HAL
open science

Mining Google and Apple mobility data: temporal anatomy for COVID-19 social distancing

Corentin Cot, Giacomo Cacciapaglia, Francesco Sannino

To cite this version:

Corentin Cot, Giacomo Cacciapaglia, Francesco Sannino. Mining Google and Apple mobility data: temporal anatomy for COVID-19 social distancing. Scientific Reports, 2021, 11 (1), pp.4150. 10.1038/s41598-021-83441-4 . hal-03150531

HAL Id: hal-03150531

https://hal.science/hal-03150531v1

Submitted on 24 Feb 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

scientific reports

Check for updates

OPEN

Mining Google and Apple mobility data: temporal anatomy for COVID-19 social distancing

Corentin Cot[1,2,5], Giacomo Cacciapaglia[1,2,5]✉ & Francesco Sannino[3,4,5]✉

We employ the Google and Apple mobility data to identify, quantify and classify different degrees of social distancing and characterise their imprint on the first wave of the COVID-19 pandemic in Europe and in the United States. We identify the period of enacted social distancing via Google and Apple data, independently from the political decisions. Our analysis allows us to classify different shades of social distancing measures for the first wave of the pandemic. We observe a strong decrease in the infection rate occurring two to five weeks after the onset of mobility reduction. A universal time scale emerges, after which social distancing shows its impact. We further provide an actual measure of the impact of social distancing for each region, showing that the effect amounts to a reduction by 20–40% in the infection rate in Europe and 30–70% in the US.

COVID-19 has disrupted our way of living with long lasting impact on our social behaviour and the world economy. At the same time, differently from earlier pandemics, a very large amount of data has been collected[1–4] thanks, also, to our smartphone dominated society. Smartphones run mobility applications, such as Google and/or Apple Maps, that help humans navigate. The mobility information stemming from these apps has been harvested by Google and Apple, which have subsequently made it publicly available on the following websites: Google[5] and Apple[20].

In this paper we mine these data to quantify and characterise the effects of social distancing measures enacted by various European countries and American states. An early study of mobility effects on the pandemic evolution in China can be found in Ref.[6]. Mobility data has also been studied in the context of the United States (US), with data collected from various sources. Google data has been shown to correlate with the political decisions taken in mid-march in each state[1] and to precede a reduction of the case growth by a test period ranging between 2 to 4 weeks. A similar study[7] at county level, based on anonymised cellular (mobile) data, found a reduction of the growth rate after 9–12 days (up to 3 weeks). Further correlations have been tested in relation to a reduction in the fever cases[8] and the income level of US counties[9], the latter drawing mobility data from multiple sources. Besides including European countries[10], in this study we will correlate the mobility reduction to a model of the infection evolution, which will allow us to extract from the data a more reliable delay between the enacting of the mobility measures and the reduction in the infection rates. Furthermore, we do not rely on political decisions, but rather define the timing of the mobility reduction based on the data provided by Apple and Google. In this way, our results do not depend on the variability and diversity of political decisions at various times and in different regions.

The Google mobility data, in Google wordings, *show movement trends by region, across different categories of places.* As categories we will use "Residential" and "Workplace", which best describe the change in people's behaviour after the implementation of social distancing measures with respect to a baseline day. The latter is defined, according to Google, as the median value of the 5-week period from the 3rd of January to the 6th of February, 2020, predating the wide spread of the virus in Europe and in the US. The data show how visitors to (or time spent in) categorised places changed with respect to the baseline day. For Apple, the available mobility data represent a relative volume of direction requests per country/region, sub-region or city, compared to a baseline volume defined on the 13th of January, 2020. We will be using, from Apple, information about "Driving" and "Walking", assuming they represent the time spent by people away from home. For the US, only "Driving" data

[1]Institut de Physique des 2 Infinis (IP2I), CNRS/IN2P3, UMR5822, 69622 Villeurbanne, France. [2]Université de Lyon, Université Claude Bernard Lyon 1, 69001 Lyon, France. [3]CP3-Origins & the Danish Institute for Advanced Study, University of Southern Denmark, Campusvej 55, 5230 Odense, Denmark. [4]Dipartimento di Fisica E. Pancini, Università di Napoli Federico II & INFN sezione di Napoli, Complesso Universitario di Monte S. Angelo Edificio 6, via Cintia, 80126 Napoli, Italy. [5]These authors contributed equally: Corentin Cot, Giacomo Cacciapaglia and Francesco Sannino. ✉email: g.cacciapaglia@ipnl.in2p3.fr; sannino@cp3.sdu.dk

are available. We identified the minimal set of mobility indicators that allowed us to time the implementation of social distancing measures. This timing is for us a crucial quantity to determine, for the first time, the correlation with a change in the infection rates. To this extent, all other categories would lead to the same results but could be of relevance for complementary studies, for instance for the economic impact.

Another set of data relevant for this work is related to the virus spreading dynamics, which we take from the websites https://ourworldindata.org/ and https://covidtracking.com. We normalise the data of each country as cases per million inhabitants.

The data relative to the total number of infected cases are effectively parameterised using the High Energy Physics inspired formalism[11], dubbed *epidemic Renormalisation Group* (eRG). The approach has been generalised to take into account the spreading dynamics across different regions of the world[12] and the evolution of the second wave pandemic across Europe[13]. The advantage of the eRG formalism resides in the limited number of coefficients needed to classify the spreading dynamics for each country. More complicated models have been used in the literature to study the effect of non-pharmaceutical interventions, including mobility, for Europe[14] and the US[15–19], with the latter mostly focusing on local communities.

Without further ado, we introduce $\alpha(t)$ below[11,12]

$$\alpha(t) = \ln(\mathcal{I}(t)) \,, \tag{1}$$

where $\mathcal{I}(t)$ is the total number of infected cases per million inhabitants in a given region and ln indicates its natural logarithm. The function $\alpha(t)$ turns out to be well described by the following logistic function:

$$\alpha(t) = \frac{ae^{\gamma t}}{b + e^{\gamma t}} \,. \tag{2}$$

Here, a represents the logarithm of the final number of infected cases per million inhabitants, b denotes the temporal shift from the start of the pandemic and γ measures the *flatness* of the curve of the number of new infected cases. Here, and in the following, we will measure the time t in weeks, so that γ is measured in inverse weeks. It has been argued[11,12] that, aside from the trivial temporal shift provided by b and for the first wave of the pandemic, two numbers are sufficient to characterise the evolution of the number of infected cases per each region, i.e. a and γ. This fact helps studying the correlation between mobility data and the virus spreading dynamics for each region. By going beyond the previous parameterisation, we will discover a finer temporal structure directly related to the effects of the imposed lockdown and social distancing measures in the different regions.

In this work we focus on a selection of European countries and all of the US states. In Europe, we considered countries with more than 3 million inhabitants and for which the data were available. Note that we will only consider the period from March to May 2020, during which the first wave of the COVID-19 was raging in Europe and in the US.

Results

Using Google and Apple data, we provide a rationale to identify the timing of the social distancing measure actualisation in each region. European countries and American states adopted different degrees of social distancing measures during the first wave of the COVID-19 pandemic. Moreover the severity of the measures changed during the spreading of the epidemic within each region of the world. This is why we defined the beginning of the impact of social distancing measures in terms of the reduction in the mobility of individuals, rather than on political decisions.

We mine Google's Residential and Workplace mobility data since they show movement trends across different places compared to a reference period before the implementation of any measure. The Residential and Workplace data are best suited to quantify when and to what extent people reduced their mobility and increased social isolation. Similarly, for Apple, we choose Driving and Walking data for Europe and Driving for the US states, expressing them in terms of a percentage reduction. Note that the Apple data refer to variations in the number of searches done on the Maps app, more details to be found on the website Apple[20].

We define an immobility indicator $\mathcal{M}$, as described in the section Methods, in terms of an average mobility reduction in the chosen categories. The average is taken over six weeks after the beginning of social distancing. This indicator allows to sort the European countries and the American states based on the hardness of social distancing. We also define regions with the highest rate of mobility reduction (*low mobility*, LM) and regions with the least reduction (*high mobility*, HM), for Europe and the US separately. The results are shown in Fig. 1, were we indicate the LM regions in cyan and the HM regions in red, with a colour gradient representing different shades of mobility being proportional to the value of the indicator. For Europe, the countries with the smallest mobility grossly correspond to those that imposed a lockdown, while the highest mobility country is Sweden, where no measures were imposed. Nevertheless, even for Sweden the mobility data show a significant variation that allows us to define the beginning of social distancing despite the political decisions. Similarly for the US states, the lowest mobility corresponds to states in the North-East, California and Hawaii, which imposed lockdown measures. We also noticed that the beginning of the measures, as defined by the mobility data in the US, corresponds to the dates when the schools were closed in each state[1].

To validate our conclusions, at the bottom of Fig. 1 we show the correlations between Google and Apple mobility data for Europe (left and central plot) and the US (right plot). Each region is represented by a tadpole-like symbol, with the head corresponding to the 6-week average and the tail to the 8-week average. We label each country and state by using the same colour code as in the maps. The plots show a clear correlation between the percentage change in each category. We also checked that the same correlation persists when comparing Google to Apple categories.

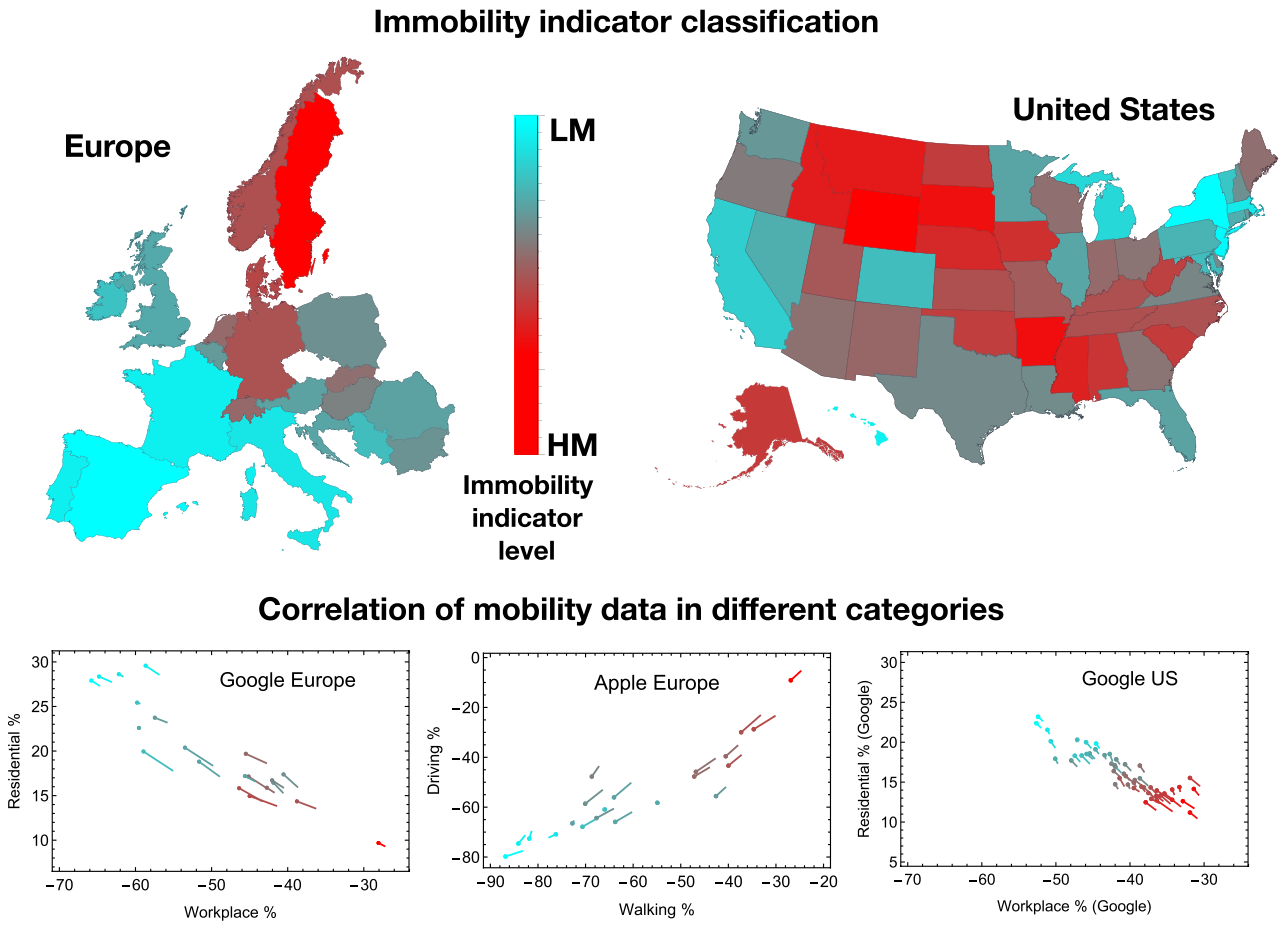

Figure 1. The COVID-19 Mobility Map for Europe and the US. The two maps represent respectively the European and US states with different shades of mobility from the highest (HM) in bright red to the lowest (LM) in cyan. At the bottom of the figure there are three tadpole-like plots showing correlations between the four mobility reduction categories: Residential and Workplace from Google, Driving and Walking from Apple. The head of the tadpoles correspond to the average over 6 weeks after social distancing begins, while the tail indicates a 8 week average. The colour code in the three plots reflects the maps one. The maps are drawn with Wolfram Mathematica.

We now analyse possible correlations between mobility data and the parameters of the logistic function $\alpha(t)$ such as the infection rate γ and the log of the total number of infected cases a. To our surprise we find that γ is uncorrelated to the degree of mobility reduction. This implies that mobility changes have little impact on the velocity of diffusion of the disease. Of course, mobility data only capture one aspect of the social distancing, thus they do not offer a complete picture of the situation in various regions. This surprising finding can be interpreted in various ways. On the one hand, the result may imply that the main factor behind a reduction of γ could lie in the behaviour of individuals in social occasions (mask wearing, proximity, greeting habits, to mention a few); on the other hand, it is quite possible that the value of γ does not represent the effect of the social distancing measures, as it derives from a global fit over a wide timescale. In other words, the fit values include both the measure and the pre-measure periods.

To push further the analysis, we now explore whether social distancing measures (as defined via the Apple/Google mobility data) lead to distinct temporal patterns in the European countries under study and the American states. In the eRG approach, γ is the natural parameter to use for this task. We assume that, after the measures are enacted, there are two distinct temporal regions describing the time dependence of the number of infected cases. These two regions, B and C in the illustrative plot in Fig. 2e, are naturally described by two different gammas. We confirm that such an analysis is possible via a MonteCarlo analysis. We then move to the actual data and discover that two distinct temporal regions with their own gammas do emerge for several regions. In Fig. 2 we show the outcome of the fit to the data in terms of the time interval Δt between the beginning of social distancing and its effects measured when the infection rate γ changes. We discover that most countries display a similar Δt. By fitting the distributions in Fig. 2c to a gaussian, we find that to the two sigma level we have $\Delta t = 2.7 \pm 1.7$ weeks for Europe and $\Delta t = 3.3 \pm 1.6$ weeks for the US. The high compatibility of the two ranges shows the emergence of a universal time scale for social distancing to be effective.

Another important result is the general and strong reduction of the infection rate measured within and after Δt both for Europe and the US, as shown in the left panels of Fig. 2 and summarised by the red histograms of Fig. 2d.

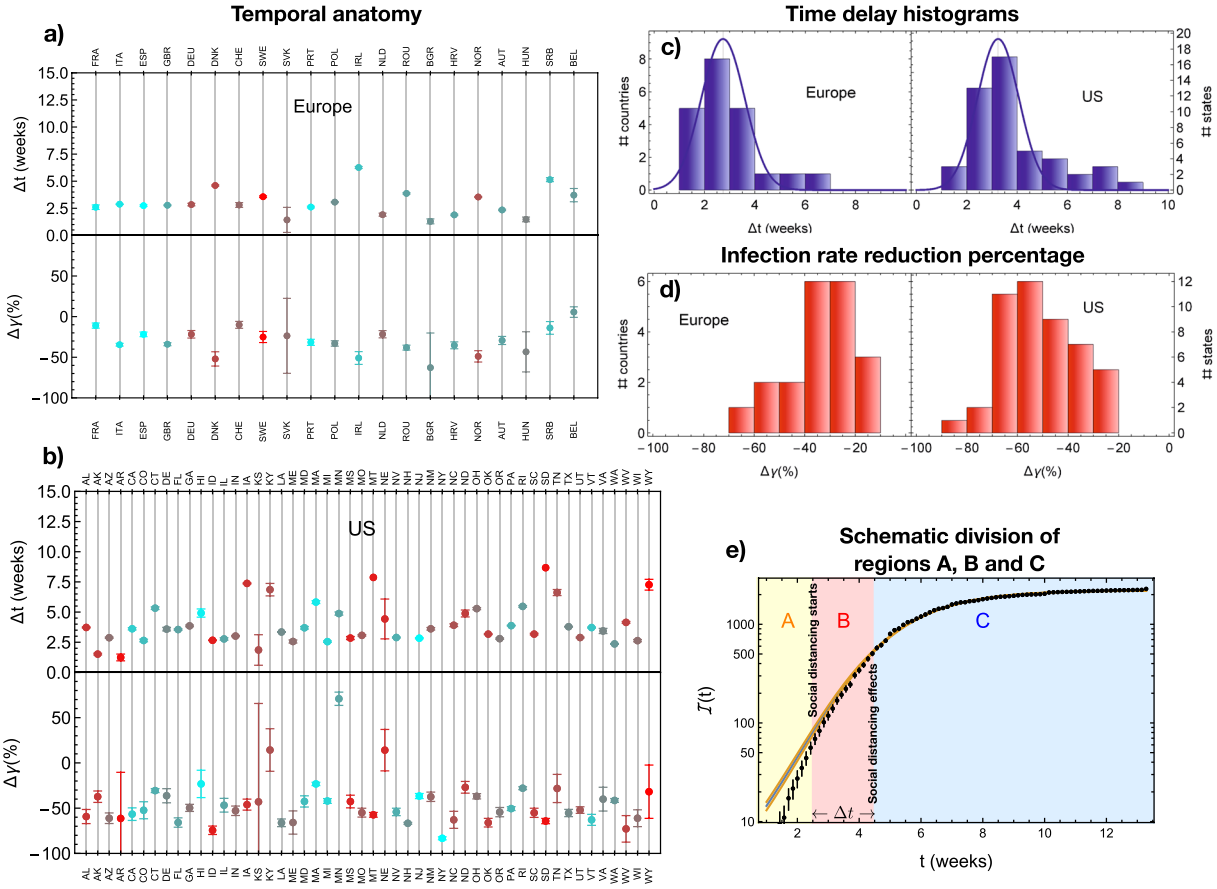

Figure 2. Temporal anatomy of COVID-19 social distancing effects. In panel (**a**), we show Δt and the percentage variation $\Delta\gamma$ in the infection rate for the European countries considered in this study. In panel (**b**), we show the same for all the US states. In panels (**c**, **d**) we display the same results in the form of histograms, for Europe and the US separately, highlighting that Δt clusters around similar values. In panel (**e**), we illustrate the subdivision of the first wave epidemic curve in three temporal regions: A before social distancing as defined via mobility data occurs, B until an effect is observed in the epidemic curve as a change in γ, and C covering the later times. Δt equals to the duration of the period B.

Methods

Immobility indicator. European countries and US states adopted different degrees of social distancing measures during the first wave of the COVID-19 pandemic. Moreover the severity of the measures changed during the spreading of the epidemic within each country or state. Rather than classifying the countries based on their political choices, we use the mobility data provided by Google and Apple as indicators of the effective hardness of the measures.

To find a measure for the *immobility* of a given population during the social distancing period, we define an average percentage variation for each of the four categories: Residential and Workplace for Google and Driving and Walking for Apple (only Driving is available for US states). For both mobility datasets, the percentage variations are defined with respect to a reference date or period predating the exponential growth of the infection cases. The data are typically very jugged, as illustrated in Fig. S1 in the supplementary material, mainly due to strong variations over the weekend. Furthermore, the mobility data feature a sharp decrease followed by a slow return to the pre-COVID-19 average. Taking into account this behaviour, it is necessary to define an average over several weeks, which would allow us to associate a single number to each category and region.

Firstly, one needs to properly define the beginning of the social distancing period for each region: we choose to identify it with the time when Google Workplace percentage first drops by 20% (at this time, typically, all mobility indicators have shown a significant variation). The ending of the measure period is harder to identify, as the social distancing measures have always been lifted progressively[3]: this appears in the mobility data, as the curves gradually return to zero, i.e. to the reference period levels, or even above. Thus, we decided to fix the same averaging period for all the regions we considered. To test the robustness of our conclusions, we determine the outcome for two choices: 6 and 8 weeks after the effective beginning of the measures. The tadpole-like plots at the bottom of Fig. 1 demonstrate that the duration of the averaging period, while changing the value of the mobility reduction, does preserve the overall trend. In the following, therefore, we will use the 6-week average as our benchmark.

To be able to classify the countries based on their immobility, we further define an immobility indicator as

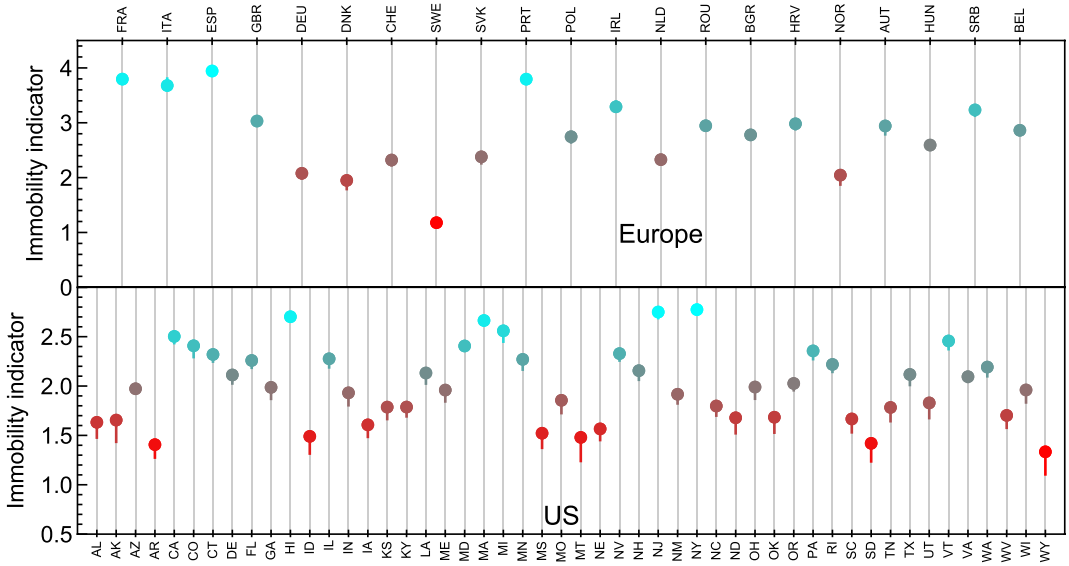

Figure 3. Immobility indicator for the European countries and the US states. Values of the immobility indicator $\mathcal{M}$ for Europe (top) and the US (bottom). The colour code corresponds to the ranking of each European country and each US state, matching the one used in Fig. 1.

$$\mathcal{M}(\text{ region}) = \sum_{j=\text{ cat.}} \frac{|p_j(\text{ region})|}{\max\left[|p_j|\right]}, \tag{3}$$

where $|p_j(\text{ region})|$ is the absolute value of the percentage variation in each category (labelled by j). For each category, we divide by the maximal value observed in the pool. Note that for European countries we have four categories, so that $\mathcal{M} < 4$, while for the US states we have 3 categories, so that $\mathcal{M} < 3$. We use this indicator to rank the European countries and the American states from the ones with *high mobility* (HM)—small $\mathcal{M}$—to the one with *low mobility* (LM)—large $\mathcal{M}$. The values of the immobility indicator we obtain for the European countries under study and US states are shown in Fig. 3. The colour code ranges from the highest mobility region in bright red to the lowest mobility one in cyan, with gradient proportional to the value of the immobility indicator.

Comparing the virus spreading parameters with mobility data. The epidemic evolution of the first wave of the COVID-19 pandemic can be effectively characterised by two parameters: the infection rate γ and the logarithm of the final number of total infected cases a[11], measured per million inhabitants. We remark, however, that it is risky to compare the number of infected for different regions due to the different procedures used when identifying the positive cases, and the different testing rates and strategies. Thus, we assign more physical meaning to the infection rates γ, which give an accurate temporal characterisation of the epidemic diffusion in each region.

It is, therefore, natural to hypothesise that regions with higher mobility may have a faster diffusion rate of the infection, i.e. larger values of γ. To test this hypothesis, in Fig. 4, we show the Workplace, Residential and Driving reductions versus the infection rates for the European countries in this study and the US states. To each country or state is associated a racecar-like symbol: the pilot seat (dot) corresponds to the 6-week average, while the tail to the 8-week average. Furthermore, the side bars indicate the error from the fits of the epidemic data. The colour codes match the immobility indicator defined above. The data used to generate the plots in Fig. 4 are reported in Tables T1 and T3 in the supplementary material, where we only report the mobility averages over 6 weeks.

Surprisingly, the data do not reveal any particular correlation between the values of γ and the mobility data. As explained in the previous section, this result can be interpreted in various ways. One possibility, which we will test, is that the γ from the fit of the first wave is not the most appropriate measure, as it averages over the infection diffusion before and after the mobility reduction occurs.

Testing the two-gamma hypothesis. We subdivide the period of the virus diffusion in three parts, as illustrated in Fig. 2e. Region A extends up to the time when the social distancing starts, $t = 0$, as defined from the mobility data; at this point Region B begins extending for a duration Δt; finally Region C starts at $t = \Delta t$. As the beginning of Region B is determined by the Google/Apple mobility data, we can probe the existence of a change in γ by fitting the data in Region B + C with the following function:

$$\alpha_{2\gamma}(t) = \begin{cases} a\frac{\exp(\gamma_B t)}{b+\exp(\gamma_B t)} & \text{for } t < \Delta t \\ a\frac{\exp(\gamma_C t)}{b\exp((\gamma_C-\gamma_B)\Delta t)+\exp(\gamma_C t)} & \text{for } t > \Delta t \end{cases} \tag{4}$$

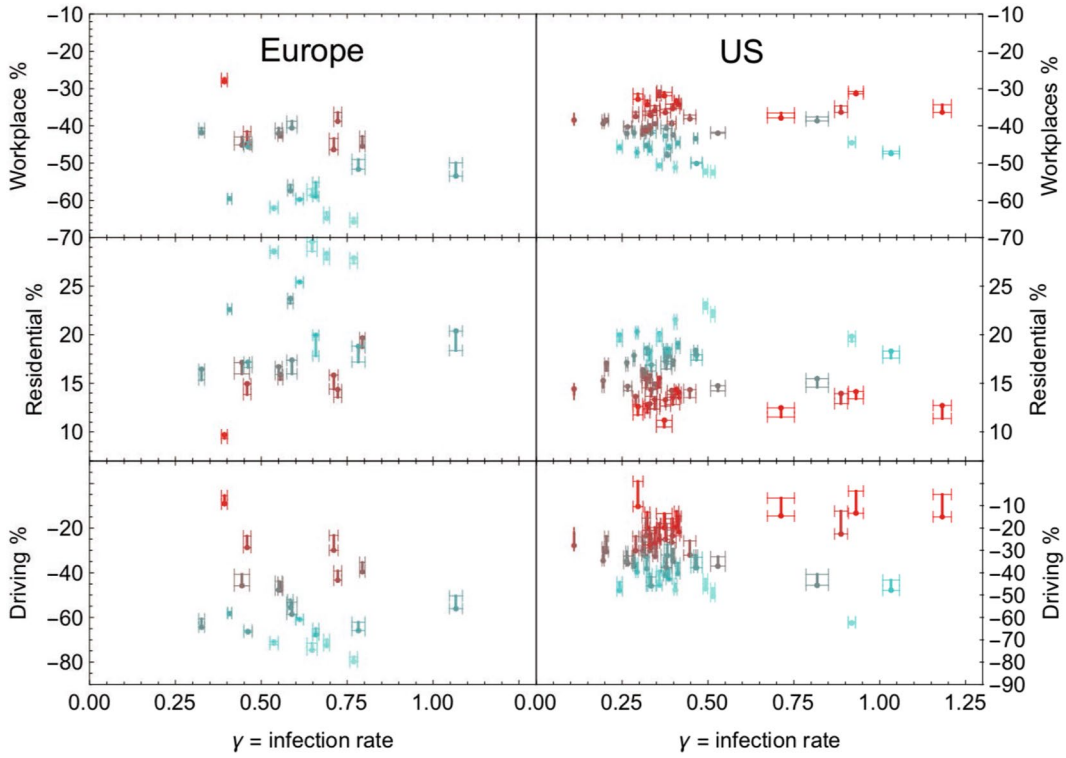

Figure 4. Infection rate compared to the mobility data. Racecar plots showing the fitted infection rates γ versus the Google/Apple mobility categories. The vertical segment indicates the difference between 6 week (dot) and 8 week averages; the horizontal bars indicate the fit error on γ.

that depends on five parameters: a, b, γ_B, γ_C and Δt. We then extract the values of the five parameters by fitting to the data.

We first test the effectiveness of our method by generating a mock set of data based on the function in Eq. (4), where we fix $\gamma_B = 0.7, \gamma_C = 0.35$ and $\Delta t = 20$ days. An example of the generated data, overlaid to the generating function, is shown in the right panel of Fig. S2 in the supplementary material: the points are randomly generated within a one standard deviation region, i.e. $[N_i - \sqrt{N_i}, N_i + \sqrt{N_i}]$, where N_i is the number of cases per day as predicted by the generating function. We generated 100 independent sets of mock data and fitted them to Eq. (4). We found that we can determine the value of Δt within a range of two weeks. Furthermore, we define the percentage variation of the infection rate as

$$\Delta\gamma = \frac{\gamma_C - \gamma_B}{\gamma_C}. \tag{5}$$

Having acquired confidence in the method, we now apply it to the real data. The results of the fits are reported in Tables T2 and T4 in the supplementary material.

Conclusions

We analysed the mobility data released by Google and Apple to quantify the effects of social distancing on the COVID-19 spreading dynamics in Europe and in the US. We:

1. Classified different shades of social distancing measures for the first pandemic wave.
2. Observed (after having identified the countries according to their level of immobility) a strong decrease in the infection rate occurring two to five weeks after the onset of mobility reduction.
3. Discovered a universal time scale after which social distancing shows its impact.
4. Provided an actual measure of the impact of social distancing for each region, showing that the effect amounts to a reduction of 20–40% of the infection rate for most countries in Europe and 30–70% in the US.

The above results lead to the first global and direct measure of the impact of social distancing. Interestingly, even countries that did not impose political measures, like Sweden, show a reduction of the infection rate similar to the ones experiencing a lockdown, suggesting that a certain degree of social restrain occurred regardless of the political decisions. Our results are compatible with early analysis of local social distancing measures taken in China[6], where mobility data inter-cities from Baidu was used within a compartmental model.

Using smartphone based open-source mobility data, we showed that it is possible to provide a temporal anatomy of social distancing. We discovered the emergence of a characteristic time scale related to when social distancing effects have a measurable impact. This timing can also be used to quantify the impact of social distancing by determining the variation in infection rate per country. Finding similar reduction, however, does not imply that the countries have a similar number of infected cases per million inhabitants. It simply means that there has been a change in social behaviour. The result of this study, based on the simple eRG approach, lays the basis for an effective tool for the authorities to evaluate the timing and impact of the imposition of social distancing measures, in particular related to movement restrictions.

Data availability

The data for the COVID-19 infected cases in Europe are extracted from the `ourworldindata.org` repository, while the data in the American states is from covidtracking.com. The mobility data are provided open source by Google and Apple.

Code availability

The results obtained in this work are based on analytic expressions presented in this section and open source data. Upon request, we can provide the Wolfram Mathematica notebook used to analyse the data via our analytical expressions.

Received: 27 September 2020; Accepted: 26 January 2021
Published online: 18 February 2021

References

1. Wellenius, G. A. *et al.* Impacts of state-level policies on social distancing in the united states using aggregated mobility data during the covid-19 pandemic (2020). arXiv:2004.10172.
2. Lurie, M. N., Silva, J., Yorlets, R. R., Tao, J. & Chan, P. A. COVID-19 epidemic doubling time in the United States before and during stay-at-home restrictions. *The J. Infect. Dis.* **222**(10), 1601–1606. https://doi.org/10.1093/infdis/jiaa491 (2020). https://academic.oup.com/jid/advance-article-pdf/doi/10.1093/infdis/jiaa491/33644763/jiaa491.pdf.
3. Islind, A. S., Óskarsdóttir, M. & Steingrímsdóttir, H. Changes in mobility patterns in europe during the covid-19 pandemic: Novel insights using open source data (2020). arXiv:2008.10505.
4. Yang, C. *et al.* Taking the pulse of covid-19: A spatiotemporal perspective. *Int. J. Digit. Earth* **13**, 1186–1211. https://doi.org/10.1080/17538947.2020.1809723 (2020).
5. COVID-19 Community Mobility Reports. *Google*, https://www.google.com/covid19/mobility/ (2020).
6. Lai, S. *et al.* Effect of non-pharmaceutical interventions for containing the covid-19 outbreak in China. *Nature* **585**, 410–413. https://doi.org/10.1038/s41586-020-2405-7 (2020).
7. Badr, H. S. *et al.* Association between mobility patterns and covid-19 transmission in the USA: A mathematical modelling study. *Lancet Infect. Dis.* **20**, 1247–1254. https://doi.org/10.1016/S1473-3099(20)30553-3 (2020).
8. Liautaud, P., Huybers, P. & Santillana, M. Fever and mobility data indicate social distancing has reduced incidence of communicable disease in the United States (2020). arXiv:2004.09911.
9. Huang, X. *et al.* The characteristics of multi-source mobility datasets and how they reveal the luxury nature of social distancing in the U.S. during the covid-19 pandemic. *medRxiv* https://doi.org/10.1101/2020.07.31.20143016 *(2020).*
10. Cacciapaglia, G., Cot, C. & Sannino, F. Mining google and apple mobility data: Twenty-one shades of european social distancing measures for covid-19 (2020). arXiv:2008.02117v1.
11. Della Morte, M., Orlando, D. & Sannino, F. Renormalization Group approach to pandemics: The COVID-19 case. *Front. Phys.* **8**, 144. https://doi.org/10.3389/fphy.2020.00144 (2020).
12. Cacciapaglia, G. & Sannino, F. Interplay of social distancing and border restrictions for pandemics (COVID-19) via the epidemic Renormalisation Group framework. *Sci. Rep.* **10**, 15828, https://doi.org/10.1038/s41598-020-72175-4 (2020). arXiv:2005.04956.
13. Cacciapaglia, G., Cot, C. & Sannino, F. Second wave covid-19 pandemics in Europe: A temporal playbook. *Sci. Rep.* **10**, 15514, https://doi.org/10.1038/s41598-020-72611-5 (2020). arXiv:2007.13100.
14. Flaxman, S. *et al.* Estimating the effects of non-pharmaceutical interventions on covid-19 in Europe. *Nature* **584**, 257–261. https://doi.org/10.1038/s41586-020-2293-x (2020).
15. Kabiri, A., Darzi, A., Zhou, W., Sun, Q. & Zhang, L. How different age groups responded to the covid-19 pandemic in terms of mobility behaviors: A case study of the United States (2020). arXiv:2007.10436.
16. Nielsen, F., Marti, G., Ray, S. & Pyne, S. Clustering patterns connecting covid-19 dynamics and human mobility using optimal transport (2020). arXiv:2007.10677.
17. Fellows, I. E., Slayton, R. B. & Hakim, A. J. The covid-19 pandemic, community mobility and the effectiveness of non-pharmaceutical interventions: The United States of America, February to May 2020 (2020). arXiv:2007.12644.
18. Hong, B., Bonczak, B., Gupta, A., Thorpe, L. & Kontokosta, C. E. Exposure density and neighborhood disparities in covid-19 infection risk: Using large-scale geolocation data to understand burdens on vulnerable communities (2020). arXiv:2008.01650.
19. Vanni, F., Lambert, D. & Palatella, L. Epidemic response to physical distancing policies and their impact on the outbreak risk (2020). arXiv:2007.14620.
20. COVID-19 Mobility Trends Reports. *Apple*, https://covid19.apple.com/mobility (2020).

Acknowledgements

G.C. and C.C. acknowledge partial support from the Labex-LIO (Lyon Institute of Origins) under grant ANR-10-LABX-66 (Agence Nationale pour la Recherche), and FRAMA (FR3127, Fédération de Recherche "André Marie Ampère").

Author contributions

This work has been designed and performed conjointly and equally by the authors. G.C., C.C. and F.S. have equally contributed to the writing of the article.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-83441-4.

Correspondence and requests for materials should be addressed to G.C. or F.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021