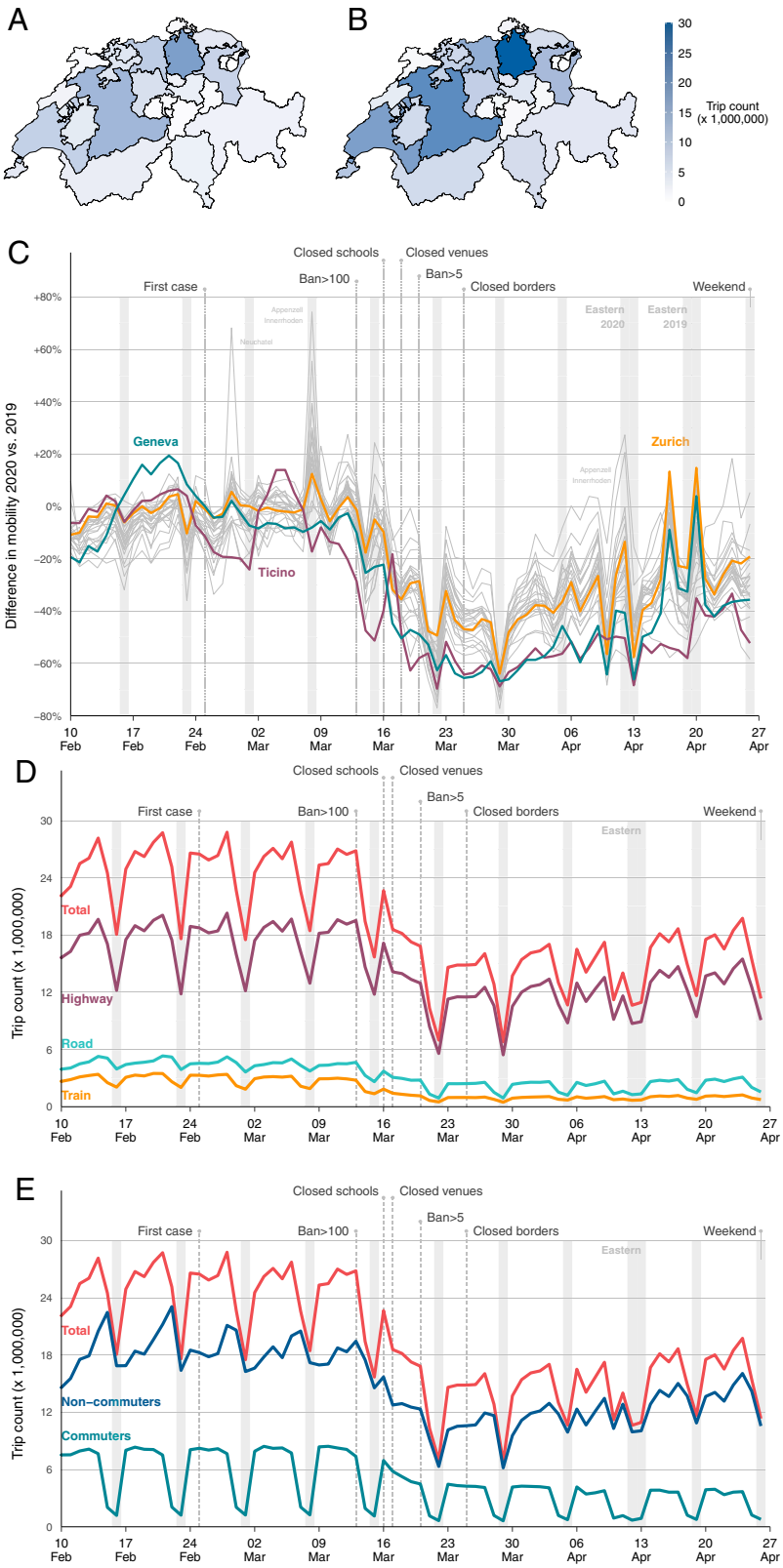

Fig. 1. Nationwide human mobility in Switzerland during the first wave of the COVID-19 epidemic. Mobility is quantified by movements ("trips") between different postcode areas. (*A*) Total number of trips per canton for the first week after all policy measures were put in effect in all cantons (25 March to 1 April 2020). (*B*) Total number of trips for the same week in 2019 (i.e., as reference year). For this week, the total number of trips dropped from ∼186 million (in 2019) to ∼95 million (in 2020), i.e., a reduction of 49.1%. (*C*) Percentage change in total trips across 26 subnational levels (cantons) for 2020 vs. 2019 (when aligned for day-of-week patterns). The reason for the comparison to 2019 is to show the reduction in mobility relative to a reference year, while accounting for seasonal changes in mobility. A higher reduction in mobility is observed for cantons that also reported a high number of COVID-19 cases (i.e., Ticino and Geneva). (*D*) Reduction in trip count by mode of mobility. (*E*) Reduction in trip count by purpose of mobility. Annotations show nationwide implementation dates of policy measures (implementation dates at cantonal level are reported in *SI Appendix*, section 1).

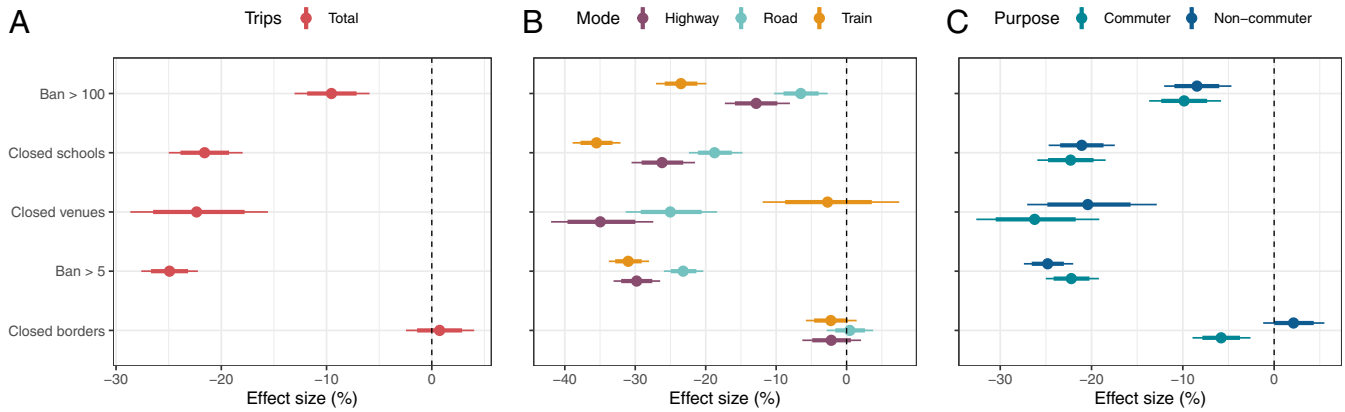

Fig. 2. The estimated effects of policy measures on mobility. Shown are the estimated effects on (*A*) the total number of trips, (*B*) trips by mode, and (*C*) trips by purpose. The dots in *A*–*C* show posterior means; the thick and thin bars represent the 80 and 95% credible intervals, respectively. Policy measures are arranged from top to bottom in the order in which they were implemented (cf. *SI Appendix*, section 1).

exhibit considerable heterogeneity across cantons (Fig. 3*A*). The strongest initial growth is observed for the cantons Ticino and Geneva, resulting in the largest number of cases toward the end of the sample. Moreover, the number of reported cases at the dates that policy measures were implemented varies greatly across cantons (Fig. 3*B*). This reflects different responses among cantons to local infection dynamics.

We use regression models to estimate the extent to which decreases in mobility predict future reductions in the reported number of new cases (Fig. 3*C*). The predicted decrease is studied with a forecast window over 7 to 13 d. The forecast window is set analogous to previous research (26) and so that it covers variations in incubation time combined with reporting delay.

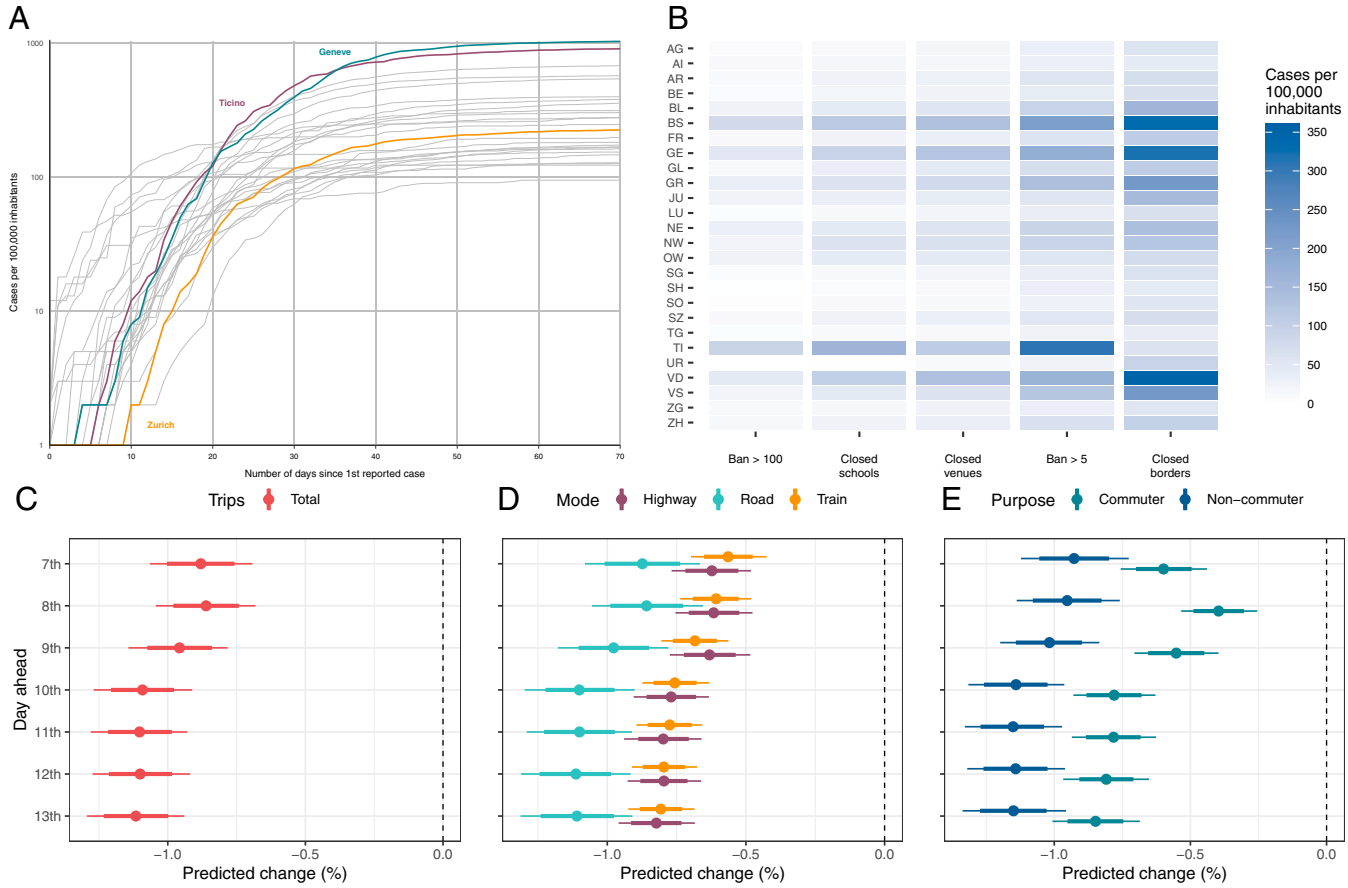

Fig. 3. Decreases in human mobility predict reductions in reported cases of COVID-19 over a forecast window of 7 to 13 d. (*A*) COVID-19 case growth since the first reported case for 26 Swiss cantons. (*B*) Number of COVID-19 cases when the policy measures were implemented for all 26 Swiss cantons. Shown are abbreviated names from https://www.bfs.admin.ch/bfs/en/home/basics/symbols-abbreviations.html. The predicted change in reported new COVID-19 cases at a given day based on mobility lagged by 7 to 13 d. (*C*–*E*) The predicted change is reported given a 1% decrease in (*C*) total trips, (*D*) mode, and (*E*) purpose. Posterior means are shown as dots, while 80 and 95% credible intervals are shown as thick and thin bars, respectively.

We find that decreases in mobility at a given day predict decreases in reported new cases 7 to 13 d later. For a 7-d ahead forecast, we find that a 1% decrease in the total number of trips predicts a 0.88% (95% CrI: 0.7 to 1.1%) reduction in the reported number of new cases. For a 13-d ahead forecast, a 1% decrease in the total number of trips predicts a 1.11% (95% CrI: 0.9 to 1.6%) reduction in the reported number of new cases. Overall, mobility predicts decreases in the reported number of new cases over the whole forecast horizon. The predicted decrease is larger for longer forecasts. This result is to be expected, as a longer time window accommodates the full distribution of incubation periods (plus reporting delays). Altogether, the regression analysis provides evidence that mobility predicts epidemic dynamics.

Our analysis also shows that the predicted change in the reported number of new cases varies across the mode and purpose of trips. In terms of mode, decreases in trips by highway and train predict reductions in the reported number of new cases of similar magnitude (Fig. 3*D*). Their estimates have comparatively narrow credible intervals, reflecting a higher degree of certainty. Trips are also categorized according to their purpose, namely commuting vs. noncommuting. The results show that decreases in trips for commuting predict smaller reductions in the number of reported new cases compared to decreases in noncommuting trips (Fig. 3*E*). Predicted reductions are nonetheless found for both modes of mobility (i.e., commuting vs. noncommuting), for all categories of purpose (i.e., highway, road, and train), and for the whole forecast window of 7 to 13 d. Again, a larger reduction is predicted for longer forecasts.

The predictive ability of mobility for reported new cases holds with alternative model specifications. For most of the 7- to 13-d forecasts, changing how we control for time-related factors still results in a predicted decrease in reported cases given decreases in mobility. Moreover, changing the dependent variable to daily hospitalizations or deaths attributed to COVID-19 leads to qualitatively the same results over a forecast horizon of 10 to 20 d. Details on these robustness checks are provided in *SI Appendix*, section 6.

Estimating the Mediating Role of Mobility. In an extended analysis, we study how decreases in reported case growth are explained by reductions in mobility due to policy measures versus other behavioral changes due to policy measures. The estimates are obtained from a mediation analysis that decomposes the total effects of the policy measures on reported case growth into 1) their direct effects not explained by changes in mobility and 2) their indirect effects through mobility. The mediation analysis is performed by combining our two regression models into a structural equation model (see *SI Appendix*, section 9 for details). Results from mediation analysis are reported for the total number of trips.

The mediation analysis shows a large direct effect for bans on gatherings of more than 5 people, bans on gatherings of more than 100 people, and school closures (Fig. 4). Pronounced indirect effects are found for all policy measures. In particular, the indirect effect of venue closures makes up about one-third of their total effect at several lags. Moreover, border closures are estimated to have reduced the reported number of new cases only indirectly through mobility. The results are discussed in further detail in *SI Appendix*, section 9C. In summary, the results show that mobility is an important mediator: The studied policy measures operate—to a large degree—through mobility. Thus, policy measures aimed at reducing mobility appear to be effective for reducing COVID-19 case growth.

Discussion

This study shows the ability of telecommunication data for near real-time monitoring of the COVID-19 epidemic. Our analysis is based on nationwide telecommunication data during 10 February to 26 April 2020 from Switzerland, which were used to infer nationwide mobility patterns. This supports monitoring of the COVID-19 epidemic as follows: 1) We first studied the link between policy measures and human mobility. In particular, we performed a difference-in-difference analysis quantifying how mobility was reduced due to five different policy measures (bans on gatherings, school closures, venue closures, and border closures). The largest reduction in total trips was linked to bans on gatherings of more than five people, followed by venue closures and school closures. Overall, the policy measures resulted in substantial reductions of human mobility. 2) We then studied the link between human mobility and reported COVID-19 cases. Reductions in mobility predicted decreases in the number of reported new cases. Specifically, a reduction in human movement by 1% predicted a 0.88 to 1.11% reduction in the daily number of new cases of COVID-19 over a forecast horizon of 7 to 13 d. Our modeling approach with telecommunication data therefore provides near real-time insights for disease surveillance. Taken together, the findings enable quantitative comparisons of the extent to which policy measures reduce mobility and, subsequently, reduce reported cases of COVID-19.

The use of telecommunication data for nationwide monitoring has several benefits (8, 42). First, telecommunication data from mobile networks provide comprehensive coverage. Specifically, such data capture all movements of individuals carrying mobile devices without explicit user interaction, including those from nonresidential and even foreign individuals. Mobile devices routinely exchange information when searching for signals from adjacent antennas; hence, metadata are retrieved regardless of the underlying mobile service provider. Second, such metadata can be collected in an anonymized manner that is compatible with data privacy laws. Third, movements at a microlevel (e.g., trips to other households, school, and work) can be inferred. Thus, compared to alternative sources of mobility information such as check-ins or smartphone apps, telecommunication data are considered to be more complete (8, 40, 41). Fourth, unlike smartphone apps, telecommunication data are also available in low-income countries (43). Finally, telecommunication data are measured with high frequency (e.g., daily), thereby enabling regularly updated monitoring as needed by decision makers. Based on these benefits, telecommunication data appear to be highly effective for policy monitoring during the COVID-19 epidemic.

This work is subject to the typical limitations of observational studies. First, the findings depend on the accuracy of the data on COVID-19 cases. Second, our models are informed by recommendations for COVID-19 modeling (44) and, therefore, follow parsimonious specifications to isolate features of the epidemic for policy-relevant insights. We cannot, however, rule out the possibility that there exist external factors beyond those that are captured by the spatially and temporally varying variables in the models. To address this, we use flexible models and conduct extensive robustness checks (*SI Appendix*, section 6). Third, the model linking policy measures to mobility estimates effects, while the model linking mobility to cases is predictive. The different objectives of the models address the needs of public decision makers: The former serves policy assessments and the latter epidemiological forecasting, respectively. Therefore, the estimates from the former are identified with a difference-in-difference analysis and may thus warrant causal interpretations under certain assumptions (*SI Appendix*, section 4A). On the other hand, the estimates from the latter are conditional associations since the model does not control for that policy measures reduce both mobility and cases. Therefore, the latter model predicts reductions in reported cases from reductions in mobility when both reductions are driven by policy measures

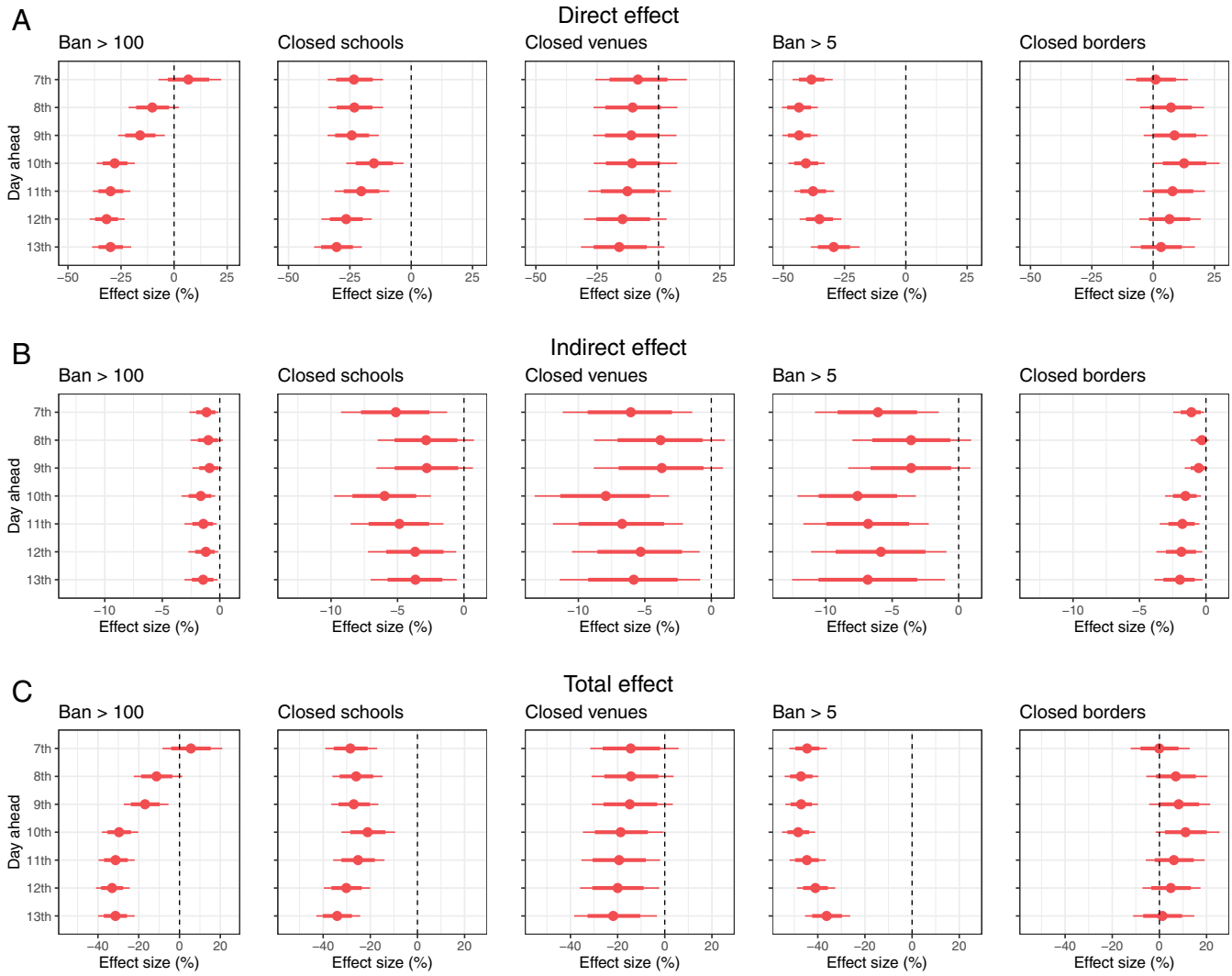

Fig. 4. Mobility mediates the effect of policy measures on the reported number of new cases. (*A*–*C*) Estimated (*A*) direct effect of policy measures, (*B*) indirect effect of policy measures via total trips, and (*C*) total effect of policy measures on the 7th to the 13th d ahead. Posterior means are shown as dots, while 80 and 95% credible intervals are shown as thick and thin bars, respectively. Policy measures are arranged from top to bottom in the order in which they were implemented (cf. *SI Appendix*, section 1).

(see *SI Appendix*, section 4B for a discussion of this approach). Fourth, our findings are limited to our study setting, that is, the first wave in Switzerland. Future research may confirm the external validity of our findings by analyzing other countries or time periods.

Inferring mobility patterns from telecommunication data is inherently coupled to the coverage of such data and our definition of trips. Only movements for individuals who carry a mobile device with a SIM card are included. In particular, trips are not included for individuals who do not carry SIM cards. Similarly, trips may be counted several times if an individual carries several SIM cards (e.g., when carrying both a phone and a SIM-based tablet). It is also possible that trips by children, elderly, or other groups of the population with less phone usage are underrepresented in the data. Furthermore, microlevel movements are not observed but inferred via triangulation between antennas through the use of a positioning algorithm achieving state-of-the-art accuracy. Despite these limitations, telecommunication data are considered to be scalable and, in particular, more complete for inferring mobility patterns compared to alternative data sources (8, 40, 41). Moreover, our objective is not to obtain accurate estimates of mobility in itself, but to evaluate the predictive ability of telecommunication data for reported case growth. Our analysis confirms telecommunication data as such a monitoring tool.

Our findings are of direct value for public decision-makers. Nationwide mobility data from mobile telecommunication networks can be leveraged for the management of epidemics. Thereby, we fill a previously noted void in the case of COVID-19 (40, 41, 45). Specifically, monitoring mobility supports public decision-makers when managing the COVID-19 epidemic in two ways. First, it helps public decision-makers in assessing the impact of policy measures targeted at mobility behavior. Second, by predicting epidemic growth, it provides a scalable tool for near real-time epidemic surveillance. Such tools are relevant for evidence-based policymaking of public authorities in the current COVID-19 epidemic.

Materials and Methods

The aim of this study is to make population inference of mobility from nationwide telecommunication data. In our study, mobility estimates derived from telecommunications data were obtained from Swisscom Mobility Insights, a

commercial data platform of a major telecommunications provider in Switzerland, and then further processed for analysis. Swisscom collects telecommunications data from routine signal exchanges (i.e., pings) with antennas, regardless of the actual service provider. Based on the telecommunication data, mobility estimates are inferred as follows (Fig. 5): 1) Telecommunication data are collected at the level of antenna. 2) Telecommunication data at antenna level are used to infer microlevel movements of individuals via triangulation. 3) Data on microlevel movements are used to count movements between postal codes (named trips) over time. This procedure is performed to capture mobility levels in the population. 4) The data are further aggregated at the cantonal level per day to link them to policy measures and COVID-19 case numbers. The following section explains Swisscom's procedure for obtaining nationwide mobility estimates from telecommunication data and how we further processed the data for analysis.

Nationwide Telecommunication Data. Telecommunication data are routinely collected from signal exchanges (i.e., pings) between mobile devices and adjacent antennas. Such signal exchanges occur for all SIM-based mobile devices (e.g., mobile phones, smartphones), regardless of the actual service provider. In particular, our data also include movements of people with a foreign SIM card and, hence, represent nationwide telecommunication. A network event between a mobile device and mobile network compromises metadata as follows: the international mobile subscriber identity (IMSI) number of the SIM card, the date and time the SIM card connected to the mobile network, and the ID of the mobile antenna to which the SIM card was connected. The IMSI number is available for all SIM cards and thus represents a unique identifier, independent of the actual service provider. The events from mobile networks are extracted from the mobile communications systems every night, and thus the mobile data are available the following day. In our analysis, we use telecommunication metadata collected by Swisscom according to the above description (46). Swisscom also ensured that IMSI numbers are stored in an anonymized format (see ref. 46 for details).

Telecommunication data hold advantages over alternative data sources for the purpose of measuring human mobility. The advantages become especially clear in comparison to location data from check-ins (9–13) or location logs from smartphone apps (14–21, 47). First, compared to smartphone applications, SIM-based devices are fairly ubiquitous. This holds for both high-income countries (such as Switzerland) and low-income countries. Second, the use of telecommunication data ensures coverage for large parts of society. Specifically, it reduces the risk of an age bias (e.g., check-ins are known to be more frequent among younger, technology-savvy people). Third, telecommunication data avoid the need for user interaction with a device. Hence, many microlevel movements are captured (e.g., school visits, commuting to work, grocery shopping) that would otherwise not be subject to monitoring.

The telecommunication infrastructure operated by Swisscom has wide coverage (46). Specifically, it covers 99.9% of the geographic area in Switzerland. The infrastructure records telecommunication metadata via almost 30,000 antennas across the country. Of these, approx. 7,000 are of the Global System for Mobile Communications (GSM) type (2G), 11,000 of the Universal Mobile Telecommunications System (UMTS) type (3G), and 12,000 of the Long-Term Evolution (LTE) type (4G) (48–50). As Switzerland has a total of 3,196 postcode areas, there are on average approx. nine antennas per postcode. Additional details on antennas are provided in *SI Appendix*, section 3.

The frequency of pings is determined by how often a mobile device connects to a new antenna or, if in between two antennas, every 5 min. Hence, if a mobile device is stationary and does not connect to a new antenna, the ping rate is once every 5 min. The rate will momentarily increase if a stationary device connects to a new antenna or if a mobile device connects to a new antenna during a trip. Importantly, the variation in ping rates across mobile devices has no influence in our analysis as the lowest possible ping rate (every 5 min) produces data of considerably higher temporal resolution than the daily aggregated data we use for the analysis. An internal algorithm by Swisscom ensures that bouncing (i.e., when a phone bounces between antennas) is correctly addressed and attributed to a single antenna. The telecommunication metadata are then used by Swisscom to infer actual locations over time via triangulation (see next section).

Inferring Positions of SIM Cards via Triangulation. The locations of SIM cards within antenna areas are determined via triangulation between antennas through the use of a positioning algorithm (51). A high-level description of the algorithm is as follows: Every signal from a SIM card in the telecommunication data is associated with a probability distribution over locations that represents the uncertainty of its actual location in a given antenna area at the time. The location is estimated from two random variables: the radius R, given by the distance of the signal from the origin of the antenna area, and its angle Θ to the antenna azimuth. Here, R is Gaussian distributed with empirical mean and variance estimated via maximum likelihood, and Θ follows a multinomial distribution depending only on the antenna azimuth and its bandwidth. The inferred locations are subject to a delay between signals between antennas and SIM cards. To address this, the location at a given point in time is estimated by marginalizing the probability distribution of the radius over the empirical distribution of signal delays estimated from all observations. Details are available in ref. 51. In sum, by tracking the location of SIM cards over time, we are able to capture microlevel movements of individuals.

The accuracy of the positioning algorithm has been empirically validated (51). The median positioning error is 132 m, making it highly accurate compared to state-of-the art methods (52). The accuracy was determined by comparing the algorithm's predicted positions to self-reported actual positions for more than 6,000 trips with over 12,000 endpoints (51).

Deriving Mobility Estimates from Telecommunication Data. Mobility has been frequently found to be helpful for understanding urban phenomena (53, 54). In this study, we use mobility estimates derived from nationwide telecommunication data.

Trips are computed as follows: A single trip is defined as the movement of a SIM card between two different postcode areas after the location has been static for 20 min (46). The trip is then counted for both postcodes. Similarly, trips that cross midnight are counted for both days. Trips from neighboring countries into Switzerland are counted for the arrival postcode.

Swisscom defines trips as movements between postal code areas as they represent the smallest spatial unit that is officially defined by the federal government. Switzerland has 3,196 postcode areas with high spatial granularity. The exact size varies between urban and rural regions, but, on average, a postcode area in Switzerland covers merely 12.9 km^2. Moreover, 71% of the Swiss working population commutes between different postcodes for work and oftentimes even between cantons (https://www.bfs.admin.ch/bfs/en/home/statistics/mobility-transport/passenger-transport/commuting.html). The average (one-way) travel distance to work is 15.0 km (see URL above), and the average daily travel distance for leisure activities

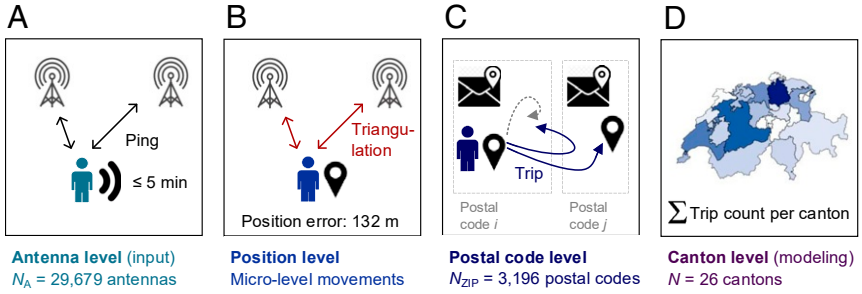

Fig. 5. Deriving mobility estimates from nationwide telecommunication data for monitoring the COVID-19 epidemic. (*A*) Telecommunication data are collected at the level of antenna. (*B*) Telecommunication data at antenna level are used to infer microlevel movements of individuals via triangulation. (*C*) Data on microlevel movements are used to count movements between postal codes (named trips) over time. (*D*) The data are aggregated at the cantonal level per day to link them to policy measures and COVID-19 case numbers.

is 36.8 km (https://www.bfs.admin.ch/bfs/en/home/statistics/mobility-transport/passenger-transport/travel-behavior.html). For both work and leisure activities, travel routinely spans several postal code areas. Hence, the use of nationwide telecommunication data combined with our definition of trips provides comparatively large-scale estimates of aggregate mobility.

For the analysis, daily trips from Swisscom Mobility Insights were further aggregated as follows: First, each trip between postcodes (including trips with departure location in a neighboring country) was mapped onto cantons at the subnational level. Here, we used cantonal shape files from the Swiss government and aggregated all daily trips within each canton. The attribution of trips to both the departure and arrival postcodes enables us to capture the number of trips between cantons as well as the number of trips between border cantons and neighboring countries. The result of the aggregation is a panel (longitudinal) dataset of trip counts in all cantons.

The reason for the aggregation to the cantonal level per day is twofold. First, policy measures are implemented within cantons, and second, COVID-19 case data are published per day only at the cantonal level. Therefore, data on policy measures and case number are not available on a more granular level.

Swisscom further labels trips according to both mode and purpose. The mode of trips was differentiated based on estimating the location of SIM cards with the positioning algorithm and the position of antennas along train, highway, and road networks. If several modes of mobility were used in the same trip, the mode with the longest leg was chosen. For comparison, public transport through train is an important mode of transportation in Switzerland (https://www.bfs.admin.ch/bfs/en/home/statistics/mobility-transport/passenger-transport/travel-behavior.html) that is relevant for explaining the results. The purpose of mobility was classified based on trips to/from work (called commuting) and all other trips (called noncommuting). This differentiation was based on the home and work locations of individuals (in terms of postal code area). These locations were derived from the most frequent geographic location of individuals between 8 PM and 8 AM for home locations and 8 AM to 5 PM for work locations. Afterward, both home and work locations were matched against the departure and arrival (postal code area) of a trip to determine whether the trip was to/from work and hence labeled commuting. The classification of trips into mode of transport is highly accurate. Specifically, a validation against self-reported data showed that 90% of all trips were correctly classified (51).

Merging of Data for Analysis. For analysis, the mobility data were merged with data on 1) policy measures, 2) the reported number of cases and hospitalizations and deaths attributed to COVID-19, 3) the number of tests conducted and their share of positive results, and 4) population sizes. All data were at the daily and cantonal level except the data on testing, which, due to lack of availability, were at the daily and country level. Data 1 to 4 are either publicly available online in the form used in the analysis or constructed from other publicly available sources. *SI Appendix*, section 1 describes how the data on policy measures were collected and validated. See *Data Availability* for details on how the remaining data were collected.

Modeling Overview. In this section, we present the regression models used to estimate the relationship between 1) policy measures and mobility and 2) mobility and reported cases. Here, the first model estimates the reduction in mobility due to policy measures. The estimates are identified with a difference-in-difference analysis and may therefore be given a causal interpretation under certain assumptions (*SI Appendix*, section 4A). The second model, in turn, estimates the extent to which reductions in mobility predict decreases in the reported number of new cases as policy measures are being implemented.

The models have parsimonious specifications recommended for isolating policy-relevant insights (44) and are informed by epidemiology. In particular, they are formalized as Bayesian hierarchical negative binomial regression models. Rather than modeling the disease dynamics themselves (as with a compartment model), our focus is on estimating the relative effect of other determinants, namely policy measures and mobility. The use of negative binomial distributions is common in epidemiological modeling, as it allows for overdispersion in the dependent variable (i.e., the number of trips and the reported number of new cases). Furthermore, each model uses a log-link between the dependent and explanatory variables. For the model of the reported number of new cases, it enables us to capture the exponential growth in cases during the initial stages of an epidemic, also observed in our data. For the model of mobility, it makes the estimates relative to the observed levels of mobility. Both dependent variables were found to follow negative binomial distributions (with overdispersion). See *SI Appendix*, section 7 for an analysis of overdispersion.

The models include further controls for 1) population size per canton, 2) unobserved heterogeneity between cantons, and 3) time effects as follows: 1) We control for differences in population size among cantons with an offset term. This is motivated by the fact that the magnitude of the estimated effects depends on the population size. Hence, the model estimates are relative to the number of inhabitants per canton. 2) Unobserved heterogeneity is estimated with a canton random effect. We thereby account for unobserved factors that affect both policy measures and mobility (for the former model) and mobility and cases (for the latter model). 3) Time effects are modeled in two ways. On the one hand, we include weekday fixed effects to control for variations in the implementation of policy measures, levels of mobility, and reporting/testing across weekdays (e.g., mobility is higher on weekdays, whereas testing is lower on weekends; thus, reported cases are lower on weekends and tests conducted on weekends may be reported on Mondays or Tuesdays). On the other hand, we incorporate a trend variable that controls for changes in case dynamics or behavioral adaptations toward social distancing that occur over time since a canton first reported a case. This could for instance occur due to unobserved changes in adherence to other policy measures (e.g., wearing masks and keeping physical distance of at least 1.5 m) over time. Here, we model the variation in when cantons reported their first case as potentially dependent on the unobserved canton heterogeneity.

The results with additional controls (e.g., testing, spatial correlation between cantons, and dependence between the different trip variables) and alternative dependent variables (e.g., hospitalizations and deaths attributed to COVID-19 instead of reported cases) are part of the robustness checks.

Model for Estimating the Reduction in Human Mobility Due to Policy Measures. A multiple time period, multiple group difference-in-difference (DiD) analysis (55) is conducted to estimate the effect of each policy measure on mobility. We restrict the analysis to the time period between 24 February and 5 April 2020, that is, starting before the first reported COVID-19 case in Switzerland and ending prior to Easter holidays. With this time period, the initial observations act as a control group in which mobility is at the baseline level (as individuals may not yet have voluntarily reduced their mobility as a response to reported cases). Furthermore, by ending at 5 April, there can be no confounding of the effects of policy measures due to Easter holidays. Such confounding would be caused by that during holidays, mobility generally changes from regular levels and, as a consequence, policy measures are more or less likely to be implemented relative to nonholidays.

Let M_{itk} denote the trip count on mobility variable $k = 1, 2, \ldots, K$ (i.e., total trips, road trips, train trips, etc.) in canton $i = 1, 2, \ldots, N$ on day $t = 1, 2, \ldots, T$. The variable M_{itk} is derived as explained in the previous section and represents the dependent variable in regression model k for the DiD analysis. The values of the model parameters depend on which mobility variable is the dependent variable of the regression; hence, we index all model parameters with k.

We model M_{itk} to follow a negative binomial distribution with conditional mean function

$$\mathbb{E}[M_{itk} \mid \eta_{it}^{(k)}, E_i] = \mu(M_{itk}) = E_i \exp \eta_{it}^{(k)}, \quad [1]$$

where E_i denotes the population size of canton i. Then, $\exp \eta_{it}^{(k)} = (\mu(M_{itk}))/E_i$ is the expected number of daily trips per inhabitant in canton i. The estimates of this model are therefore adjusted according to the variation in canton population sizes. The term $\eta_{it}^{(k)}$ is the linear predictor, specified in hierarchical form as

$$\eta_{it}^{(k)} = \alpha_i^{(k)} + \delta_{w(t)}^{(k)} + \gamma^{(k)} \log z_{it} + \sum_{l=1}^{L} \beta_l^{(k)} d_{itl}, \quad [2]$$

$$\alpha_i^{(k)} = \alpha^{(k)} + \theta_i^{(k)} + \gamma_B^{(k)} \overline{\log z_i}, \quad [3]$$

whose variables and parameters are explained in the following.

The first term $\alpha_i^{(k)}$ is a time-invariant effect specific to canton i. We model $\alpha_i^{(k)}$ as a function of several variables that vary across cantons; see Eq. **3**. Here, $\alpha^{(k)}$ is the intercept among all cantons, which represents the overall baseline relative mobility on Mondays before any policy measure was implemented and any COVID-19 cases were reported. The term $\theta_i^{(k)}$ is a random effect that captures unobserved time-invariant factors for canton i (e. g., population density) that confound the effect of policy measures on mobility.

The final variable $\overline{\log z_i}$ is discussed in detail below. The subscript B on the associated parameter denotes that it measures between-canton effects; that is, the parameter measures only the effect of increases in the variable across cantons.

The variable d_{itl} is a dummy variable that takes a value of 1 if policy measure $l \in \{1, 2, \ldots, L\}$ is implemented by canton i at day t and 0 otherwise. Hence, $\exp(\beta_l^{(k)})$ measures the multiplicative effect of policy measure l on the expected number of daily trips on mobility variable k per canton inhabitant. Note that all policy measure variables are included in Eq. **2**. Hence, the effect of each policy measure is conditional on the other policy measures being held fixed. Fig. 1*C* shows that the reduction in mobility is similar across cantons; therefore, we do not estimate the heterogeneity in the effect of the policy measures on mobility across cantons.

The term $\delta_{w(t)}^{(k)}$ represents the fixed effect of weekday w on the relative (log-transformed) mobility compared to the reference weekday (here, Monday). The term controls for the confounding factor that aggregate mobility and the probability of implementing a policy measure are likely higher on, e.g., Mondays than Sundays.

The variable z_{it} captures other sources of time-related confounding and is derived as follows: Let q_{it} be the number of days since the first reported COVID-19 case in canton i. The variable is calculated as

$$q_{it} = \begin{cases} t - t_i', & \text{if } t > t_i', \\ 0, & \text{if } t \leq t_i', \end{cases} \quad [4]$$

where t_i' is the date the first case was reported in canton i. Since the logarithm of zero is undefined, we then set $z_{it} = q_{it} + 1$ and include the logarithm of z_{it} in the model. The associated parameter $\gamma^{(k)}$ is therefore interpreted as the percentage increase in relative mobility given a 1% increase in the number of days since a canton first reported a case. The rationale for including z_{it} is that individuals may adapt their mobility behavior over time irrespective of social distancing policies. Therefore, the variable captures how mobility would trend over time even if the policy measures were not implemented.

The variable $\overline{\log z_i} = T^{-1} \sum_{t=1}^{T} \log z_{it}$ is the time average of $\log z_{it}$ in canton i (the bar over the expression denotes an average value). The variable is included in the model to allow the canton-specific effect $\theta_i^{(k)}$ to be correlated with $\log z_{it}$ over the cantons. Such correlation would, for instance, arise if the date that the first COVID-19 case is reported in each canton depends on the unobserved canton-specific factors. As an example, the date the first case is reported in a canton could depend on the (unobserved) adherence of inhabitants to social distancing recommendations. If such correlation exists but is ignored, it would instead enter the error term of the model, leading to a violation of the exogeneity assumption and incorrect parameter estimates. By including $\overline{\log z_i}$, we essentially make $\theta_i^{(k)}$ a Mundlak-type correlated random effect (56). The benefit of correlated random effects over fixed effects or standard random effects is that they use only the within-unit variation to estimate parameters (and, hence, give identical estimates to those of fixed-effects models) while also having the random effects property of estimating the variation in the unobserved heterogeneity via partial pooling. Note that time averages of the policy measure variables or weekday effects are not included in the model since those are conditionally exogenously determined and, therefore, uncorrelated with the model errors. We refer to ref. 57 for a detailed discussion of the underlying benefits of this approach relative to fixed effects.

By substituting the linear predictor Eq. **2** into the conditional mean function Eq. **1** and expanding $\alpha_i^{(k)}$, the full model of mobility variable k becomes

$$\log \mathbb{E}[M_{itk} \mid \eta_{it}^{(k)}, E_i] = \log E_i + \alpha^{(k)} + \theta_i^{(k)} + \delta_{w(t)}^{(k)} + \gamma^{(k)} \log z_{it} + \gamma_B^{(k)} \overline{\log z_i} + \sum_{l=1}^{L} \beta_l^{(k)} d_{itl}. \quad [5]$$

The conditional variance of M_{itk} is given by

$$\mathbb{V}[M_{itk} \mid \eta_{it}^{(k)}, E_i, \zeta^{(k,M)}] = \mu(M_{itk}) \left(1 + \frac{\mu(M_{itk})}{\zeta^{(k,M)}}\right), \quad [6]$$

where $\zeta^{(k,M)}$ is the overdispersion parameter (the superscript M distinguishes the overdispersion parameter of the mobility model from the model of reported cases).

We specify one regression equation in the form of Eq. **5** for each mobility variable and estimate them separately. Each regression has the same explanatory variables but a different mobility variable as the dependent variable (i.e., total trips or one of the mobility variables based on mode or purpose).

Model for Estimating the Relationship between Mobility and COVID-19 Cases. The model for estimating the relationship between mobility and reported COVID-19 cases is similar in structure to the model used to link policy measures to mobility. To accommodate the forecast horizon, we lag the mobility variables to estimate how a decrease in mobility at a given day predicts reductions in the reported number of new cases at a later day. This enables forecasting of future reported case growth by evaluating the model at daily observed mobility levels.

Let C_{it} denote the cumulative number of reported cases in canton $i = 1, 2, \ldots, N$ until and including day $t = 1, 2, \ldots, T_i$. Then, $Y_{it} = C_{it} - C_{i,t-1}$ is the number of new cases that are reported in canton i on day t. Note that the time period (i.e., total number of days T_i) varies across cantons $i = 1, \ldots, N$. The reason for this is that the dependent variable of this regression model is the reported number of new cases, and as such we restrict the data to start at the date of each canton's first reported case. Hence, the data for this model start between 24 February and 16 March (depending on the canton) and end at 5 April (for all cantons). Our modeling approach accounts for the resulting unbalanced number of observations per canton.

We model Y_{it} as following a negative binomial distribution with conditional mean function

$$\mathbb{E}[Y_{it} \mid \eta_{it}^{(k,s)}, E_i] = \mu^{(k,s)}(Y_{it}) = E_i \exp \eta_{it}^{(k,s)}, \quad [7]$$

where

$$\eta_{it}^{(k,s)} = \alpha_i^{(k,s)} + \delta_{w(t)}^{(k,s)} + \gamma^{(k,s)} \log z_{it} + \xi_{ks} \log m_{i,t-s,k}, \quad [8]$$

$$\alpha_i^{(k,s)} = \alpha^{(k,s)} + \theta_i^{(k,s)} + \gamma_B^{(k,s)} \overline{\log z_i} + \xi_{ks,B} \overline{\log m_{ik}} \quad [9]$$

is the hierarchical linear predictor. In this model, $\exp \eta_{it}^{(k,s)} = (\mu^{(k,s)}(Y_{it}))/E_i$ is the expected number of reported positive cases in canton i on day t relative to the canton population size (sometimes called the relative risk in spatial epidemiology) (58, 59). The model for the daily growth in reported cases can then be written as

$$\log \mathbb{E}[Y_{it} \mid \eta_{it}^{(k,s)}, E_i] = \log E_i + \alpha^{(k,s)} + \theta_i^{(k,s)} + \delta_{w(t)}^{(k,s)} + \gamma^{(k,s)} \log z_{it} + \gamma_B^{(k,s)} \overline{\log z_i} + \xi_{ks} \log m_{i,t-s,k} + \xi_{ks,B} \overline{\log m_{ik}}. \quad [10]$$

The conditional variance of Y_{it} given mobility variable k lagged by s days is given by

$$\mathbb{V}[Y_{it} \mid \eta_{it}^{(k,s)}, E_i, \zeta^{(k,s,Y)}] = \mu^{(k,s)}(Y_{it}) \left(1 + \frac{\mu^{(k,s)}(Y_{it})}{\zeta^{(k,s,Y)}}\right), \quad [11]$$

where $\zeta^{(k,s,Y)}$ is the overdispersion parameter.

For simplicity, we use the same notation for variables and parameters in this model as in the model used for the effect of policy measures on mobility (but the estimated parameters have, of course, a different interpretation and values). The superscript (k, s) is attached to parameters to indicate that their values depend on the choice of mobility variable k and its lag s.

The parameter of interest is ξ_{ks}. It measures the expected percentage change in the reported number of new cases per canton inhabitant s days after a 1% increase in mobility variable k. Hence, the parameter shows how the relative growth rate in reported cases changes as a function of lagged mobility, after adjusting for relevant factors but where mobility varies according to which policy measures are implemented. Note that we intentionally include only a single lag s (and then refit the model for different lags) rather than including multiple lags at the same time. We follow this approach to avoid issues related to multicollinearity between the mobility variables and because one cannot condition on mobility in future days when predicting future reported cases from current mobility. *SI Appendix*, section 4B further explains our rationale.

The intercept $\alpha^{(k,s)}$ gives the baseline number of reported cases relative to the canton population for Mondays.

The parameter $\delta_{w(t)}^{(k,s)}$ is the effect of weekday w relative to the Monday effect. By including weekday effects in the model, we control for confounding differences in the number of trips and the number of reported COVID-19 cases between weekdays that would bias the parameter estimate

on the lagged mobility variable. For instance, people travel to schools and work primarily on weekdays and, similarly, there are fewer COVID-19 tests on weekends and thus fewer reported cases on Monday/Tuesday (due to reporting delays).

The variable $\log z_{it}$ is also included in this model. It now controls for the fact that mobility and the reported case growth both depend on when the first case was reported in a canton.

The Mundlak-style random effects $\theta_i^{(k,s)}$ estimate the impact of unobserved canton-specific factors that may be correlated with both the variation in mobility across cantons and the logarithm of the number of days since the first case was reported in each canton. We achieve this by including variables of the time averages of $\log z_{it}$ and lagged mobility; that is, $\overline{\log m_{ik}} = T^{-1} \sum_{t-s} \log m_{i,t-s,k}$. Then, any potential cross-canton correlation between the random effect and $\log m_{i,t-s,k}$ or $\log z_{it}$ via the model's error term is controlled for.

The above regression model is fitted separately for each (k, s), that is, each pair of mobility variable and lag. This allows us to investigate to what extent different lags of each mobility variable predict the number of reported cases.

Estimation Details. We estimate our models in a fully Bayesian framework. We run four Markov chains for every model, each with 2,000 warm-up samples and another 2,000 samples from the posterior distributions. Since our models are fitted with a log-link, we transform the posterior parameter samples so that they give estimates for the original scale of the dependent variable. For each parameter, we report in our plots the posterior mean and the associated 80 and 95% CrIs of the transformed posterior distribution.

The software used for estimation is the R package brms (60, 61) version 2.11.1 built upon the statistical modeling platform Stan (62). Parameter estimates are obtained by Markov chain Monte Carlo sampling in Stan version 2.19.2, using the Hamiltonian Monte Carlo algorithm (63, 64) and the No-U-Turn sampler (NUTS) (65).

Table 1 presents our choices of priors for the variables in the models. We use weakly informative priors to stabilize the computations and provide some regularization. Our prior on each β_l reflects that we believe that each policy measure reduces the logarithm of expected mobility by 25%, on average, but that effects between 0 and 50% are relatively probable. The prior on each ξ_{ks} implies that we expect that a 1% decrease in the lagged mobility variable predicts a 1% decrease in reported cases for each of the considered lags, with negative effect sizes or effect sizes exceeding 2% being unlikely. The prior on γ implies that we expect the relative outcome to increase by 1% for each 1% increase in the number of days since the first reported case. The intercept, overdispersion parameter, and standard deviation of the canton random effects are given weakly informative priors. The prior on $\delta_{w(t)}$ states that the effect of a given weekday that is not Monday should fall within 50 to 150% of the Monday effect. The parameters for the variables of between-canton averages are assigned vague priors since we have no a priori belief of their effects.

Model Diagnostics. We followed common practice for model diagnostics of Bayesian models (66). For each of the models, we inspected 1) posterior predictive checks, 2) divergent transitions, 3) effective sample size and convergence of the Markov chains, 4) overdispersion in the dependent variables, 5) influential observations, and 6) correlation between the policy parameters. All model diagnostics indicate a good fit. Details are provided in *SI Appendix*, section 7.

Robustness Checks. First, we checked the robustness of the model estimates against alternative specifications of time effects: 1) A model is specified as in the main text but where the logarithmic trend is replaced with corresponding linear and quadratic trends (of the number of days since the first reported case in each canton) to capture nonlinearities in both the reported case dynamics and general behavior toward social distancing. 2) A model is specified with additional week fixed effects (i.e., a weekday fixed effect, a week fixed effect, and a trend variable in logarithmic form). This model allows us to control for weekly exogenous shocks (e. g., media reports about the shortage of critical care in Italy) but acknowledge that such fixed effects would be unknown at the time of forecasting and, therefore, cannot be used to predict reported case growth at a future date. All models yield similar estimates and hence confirm the predictive ability of the mobility variables (*SI Appendix*, section 6A).

Second, the number of reported cases could potentially depend on the number of conducted tests per canton and day. When controlling for this, we obtain similar estimates (*SI Appendix*, section 6B).

Third, we extend the models by including a spatial random effect, as commonly used in the spatial epidemiology and disease mapping literature (67, 68). This approach allows us to account for the spatial dependence in mobility between neighboring cantons. We find that the spatial dependence is low and retrieve estimated effects of policy measures that are practically identical to those of the main analysis (*SI Appendix*, section 8A).

Fourth, we account for potential dependence between different mobility variables by estimating their models jointly (that is, by modeling the covariance of the canton random effects for the different mobility variables). This model yields slightly narrower credible intervals but almost identical point estimates for the effects of the policy measures (*SI Appendix*, section 8B).

Data Availability. Human mobility data presented in this work are available from the Swisscom Mobility Insights platform (https://mip.swisscom.ch). Cantonal geographic boundaries can be found as shape files at the Federal Office of Topography, swisstopo (https://www.swisstopo.admin.ch/en/geodata/landscape/boundaries3d.html). Data on reported COVID-19 cases per canton and relative to the cantonal population size (i.e., cases per 100,000 inhabitants) come from the Federal Office of Public Health of the Swiss Confederation: BAG (https://www.covid19.admin.ch/en/overview). We also use this source to obtain data on deaths and hospitalizations attributed to COVID-19 per canton and data on testing conducted in Switzerland. Additional information on the Swiss population comes from the Swiss Federal Statistical Office: BFS (https://www.bfs.admin.ch/bfs/en/home/statistics/catalogues-databases/data.assetdetail.14087625.html). Data on antenna locations are obtained from a web tool (https://map.geo.admin.ch/) at the Swiss federal geoportal, which is based on data provided by the Federal Office of Topography, swisstopo (https://www.swisstopo.admin.ch/en/home.html), and the Federal Office of Communications, OFCOM (direct links to

Table 1. Choice of priors

Parameter	Description	Prior	Model
β_l	Effect of policy measure l	N(−0.25, 0.25)	Eq. **5**
ξ_{ks}	Effect of log mobility variable k with a lag of s	N(1, 1)	Eq. **10**
α	Intercept	Half-t(3, 1.8, 2.5)	Eqs. **5** and **10**
θ_i	Canton random effect	N(0, σ_θ)	Eqs. **5** and **10**
σ_θ	Standard deviation for canton random effect	Half-t(3, 0, 2.5)	Eqs. 5 and **10**
$\delta_{w(t)}$	Effect of weekday w (compared to Monday)	N(0, 0.5)	Eqs. **5** and **10**
γ	Effect of log no. of days since first reported case	N(1, 1)	Eqs. **5** and **10**
γ_B	Effect of between-canton average of log no. of days since first reported case	N(0, 5)	Eqs. **5** and **10**
$\xi_{ks,B}$	Effect of between-canton average of log mobility with a lag of s	N(0, 5)	Eq. **10**
ζ	Overdispersion in dependent variable	Gamma(0.01, 0.01)	Eqs. **6** and **11**

The superscripts (k) and (k, s) are omitted as the same priors are assigned to each model. The column "Description" states what effect the associated parameter represents (except for the overdispersion parameter).

the maps: https://s.geo.admin.ch/8f7aa435b8, https://s.geo.admin.ch/8f7aa5536b, and https://s.geo.admin.ch/8f7aa69498). Details on data collection for policy measures are provided in *SI Appendix*, section 1. When referring to cantons, we use abbreviations instead of full canton names (https://www.admin.ch/ch/d/gg/pc/documents/1336/Abkuerzungsverzeichnis.pdf).

Anonymized algorithms, code, and data for reproducing our results are available at our GitHub page (https://github.com/jopersson/covid19-mobility).

Ethics Declarations. S.F. declares membership in a COVID-19 Working Group by the World Health Organization (WHO) but without competing interests. Ethics approval (2020-N-41) was obtained from the institutional review board at ETH Zurich (Swiss Federal Institute of Technology).

ACKNOWLEDGMENTS. We thank Dominik Hangartner and Achim Ahrens for invaluable feedback. We also thank Swisscom for their extensive support. J.P. and S.F. acknowledge funding from the Swiss National Science Foundation on data-driven health management (Grant 186932).

1. World Health Organisation coronavirus disease (COVID-19) situation report (2020). https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports. Accessed 10 December 2020.
2. S. Flaxman *et al.*, Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* **584**, 257–261 (2020).
3. S. Hsiang *et al.*, The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature* **584**, 262–267 (2020).
4. N. W. Ruktanonchai *et al.*, Assessing the impact of coordinated COVID-19 exit strategies across Europe. *Science* **369**, 1465–1470 (2020).
5. S. Lai *et al.* Effect of non-pharmaceutical interventions to contain COVID-19 in China. *Nature* **585**, 410–413 (2020).
6. H. J. T. Unwin *et al.*, State-level tracking of COVID-19 in the United States. *Nat. Commun.* **11**, 6189 (2020).
7. N. Banholzer *et al.*, Estimating the effects of non-pharmaceutical interventions on the number of new infections with COVID-19 during the first epidemic wave. *PLoS ONE* **16**, e0252827 (2021).
8. K. H. Grantz *et al.*, The use of mobile phone data to inform analysis of COVID-19 pandemic epidemiology. *Nat. Commun.* **11**, 4961 (2020).
9. S. G. Benzell, A. Collis, C. Nicolaides, Rationing social contact during the COVID-19 pandemic: Transmission risk and social benefits of US locations. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 14642–14644 (2020).
10. S. Gao *et al.*, Mobile phone location data reveal the effect and geographic variation of social distancing on the spread of the COVID-19 epidemic. arXiv [Preprint] (2020). https://arxiv.org/abs/2004.11430 (Accessed 16 June 2021).
11. S. Chang *et al.*, Mobility network models of covid-19 explain inequities and inform reopening. *Nature* **589**, 82–87 (2020).
12. D. M. Dave, A. I. Friedson, K. Matsuzawa, J. J. Sabia, When do shelter-in-place orders fight COVID-19 best? Policy heterogeneity across states and adoption time. *Econ. Inq.* **59**, 29–52 (2020).
13. S. Gupta *et al.*, Tracking public and private response to the COVID-19 epidemic: Evidence from state and local government actions. https://www.nber.org/system/files/working_papers/w27027/w27027.pdf. Accessed 16 June 2021.
14. A. Adiga *et al.*, Interplay of global multi-scale human mobility, social distancing, government interventions, and COVID-19 dynamics. medRxiv [Preprint] (2020). https://doi.org/10.1101/2020.06.05.20123760 (Accessed 16 June 2021).
15. M. Chinazzi *et al.*, The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* **368**, 395–400 (2020).
16. A. Galeazzi *et al.*, Human mobility in response to COVID-19 in France, Italy and UK. arXiv [Preprint] (2020). https://arxiv.org/abs/2005.06341 (Accessed 16 June 2021).
17. H. Fang, L. Wang, Y. Yang, Human mobility restrictions and the spread of the novel coronavirus (2019-nCoV) in China. *J. Public Econ.* **191**, 104272 (2020).
18. M. U. G. Kraemer *et al.*, The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* **368**, 493–497 (2020).
19. R. Li *et al.*, Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* **368**, 489–493 (2020).
20. H. Tian *et al.*, An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* **368**, 638–642 (2020).
21. G. Bonaccorsi *et al.*, Economic and social consequences of human mobility restrictions under COVID-19. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 15530–15535 (2020).
22. P. Nouvellet *et al.* Imperial College London COVID-19 response team – Report 26: Reduction in mobility and COVID-19 transmission. (2020). https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-26-mobility-transmission/. Accessed 16 June 2021.
23. Y. Kang *et al.*, Multiscale dynamic human mobility flow dataset in the U.S. during the COVID-19 epidemic. *Sci. Data* **7**, 390 (2020).
24. N. E. Kogan *et al.*, An early warning approach to monitor COVID-19 activity with multiple digital traces in near real-time. arXiv [Preprint] (2020). https://arxiv.org/abs/2007.00756 (Accessed 16 June 2021).
25. J. Huang *et al.*, "Understanding the impact of the COVID-19 pandemic on transportation-related behaviors with human mobility data" in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, R. G. Gupta, Y. L. Liu, J. T. Tang, B. A. P. Prakash, Eds. (Association for Computing Machinery, New York, NY, 2020), pp. 3443–3450.
26. C. Xiong, S. Hu, M. Yang, W. Luo, L. Zhang, Mobile device data reveal the dynamics in a positive relationship between human mobility and COVID-19 infections. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 27087–27089 (2020).
27. H. S. Badr *et al.*, Association between mobility patterns and COVID-19 transmission in the USA: A mathematical modelling study. *Lancet Infect. Dis.* **20**, 1247–1254 (2020).
28. J. S. Jia *et al.*, Population flow drives spatio-temporal distribution of COVID-19 in China. *Nature* **587**, 389–394 (2020).
29. B. Jeffrey *et al.*, Imperial College London COVID-19 response team – Report 24: Anonymised and aggregated crowd level mobility data from mobile phones suggests that initial compliance with COVID-19 social distancing interventions was high and geographically consistent across the UK. *Wellcome Open Res.* **5**, 1–14 (2020).
30. G. Pullano, E. Valdano, N. Scarpa, S. Rubrichi, V. Colizza, Population mobility reductions during COVID-19 epidemic in France under lockdown. medRxiv [Preprint] (2020). https://www.medrxiv.org/content/10.1101/2020.05.29.20097097v1 (Accessed 16 June 2021).
31. M. Vinceti, T. Filippini, K. J. Rothman, F. Ferrari, A. Goffi, Lockdown timing and efficacy in controlling COVID-19 using mobile phone tracking. *EClinicalMedicine* **25**, 100457 (2020).
32. F. Schlosser *et al.*, COVID-19 lockdown induces disease-mitigating structural changes in mobility networks. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 32883–32890 (2020).
33. N. W. Ruktanonchai *et al.*, Identifying malaria transmission foci for elimination using human mobility data. *PLoS Comput. Biol.* **12**, e1004846 (2016).
34. A. Wesolowski *et al.*, Quantifying the impact of human mobility on malaria. *Science* **338**, 267–270 (2012).
35. C. Viboud, A. Vespignani, The future of influenza forecasts. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 2802–2804 (2019).
36. L. Bengtsson *et al.*, Using mobile phone data to predict the spatial spread of cholera. *Sci. Rep.* **5**, 8923 (2015).
37. A. Wesolowski *et al.*, Quantifying seasonal population fluxes driving rubella transmission dynamics using mobile phone data. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 11114–11119 (2015).
38. A. Wesolowski *et al.*, Impact of human mobility on the emergence of dengue epidemics in Pakistan. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 11887–11892 (2015).
39. E. Pollina, D. Busvine, European mobile operators share data for coronavirus fight (2020). *Reuters*, https://www.reuters.com/article/us-health-coronavirus-europe-telecoms-idUSKBN2152C2. Accessed 5 August 2020.
40. C. O. Buckee *et al.*, Aggregated mobility data could help fight COVID-19. *Science* **368**, 145 (2020).
41. N. Kishore *et al.*, Measuring mobility to monitor travel and physical distancing interventions: A common framework for mobile phone data analysis. *Lancet Digit. Health* **2**, e622–e628 (2020).
42. A. N. Desai *et al.*, Real-time epidemic forecasting: Challenges and opportunities. *Health Security* **17**, 268–275 (2019).
43. J. Blumenstock, G. Cadamuro, R. On, Predicting poverty and wealth from mobile phone metadata. *Science* **350**, 1073–1076 (2015).
44. A. L. Bertozzi, E. Franco, G. Mohler, M. B. Short, D. Sledge, The challenges of modeling and forecasting the spread of COVID-19. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 16732–16738 (2020).
45. N. Oliver *et al.*, Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. *Sci. Adv.* **6**, eabc0764 (2020).
46. Swisscom, Swisscom mobility insights. https://www.swisscom.ch/en/business/enterprise/offer/enterprise-mobile/mobility-insights.html. Accessed 10 December 2020.
47. M. U. G. Kraemer *et al.*, Mapping global variation in human mobility. *Nat. Hum. Behav.* **4**, 800–810 (2020).
48. Standorte mobilfunkmasten GSM. https://opendata.swisscom.com/explore/dataset/standorte-mobilfunkmasten-gsm/table/?disjunctive.powercode&sort=-id. Accessed 10 December 2020.
49. Standorte mobilfunkmasten UMTS. https://opendata.swisscom.com/explore/dataset/xy_pwr_umts_170101/information/?disjunctive.powercode. Accessed 10 December 2020.
50. Standorte mobilfunkmasten LTE. https://opendata.swisscom.com/explore/dataset/standorte-mobilfunkmasten-lte/information/?disjunctive.powercode. Accessed 10 December 2020.
51. M. Kafsi, Quantifying the accuracy of mobility insights from cellular network data. https://mkafsi.medium.com/quantifying-the-accuracy-of-mobility-insights-from-cellular-network-data-e5b83437a609. Accessed 23 November 2020.
52. I. Leontiadis *et al.*, "From cells to streets: Estimating mobile paths with cellular-side data" in *ACM International on Conference on Emerging Networking Experiments and Technologies*, A. S. Seneviratne, Ed. (Association for Computing Machinery, New York, NY, 2014), pp. 121–132.
53. A. Bassolas *et al.*, Hierarchical organization of urban mobility and its connection with city livability. *Nat. Commun.* **10**, 4817 (2019).
54. M. C. Gonzalez, C. A. Hidalgo, A.-L. Barabasi, Understanding individual human mobility patterns. *Nature* **453**, 779–782 (2008).

55. A. Goodman-Bacon, Difference-in-differences with variation in treatment timing https://www.nber.org/system/files/working_papers/w25018/w25018.pdf. Accessed 16 June 2021.
56. Y. Mundlak, On the pooling of time series and cross section data. *Econometrica: J. Econ. Soc.* **46**, 69–85 (1978).
57. A. Bell, K. Jones, Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. *Pol. Sci. Res. Method* **3**, 133–153 (2015).
58. A. Riebler, S. H. Sørbye, D. Simpson, H. Rue, An intuitive Bayesian spatial model for disease mapping that accounts for scaling. *Stat. Methods Med. Res.* **25**, 1145–1165 (2016).
59. N. Asmarian, S. M. T. Ayatollahi, Z. Sharafi, N. Zare, Bayesian spatial joint model for disease mapping of zero-inflated data with R-INLA: A simulation study and an application to male breast cancer in Iran. *Int. J. Environ. Res. Publ. Health* **16**, 4460 (2019).
60. P. C. Bürkner, BRMS: An R package for Bayesian multilevel models using Stan. *J. Stat. Software* **80**, 1–28 (2017).
61. P. C. Bürkner, Advanced Bayesian multilevel modeling with the R package BRMS. *R J.* **10**, 395–411 (2018).
62. B. Carpenter *et al.*, Stan: A probabilistic programming language. *J. Stat. Software* **76**, 1–32 (2017).
63. R. M. Neal *et al.*, "MCMC using Hamiltonian dynamics" in *Handbook of Markov Chain Monte Carlo*, S. B. Brooks, A. G. Gelman, G. J. Jones, X.-L. M. Meng, Eds. (Chapman and Hall/CRC, 2011), vol. 2, pp. 113–162.
64. S. Duane, A. D. Kennedy, B. J. Pendleton, D. Roweth, Hybrid Monte Carlo. *Phys. Lett. B* **195**, 216–222 (1987).
65. M. D. Hoffman, A. Gelman, The No-U-Turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.* **15**, 1593–1623 (2014).
66. A. Gelman *et al.*, *Bayesian Data Analysis* (CRC Press, 2013).
67. J. Besag, J. York, A. Mollié, Bayesian image restoration, with two applications in spatial statistics. *Ann. Inst. Stat. Math.* **43**, 1–20 (1991).
68. J. C. Wakefield, N. G. Best, L. Waller, "Bayesian approaches to disease mapping" in *Spatial Epidemiology: Methods and Applications*, P. E. Elliott, J. W. Wakefield, N. B. Best, D. B. Briggs, Eds. (Oxford University Press, 2000), pp. 104–127.