

## Modelling virus spreading in ride-pooling networks

Kucharski, Rafał; Cats, Oded; Sienkiewicz, Julian

**DOI**

[10.1038/s41598-021-86704-2](https://doi.org/10.1038/s41598-021-86704-2)

**Publication date**

2021

**Document Version**

Final published version

**Published in**

Scientific Reports

**Citation (APA)**

Kucharski, R., Cats, O., & Sienkiewicz, J. (2021). Modelling virus spreading in ride-pooling networks. *Scientific Reports*, 11(1), Article 7201. <https://doi.org/10.1038/s41598-021-86704-2>

**Important note**

To cite this publication, please use the final published version (if applicable). Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.



OPEN

## Modelling virus spreading in ride-pooling networks

Rafał Kucharski<sup>1,2✉</sup>, Oded Cats<sup>1</sup> & Julian Sienkiewicz<sup>3</sup>

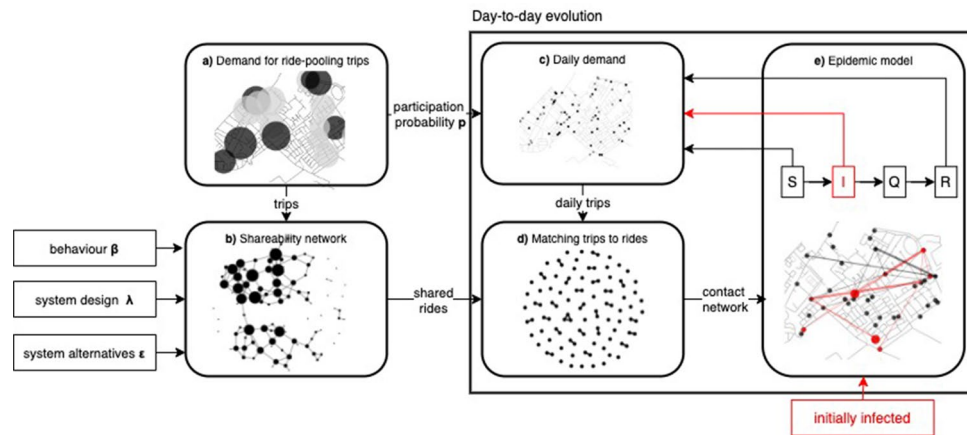
Urban mobility needs alternative sustainable travel modes to keep our pandemic cities in motion. Ride-pooling, where a single vehicle is shared by more than one traveller, is not only appealing for mobility platforms and their travellers, but also for promoting the sustainability of urban mobility systems. Yet, the potential of ride-pooling rides to serve as a safe and effective alternative given the personal and public health risks considerations associated with the COVID-19 pandemic is hitherto unknown. To answer this, we combine epidemiological and behavioural shareability models to examine spreading among ride-pooling travellers, with an application for Amsterdam. Findings are at first sight devastating, with only few initially infected travellers needed to spread the virus to hundreds of ride-pooling users. Without intervention, ride-pooling system may substantially contribute to virus spreading. Notwithstanding, we identify an effective control measure allowing to halt the spreading before the outbreaks (at 50 instead of 800 infections) without sacrificing the efficiency achieved by pooling. Fixed matches among co-travellers disconnect the otherwise dense contact network, encapsulating the virus in small communities and preventing the outbreaks.

In the era of widespread concerns about personal safety and exposure to virus transmission, urban mobility faces an unprecedented challenge<sup>1–3</sup>. While mass transit, a crowded backbone of pre-pandemic megacities' mobility systems, is under societal pressure due to health concerns related to its potential role in virus spreading<sup>4,5</sup>, people search for other travel alternatives that reduce one's exposure. The natural reaction of risk-averse travellers is to opt for individual transport modes, such as private cars<sup>6</sup>, which can be devastating for the sustainability of pandemic urban mobility systems<sup>7</sup>. To counteract this, we explore whether shared mobility may offer an attractive alternative by efficiently serving travel demand using a shared fleet while allowing users to avoid the crowd. **Ride-pooling**<sup>8,9</sup>, available via two-sided mobility platforms (such as UberPool and Lyft), has recently emerged as a travel alternative in cities worldwide and gained attention or researchers studying both specific systems (like Singapore<sup>10</sup> and New York City<sup>11</sup>) as well as uncovering universal scaling laws governing cities<sup>12–14</sup> which in turn allow for generalizing the results to any urban system. Travellers, requesting rides are offered a pooled ride, where they share a single vehicle with co-travellers riding in a similar direction. Despite being perceived by policymakers as a solution for improving mobility and sustainability by leveraging on the platform economy revolution, the COVID pandemic led to safety concerns among sharing travellers (worried about their health), policymakers (concerned about public health and epidemic outbreak) and operators (uncertain about the future of their business).

While preliminary findings on COVID-19<sup>15</sup> suggest transmission taking place in proximity (e.g. among co-travellers within the same vehicle), evidence on how and if the virus transmits beyond a single vehicle is lacking. Indeed, the potential of shared rides to serve as an alternative, in-between the mass transit (where - perceived or real - virus exposure may be high) and private cars (which generate negative externalities) remains largely unknown. Will the random infected passenger spread the virus across a large number of travellers across the network, or will it be encapsulated and thus confined to a distinct community? How many other travellers will get infected and how will the epidemiological process evolve? Finally, can we mitigate it by effective control and design measures and thus introduce it to policymakers as a safe alternative? Such questions are valid as COVID-19 pandemic constantly challenge our current policies<sup>16,17</sup> calling for a long-term preparedness<sup>18</sup>.

The propagation of different types of epidemics (biological and social) has long been a playground of network science community (summed up in the seminal work of Pastor-Satorras et al.<sup>19</sup> and reaching as far as to propose the idea of “physics of vaccination”<sup>20</sup>). Addressing also mobility network (e.g. public transport) structures<sup>21,22</sup> with their complex topology and temporal evolution studies<sup>23</sup>. Recent COVID-19 propagation studies either follow a coarse-grained level of aggregated cases<sup>24–26</sup>, adopt purely synthetic network structures<sup>27,28</sup> or lack emerging mobility modes (like ride-pooling) in the picture<sup>3,6</sup>. This study brings to the front the network evolution (crucial

<sup>1</sup>Department of Transport and Planning, Delft University of Technology, Delft, The Netherlands. <sup>2</sup>Department of Transport Systems, Cracow University of Technology, Cracow, Poland. <sup>3</sup>Faculty of Physics, Warsaw University of Technology, Warsaw, Poland. ✉email: r.m.kucharski@tudelft.nl



**Figure 1.** Methodology at glance: We consider travel demand for ride-pooling trips (a), for which we compute a shareability network (b) with a given behavioural parameters  $\beta$ , system design  $\lambda$  and alternatives' attractiveness  $\epsilon$ . We simulate the day-to-day evolution of spreading until the virus is halted. Each day we obtain the daily demand (c), consisting of those who want and can travel (decided to travel with probability  $p$  and are not quarantined). Daily trip demand is optimally assigned to shared rides, which forms the contact network (d) on which virus spreading is then modelled (e). Starting from initially infected travellers, each day we simulate epidemic transitions: susceptible travellers are infected by infected co-travellers who quarantine after 7 days and return immune to the system after 14 days.

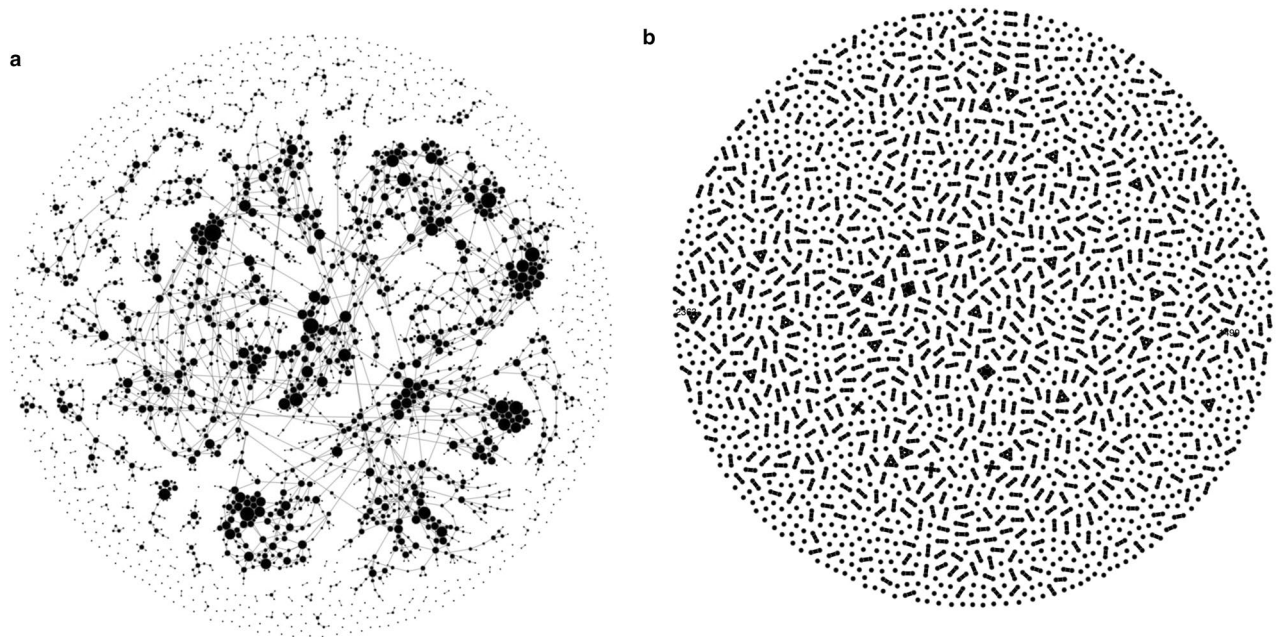
in the context of epidemic spreading<sup>29,30</sup>) and couples it with empirical, behaviour driven contact network<sup>9</sup>, specific to ride-pooling, yielding a simulation framework (see Methods for details) capable to provide rich and realistic insights into possible epidemic outbreaks specific to ride-pooling networks.

We model the evolution of virus spreading in ride-pooling systems through an extensive set of experiments with demand sampled from actual mobility patterns of afternoon commuters in Amsterdam. The underlying **shareability network** (see Methods for details how such network is set up) is the outcome of travellers' willingness to share, which depends on whether they are sufficiently close to each other in terms of induced detour (compatibility of origins and destinations) and delay (compatibility of departure times) compared to a private ride-hailing. The resulting dynamic, time-dependent contact network<sup>31</sup> is subject to day-to-day variations as well as the results of the iterative **SIQR epidemiological model**<sup>32</sup> (see Methods section for explanation). We examine the resulting spreading process, i.e. number of infections along with its temporal and spatial evolution. To instantly show the methodology at glance, Fig. 1 presents all steps of our framework and their inter-dependencies, further detailed in Methods section.

Findings from our extensive simulation study are on first sight devastating, with only few initially infected travellers needed to spread the virus to hundreds of ride-pooling users. Even under conservative assumptions where the driver is not a spreader and mobility pattern is restricted to only two trips per day, the virus makes its way to infect the majority of the giant component. Introducing natural stochasticity and non-recurrence inherent to travel demand triggers a virus transmission. Despite slow temporal evolution virus gradually makes its way towards new communities, neglecting natural spatial barriers. There seems to be no epidemic threshold and even two initial infections may trigger an outbreak and reach high transmissivity. This is a very alarming finding, suggesting that ride-pooling system without intervention may substantially contribute to virus spreading. Nonetheless, we identified effective control measure allowing to halt the spreading before the outbreaks. Namely, if we trade-off spontaneity of platform-based ride-pooling service and let the operator fix matches with co-travellers, radically different image appears. Such setting disconnects the otherwise dense contact network, containing the virus in small communities and preventing the outbreaks. Notably, this trade-off is not at the cost of system effectiveness, most importantly not at the cost of occupancy rate, which may remain at the original level. We argue that under strict demand control measures, mobility platforms may provide an appealing alternative service in-between public and private transport modes for pandemic reality. Universal properties of ride-pooling networks<sup>12,14,33</sup> allow us to generalise our Amsterdam findings to a generic systems for which the critical mass needed to induce sharing comes along with a highly connected shareability network, whereas fixing the matches disconnects it into isolated communities.

## Application and Results

To understand how the virus spreads among travellers sharing rides, we conduct a series of experiments within a realistic travel demand setting of Amsterdam. Afternoon commuters, sampled from the actual trip demand dataset<sup>35</sup>, hail a ride from a mobility platform to reach their destination. They may opt for a shared ride if reduced trip fare will compensate for any detour and delay imposed by sharing. We consider a system where a 30% discount is offered for sharing and we specify the private ride-hailing ride as an alternative  $\epsilon$ . We employ behavioural parameters (value-of-time and willingness-to-share  $\beta$ 's) in-line with recent findings<sup>36,37</sup> and apply *ExMAS* algorithm<sup>9</sup> to reproduce a behaviourally rich shareability network connecting 3200 travellers to 11,000



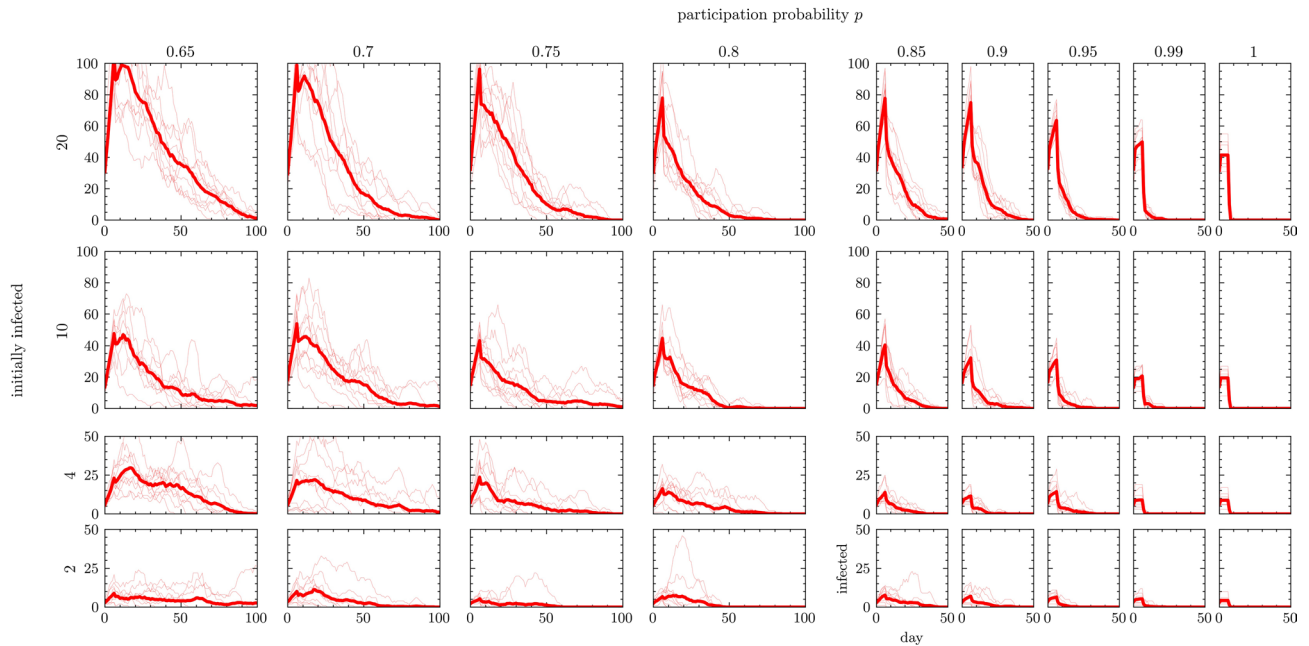
**Figure 2.** Shareability graph linking 3 200 travellers to 11,000 pooled rides feasible for them (a). Size of nodes is proportional to degree (number of travellers for shared rides and number of feasible rides for travellers). The graph structure includes a giant component and a high degree nodes, which may become a super-spreaders, as well as isolated peripheral nodes, where travellers either cannot find a feasible match or form a small, isolated communities from which virus will not outbreak. The actual matching of travellers to shared rides on a single day (b) has a substantially different structure. Here (b) nodes denote travellers, linked if they share a ride. Single dots are unmatched travellers riding alone, while lines, triangles and squares denote pooled rides of higher degree (2, 3 and 4, respectively). While the potential shareability (a) is densely connected, matching on a single day (b) is disintegrated. Each pooled-ride forms an isolated community (i.e. co-travellers within a single vehicle), with a clique of size bounded with vehicle capacity (four in our case). The virus will spread within each clique but will not reach beyond it on a given day. However, infected traveller may be assigned to a new ride on successive day, resulting with virus propagation beyond the single vehicle. Networks visualized with `newtulf`<sup>34</sup>.

feasible shared rides (see Methods for algorithm description). The size of travel demand sample is such that, on one hand, the critical mass needed to induce sharing can be attained and on the other hand, it represents a relatively low demand levels reached by ride-pooling services so far<sup>38</sup>. Notably, we model demand for shared rides which is non-stable and fluctuating from one day to the other<sup>39</sup>, hence each afternoon is comprised of a slightly or significantly different pool of travellers, controlled through a *demand stability* parameter  $p$  (i.e., the participation probability, see Methods). To allow for comparisons, while experimenting with  $p$  we keep the total daily number of travellers in the system fixed (to 2000), yet we adjust the pool of passenger from which we draw them on any given day. In our series of experiments we explore demand stability varying from  $p = 0.65$  (where each day we draw from the pool of  $3075 = 2000/0.65$  travellers) up to  $p = 1$  (where the total demand is assumed constant). We vary the number of initially infected travellers from 2 to 20. To assess the impact of demand level, we conduct experiments where we gradually increase it up to 3200 travellers. In order to account for the impact of their random location, we replicate each scenario 20 times.

We present the results through epidemic evolution plots (Fig. 3) for various settings of demand stability and number of initial infections, summarised with boxplots of the total number of infections in Fig. 4a. On Fig. 4c we trace the efficiency of ride-pooling across scenarios. In Fig. 4b we explore spreading for increasing demand levels and reveal exponential growth of the share of infected characterized by fitting coefficients scaling linearly with participation probability  $p$  (Fig. 4c). To understand the impact of demand stability on the course of outbreaks we plot node degree evolution in Fig. 5a and the distribution of transmissivity in Fig. 5b, further visualised in terms of its spatial distributions in Fig. 6. To demonstrate the potential to control the outbreak, we display its first phase in Fig. 5c.

**Outbreaks.** As long as the demand is unstable and varies considerably from one day to the other, the virus may outbreak even when only a small number of passengers are initially infected. Outbreaks starting from two infections are highly variable (Fig. 3 lower left). For initial spreaders located centrally in a highly connected giant component of the contact network (Fig. 2) the virus outbreaks and eventually infects over 250 travellers during the course of the spreading, whereas outbreaks from disconnected part of network can be naturally contained and halted already after 7 infections (Fig. 4a). An outbreak starting from 20 initial spreaders is always devastating and reaches from 450 up to almost 800 travellers, it needs only few days to reach 100 cases (Fig. 3 upper





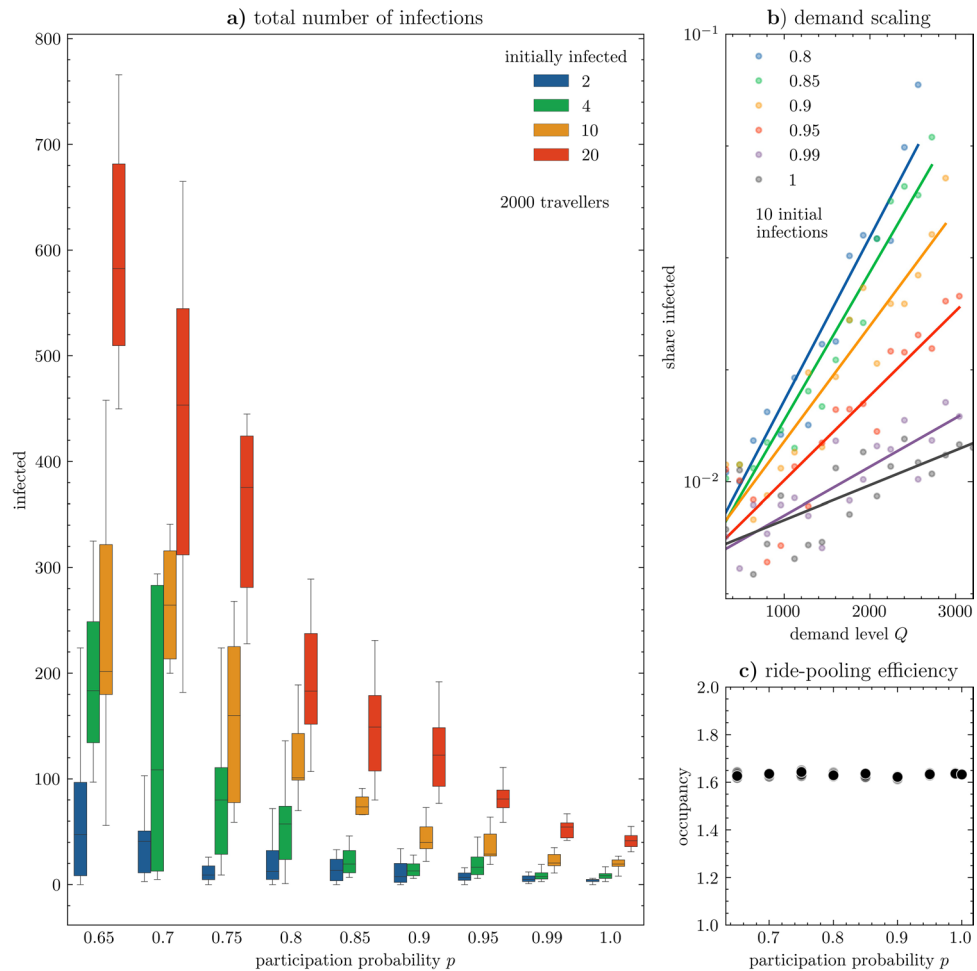
**Figure 3.** Number of infected travellers over the course of epidemic outbreaks, with various settings of initially infected (rows) and demand stability (columns), bold lines denote averages over all experiments (shown individually using thin lines). With an unstable demand (0.65), 20 initially infected always triggers transmission reaching at least 60 travellers (out of 2000) and lasting at least 60 days. Yet in most other configurations results are less stable and actual outbreaks strongly depends on the location of initially infected, revealing a strongly heterogeneous structure of the underlying contact network. In most cases we can observe a smooth, log-normal shape with a strong outbreak in the first phase, exponentially decaying in the latter phases. Mean temporal profiles of outbreaks are consistently following the trend of decreasing when the number of initial infections is lower and a demand pattern becomes more stable. For stable demand patterns, we can observe that the number of infected drops when initially infected quarantine, followed by a smooth transmission in the second phase when demand still fluctuates ( $p = 0.85$ ) or halted immediately ( $p > 0.95$ ). Typically, stabilising the demand halts the epidemic faster. For  $p > 0.85$  epidemic is over in less than 50 days, while for  $p = 0.65$  it can remain active after 100 days (regardless number of initially infected). Despite a clear and strong trend, some simulated outbreaks do not follow the same patterns. We can observe for example an exceptionally high number of infections for  $p = 0.8$  starting from 2 infections when a highly connected hub got infected; quickly halted spreading from 10 infections at  $p = 0.65$ ; or second wave at  $p = 0.75$  and 10 initial infections.

left). Such fast and prevalent spreading can be attributed to a gradually evolving contact network, where each additional day may bring opportunities to be pooled with a new set of travellers, extending the accumulated contact network (Fig. 5a). Consequently, despite having a low mean node degree (i.e. sharing with few travellers at the time), some travellers become hubs, spreading the virus to over 10 travellers, resulting with a long tail of the transmission distribution (Fig. 5b). With low demand stability even two infections may spatially penetrate to all parts of the Amsterdam area, whereas for stable demand the virus may be contained spatially and not spread beyond its original community (Fig. 6).

**Scaling for the demand level.** We experiment with changing the demand levels, gradually increasing it from 100 to 3000 travellers (Fig. 4b). For low demand levels the virus cannot spread since the potential shareability network remains disconnected, i.e. with no giant component (few matches between sparse travellers are found and trips are rarely pooled). Thus, below the critical mass of ride pooling, stabilizing demand has a limited impact. However, as soon as the shareability network includes a larger number of connections (thanks to more compatible trip groups in the demand set) spreading is triggered and the importance of controlling becomes evident.

The relation between the share of infected individuals  $n_i$  and the demand level  $Q$  can be fitted with an exponential function  $n_i = A \exp(\alpha Q)$ , allowing to make predictions about the number of people reached by the virus for higher values of  $Q$  than those explored in our study. For all values of  $p$  shown in Fig. 4b we get high statistical significance of coefficients  $A$  and  $\alpha$  (in all cases  $p$ -value  $< .001$ ), with the coefficient of determination  $R^2$  of 0.92, 0.93, 0.87, 0.81, 0.73 and 0.52 for demand stabilities  $p$  of 0.8, 0.85, 0.9, 0.95, 0.99 and 1 respectively. For less stable demand spreading is ubiquitous and thus less variable, while for stable demand spreading can still remain contained, leading to significant variability in the results and lower goodness-of-fit.

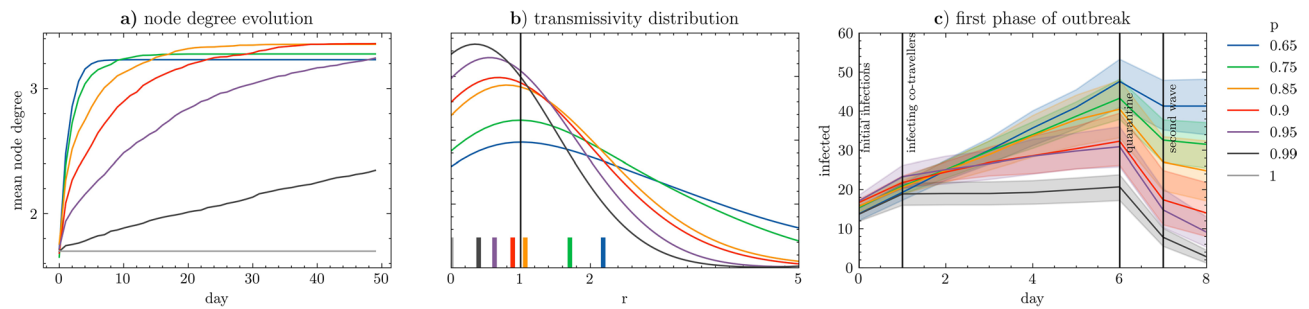
**Ride-pooling efficiency.** Ride-pooling needs a critical mass of demand to become efficient and sustainable. We report ride-pooling efficiency by means of the average occupancy  $o$ , i.e. ratio of passenger-kilometer hours to vehicle kilometer hours. In line with previous studies<sup>9</sup>, we find that occupancy is a function of demand



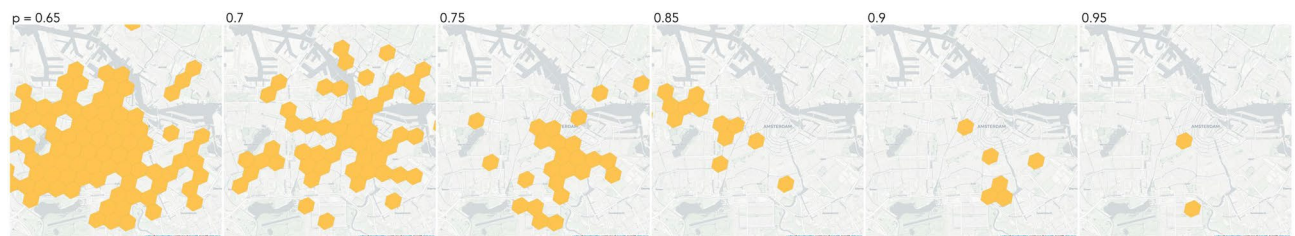
**Figure 4.** (a) Number of eventually infected travellers for varying demand stability  $p$  and initially infected travellers. Distributions based on 20 replications (mean within interquartile box and whiskers from min to max values). Initially infected 20 travellers may spread the virus to almost 40% of the population (800 out of 2000 travellers). Yet as long as demand becomes stabilised, outbreaks start being contained. Even a large number of initially infected does not reach more than 10% of the population if demand stability is set to 90% and is eventually contained below 60 travellers (3%) for fully stable demand. The variability of outbreaks also decreases as the demand stabilises: 10 initial infections may reach between 40 and 100 travellers if  $p = 0.9$ , while the range expands from 50 to over 400 for  $p = 0.7$ . The lower bound increases when the number of initially infected is high, making outbreaks more predictable, unlike the ones starting from a small number of infections, for which variability is greater. Importantly, stabilizing the demand does not reduce the efficiency, as we report in (c), where the mean occupancy (key efficiency indicator of ride-pooling) remains stable as demand stabilizes. Notably, the importance of demand stabilization increases with the demand level as we demonstrate on panel (b) which shows share of infected travellers changing with a demand for various  $p$ 's. Each dot is the average from 20 replications. For all the values of  $p$  share of infected individuals scales with the demand level as  $A \exp(\alpha Q)$ , marked as trendlines on (b), with  $p$ -values and  $R^2$  reported in the text. For demand levels below 1500, the virus rarely reaches more than 4% of the travellers. In contrast, when the demand level is 2500, the epidemic may reach up to 10% when  $p = 0.8$  or be contained below 2% for a stable demand ( $p > 0.9$ ), which underlines the importance of control measures for higher demand levels.

levels, yet, notably, stabilizing the demand with our control parameter  $p$  does not affect it. As long as the same number of travellers participates in pooling everyday, the efficiency remains more or less stable, as we report in Fig. 4c, where 2000 travellers participating daily in the system yield the same occupancy regardless of the replication (dots) and demand stability (x-axis).

**Control and mitigation.** Results show that the virus may easily spread through the ride-pooling networks infecting the majority of the population with a low epidemic threshold (20 initial spreaders may infect up to 800 out of 2000 travellers). While the ride-pooling service provider cannot control for the initial share of infected traveller, nor the incubation and recovery periods, the ride-pooling demand may be controlled to mitigate the virus spreading. Specifically, we show that imposing fixed matches by means of a more stable demand level - solely by controlling for  $p$ , without making amendments to the matching algorithm itself - can mitigate the



**Figure 5.** (a) Average node degree in the evolving contact networks. Regardless demand stability  $p$ , an average traveller is linked to 1.7 other travellers each day. Yet if the demand is unstable it evolves, after 10 days it reaches 1.9 if  $p = 0.99$ , 2.5 if  $p = 0.95$  and goes beyond 3 if  $p < 0.8$ . (b) Mean transmission rate  $r$  (number of new infections per infected) distributions. The long tails for low demand stability reveal the super-spreaders (transmitting to 5 and more travellers). For a stable demand initial infections does not manage to transmit a disease effectively, eventually reducing transmissivity below 1 when  $p > 0.9$ . (c) Insights into the first phase of the epidemic outbreak in the case of 10 initial infections. When first infected travellers are diagnosed after 7 days, their accumulated contact network may vary from 18 to over 60 infected travellers. If contact tracing and mitigation strategies are put in place, already infected travellers may be identified and quarantined before the second outbreak after day 7.



**Figure 6.** An illustration of the spatial extent of epidemic outbreaks originating from two initial infections. A major part of Amsterdam becomes infected for spontaneous demand (left), while it remains spatially contained as the demand stabilises (right). For stable demand ( $p > 0.8$ ) the geographical boundaries are confined, while otherwise, the virus crosses the river IJ and reaches also the north parts of Amsterdam.

spreading and bring it to halt (Figs. 3, 4, 5, 6). We demonstrate that matching and its stability is key to halt epidemics in ride-pooling networks. It can be used proactively in the design of the real-time matching algorithm.

Moreover, if contact tracing apps are used, when a traveller is diagnosed not only s/he has to quarantine, but also his/her traced contact network over the relevant period of time can be identified and eventually isolated. As we show in Fig. 5c, 10 initial infections will spread to a maximum of 60 travellers prior to diagnosis, which seems to be feasible to trace back, isolate and halt spreading.

**Study limitations and caveats.** We aim at revealing the universal patterns characterising the spreading of a virus in ride-pooling networks, yet our findings shall be considered with caveats. Namely, we simulate only a subset of daily mobility patterns (afternoon commute), from many sources of non-recurrence present in travel patterns we picked-up one (participation probability), which we found sufficient to reproduce its impact, while role of others (varying and fluctuating travel modes, destinations, departure times etc.) may be similar or potentially even stronger. While the shareability network in the morning commute is likely to be similar (inverse of afternoon), other, non-commuting trips, will likely yield a different shareability network, catalysing the spreading to the new co-travellers. Nonetheless, our results are valid for systems with regular users with symmetric demand patterns in the morning and afternoon. Moreover, drivers are assumed not to be spreaders (which seems plausible in the context of ad-hoc made shields isolating many of ride-sourcing drivers from travellers). We applied a fixed and deterministic epidemiological model in terms of the infection probability, incubation period and quarantining, since reliable estimates of those parameters distribution are not reported yet. Despite, we claim that the main message holds true for general urban networks: without intervention ride-pooling significantly contributes to virus spreading, while fixing matches between co-travellers dramatically reduces transmission.

## Discussion and conclusions

Sharing a single vehicle with co-travellers during pandemics induces a risk to become exposed to viruses. Sadly, the risk extends beyond the fellow travellers one shares the ride within a single vehicle, mainly due to the accumulated contact graph resulting from day to day variations. Regardless of the number of initial infections, the

upper bound of outbreaks in spontaneous ride-pooling networks is high. Even two initial infections may lead to hundreds of cases across the network. We did not observe spatial nor topological limits to the spreading and disease starting from only two initial infections managed to reach most parts of the Amsterdam's area.

In plausible demand and behavioural settings, if a generic ride-pooling system reaches a critical mass, the travellers become densely connected through the shareability network so that the virus transmits easily through the giant component without clearly visible epidemic thresholds. Only travellers belonging to isolated communities are left unaffected. The pace at which it spreads, however, is low, requiring a long time until virus penetrates across the network. Nonetheless, the daily contact network with its low node degree and hub-free, will evolve due to spontaneity in the demand patterns. If each day, a slightly different pool of travellers decides to travel, this will yield a new matching and resulting with a new shareability, contributing to the accumulated number of contacts steadily growing over time. The slow pace evolution of contact networks becomes beneficial when tracking measures are applied and we can trace past co-travellers for each diagnosed spreader. Even in a highly spontaneous ride-pooling networks, 10 initial infections manage to transmit to only 60 within the 7 day incubation period, which seems to be feasible to trace, isolate and halt the spreading, specifically given the app-based operations of the mobility platform, presumably storing travellers' traces anyhow. Otherwise, if not halted early, epidemics may evolve unhampered and randomly. Depending on the location of initial infections, the epidemic may die-out as well as outbreak, making it potentially risky and uncertain.

Notably, we can substantially limit spreading by sacrificing the spontaneity offered by the ride-pooling service. If we enforce the same matching and fix the pools of co-travellers, spreading is efficiently mitigated and even 20 initial infections remain manageable to be contained. With one, clearly controllable parameter we can reduce the outbreak of viruses. If translated to platform operations, this can become an efficient management and control measure, adjustable along with other country-wide pandemic measures. This may contribute to the provision of a safe shared-mobility alternative in the presence of public health fears and risks. Future research may modify the matching algorithm itself so as to favor the matching of travellers that have already travelled together in past rides. Such an approach is expected to allow for virus spreading reductions even when the demand pattern is subject to large day-to-day variations. Finally, if tracing is combined with fixed pooling, the system's safety may further improve, making ride-pooling a promising intermediate mobility solution for the pandemic world.

Although the presented methodology has been illustrated on the case of Amsterdam ride-pooling it is essential to emphasise its general applicability to examine, in a non-invasive way, the likely outcomes of different underlying topologies on the way the virus spreads through the network. In this sense it opens space for discussion of potential alterations of practical ride-pooling systems on one side and theoretical studies on the other one. Although our study considers a limited number of rides, it is widely acknowledged that cities and their properties are connected with scaling laws<sup>40</sup> even if the form and the details of the methodology behind these relations is questioned<sup>41,42</sup>. Typically such laws<sup>21</sup> may associate a certain index  $x$  with city population size  $M$  by an allometric scaling  $x \sim M^\alpha$ . More importantly the existence of scaling laws has also been proven in the case of ride-sharing networks with respect to shareability<sup>12</sup>, visitation frequency<sup>13</sup> or lately ride-sharing efficiency<sup>33</sup>. In view of this we may assume that our results regarding the number of infected individuals as a function of the demand level  $Q$  should hold for larger systems emphasising the necessity to stabilize the demand.

## Methods

**Travel demand data.** We run a series of experiments on a travel dataset available for Amsterdam from a nation-wide activity-based model<sup>35</sup>, with a single trip defined as a combination of its origin  $o_i$ , destination  $d_i$  and desired pick-up time  $t_i^p$ :

$$Q_i = (o_i, d_i, t_i). \quad (1)$$

The dataset contains over 240 thousand trips conducted within the boundaries of Amsterdam on a representative working day, which we filter to afternoon (2PM–6PM) trips longer than one kilometer. We use 3200 passenger trip requests for the experiments, 2000 of which participates in the pooling on any given day. The pool of travellers from which we sample the daily demand is controlled using  $p$ , based on which each day we draw from the pool of  $2000/p$  travellers.

**Ride-pooling algorithm.** To identify attractive pooled rides we use the EXMAS<sup>9</sup> algorithm (publicly available python library), which for a given network (osmnx graph), travel demand, behavioral parameters (like willingness-to-share) and system parameters (pooling discount) identifies all feasible shared rides and then constructs a shareability network (Fig. 2a) to finally optimally match trips into shared rides (Fig. 2b).

It generates the so-called shareability network, linking two kinds of nodes: travellers and rides. Traveller  $i$  is linked to a feasible ride  $r$  if and only if s/he finds it attractive, which we express as the probability that ride utility  $U_{i,r}$  - reflecting the extent to which delays and detours  $\delta_{r,i}$  imposed by sharing are compensated by a discounted ride fare  $\lambda$  under traveller's behavioural parameters  $\beta_i$  (value of time and willingness-to-share) - is greater than travellers' attractiveness threshold  $\epsilon_i$ . The theoretical number of shared-rides explodes combinatorically with the number of travellers (e.g. 2000 travellers can be matched into  $4.65 \times 10^{20}$  theoretically feasible trips shared by up to five passengers). This can be made tractable by considering only attractive rides, which is governed on one hand by travellers preferences  $\beta_i$  (i.e. individual trade-offs between longer ride and discounted price) and on the other hand by service design  $\lambda$  (controlled through the discount offered by the platform for sharing) and  $\epsilon_i$  expressing the quality of alternatives for ride-pooling (private ride-hailing, or public transport and/or bike), further detailed in<sup>9</sup>. Importantly, the shareability network is composed of feasible rides only, expressed with  $F_r$ , being one if the ride is attractive for all travellers sharing it and zero otherwise. We formalize the shareability network with a link formation  $l_{i,r}$  formula, combining ride feasibility and attractiveness as follows:



$$I_{i,r} = F_r \cdot \Pr(U_{i,r} = U(i, \beta_i, \delta_{r,i}, \lambda) > \epsilon_i) \quad (2)$$

**Matching travellers to attractive shared rides.** Each traveller may be linked to multiple rides and the resulting shareability network is typically highly connected, characterized by the formation of communities and hubs (Fig. 2a). While the shareability network denotes the potential to share a ride, on any given day travellers are matched to exactly one particular shared-ride (Fig. 2b).

To address this, we formulate a binary traveller-ride assignment problem, where each traveller  $i$  is unilaterally assigned to a ride  $r$  and the assignment yields the minimal costs. It is formulated as a problem of determining a binary vector  $x_r$ , an assignment variable equal to one if a ride is selected and zero otherwise (eq. 3c). The objective of this deterministic assignment are ride costs  $c_r$ , multiplied by the assignment variable  $x_r$ , aggregated for all rides (eq. 3a).

Such an assignment satisfies the constraint of assigning each traveller to exactly one ride, obtained through the row-wise sum for assignment variable  $x_r$  and traveller-ride incidence matrix  $I_{i,r}$ . The latter is a binary matrix, where each entry is one if ride  $r$  serves traveller  $i$  and zero otherwise (eq. 3b). Eventually, the solution to the problem (eq. 3a) is the subset  $\mathbf{R}^*$  of feasible rides  $\mathbf{R}$  such that  $x_r = 1 \forall r \in \mathbf{R}^*$ . We express the shareability problem as the following program:

$$\min \quad \sum_{r \in \mathbf{R}} c_r x_r \quad (3a)$$

$$\text{subject to} \quad \sum_{i \in \mathbf{Q}} I_{i,r} x_r = 1, \forall i \quad (3b)$$

$$x_r \in \{0, 1\}. \quad (3c)$$

Although matching problem (Eq. 3a) can be read as the set cover problem<sup>43</sup>, which is known to be NP-Hard, real-life ride-pooling situations usually yield configurations managed by standard solvers (like in<sup>8,9</sup>).

**Contact network.** On any given day, the contact graph is composed of connecting each ride to all travellers that have shared (part of) it. Notably, the contact network evolves over time, primarily due to the different pool of travellers being matched on any given day. Hence, this representation allows simulating an epidemic outbreak by analyzing potential transmissions between travellers that have shared rides with other travellers over the course of the analysis period. In our model the contact network changes from day to day due to one or more of the following reasons: (i) infected travellers quarantine (which may catalyse spreading as quarantined travellers are replaced by susceptible ones, who will get infected) (ii) recovered travellers return to the system (which impedes spreading as recovered travellers restore to previous, optimal matches, already penetrated by the virus) or (iii) daily variations in travel demand as travellers decide not to use ride-hailing on a given day (for example because they opt for an alternative mode). We represent the daily participation, central endogenous variable of the model, through the *demand stability* parameter  $p$  in our experiments. Each day we update the pool of travellers (using the daily participation formula  $F_i^d = \Pr(p) \cdot (1 - K_i^d)$  which combines the participation probability  $p$  and quarantined travellers on day  $K_i^d$ ). This, in turn, results with updating the pool of rides feasible on a given day (composed only of travellers present in the daily pool). The contact network will then evolve as travellers are matched to new rides when their co-travellers are quarantined or absent.

**Epidemic model.** We adopt a SIQR model to represent the four compartments characteristic of the COVID-19 pandemic: Susceptible ( $S$ ), Infected ( $I$ ), Quarantined ( $Q$ ) and Recovered ( $R$ ), recently, directly applied to tackle COVID-19 propagation in other studies (Italy<sup>44</sup> and Japan<sup>45</sup>). Following the argumentation of Pedersen and Meneghini<sup>44</sup> we do not explicitly designate the E state, given the evidence suggesting that the COVID-19 virus can be propagated without first exhibiting visual symptoms. The SIQR model was first introduced by Feng and Thieme in 1995<sup>32</sup> and then examined in detail by Hethcote *et al.*<sup>46</sup>. Previous studies focused on mathematical aspects of the model (e.g., oscillations<sup>32</sup>, stability analysis<sup>46</sup> or the role of stochastic noise<sup>47</sup>). While the aggregate epidemiological properties of the SIQR model are well studied, studies taking into account the underlying network structure and its evolution are scarce.

The phenomenon central to this paper is driven by the structure and evolution of the contact network, rather than by the parameters of the epidemic model. We, therefore, adopt a deterministic model where infected travellers infect all of their co-travellers with a probability of 1. For the sake of clarity, unlike SEIR models, we assume that all exposed inevitably become infected, all of which quarantine and recover after certain incubation and recovery periods (we use here the latest reliable findings suggesting, respectively, 7 and 14 for COVID-19<sup>48</sup>). This ubiquitous spread over the contact network may be seen as a pessimistic upper-bound of the spreading process, yet in the view of recent pandemics<sup>15</sup>, sharing a vehicle with infected co-traveller is expected to yield a high contagion risk. Furthermore, the focus of this study is on spreading across the network and over multiple vehicles and rides rather than within vehicle transmission probabilities. Future medical estimates of the latter can be embedded into the analysis performed in this study as soon as those are made available to refine our model specifications and thus obtain more precise estimates. Our findings should therefore be considered an upper-bound of the epidemiological consequences of virus spreading in ride-pooling systems.

**Modelling framework.** The ExMAS ride-pooling algorithm is embedded within the day-to-day loop characteristic to epidemiological model. The simulation initializes with a trip demand set composed of all the travellers that may consider ride-pooling on any given day during the course of the simulated epidemic outbreak, to further allow embedding the participation probability  $p$ . Before entering the main epidemic loop, we identify all feasible pooled rides (Fig. 2a)—to determine potential co-travellers that any given traveller may encounter during the course of an epidemic outbreak. We create the complete shareability graph by applying equation (2) with  $\epsilon$  corresponding to a private, non-shared ride alternative in a deterministic model.

Following this initialisation phase, we then enter the main simulation loop. We start with assigning initial infectors - drawn in random by sampling a pre-defined number of initially infected, which is treated as random input and vary from one replication to another. Next, we enter the day-to-day simulation: every day we first determine the daily ride-pooling demand. We assume that only a subset of travellers actually participates in the ride-pooling system on any given day, i.e. every day we sample a given number of travellers from the total latent demand. We fix the demand to 2000 everyday in our experiments to ease comparisons (except Fig. 4b where we experiment with various demand levels). Those travellers are then matched to identify the realization of shared-rides on a given day. Everyday we apply the SIQR model with transitions taking place when:

- (a) infected travellers infect their susceptible co-riders ( $S \rightarrow I$ ),
- (b) infected travellers are quarantined after the incubation period ( $I \rightarrow Q$ ),
- (c) travellers recover after the quarantine and acquire complete immunity to the virus ( $Q \rightarrow R$ ).

For any given day, the model outputs information about the number of travellers in each state (S-I-Q-R) and newly infected travellers, based on which we can reproduce epidemic spreading profiles. The loop terminates when all the infected travellers are quarantined (there are no active infections).

## Code and data availability

The code to generate the shareability network from a given demand pattern and then to reproduce the epidemic simulations is available at the public GitHub repository (<http://www.github.com/rafalkucharskiPK/ExMAS>). The experimental results data is available under <https://doi.org/10.4121/14140616.v1>. The network data was obtained from Open Street Map with `osmnx`, Amsterdam travel demand was derived from Albatross data set<sup>35</sup>.

Received: 30 October 2020; Accepted: 17 March 2021

Published online: 30 March 2021

## References

1. Acuto, M. *et al.* Seeing COVID-19 through an urban lens. *Nat. Sustain.* <https://doi.org/10.1038/s41893-020-00620-3> (2020).
2. He, G., Pan, Y. & Tanaka, T. The short-term impacts of COVID-19 lockdown on urban air pollution in China. *Nat. Sustain.* <https://doi.org/10.1038/s41893-020-0581-y> (2020).
3. Muller, S. A., Balmer, M., Neumann, A. & Nagel, K. Mobility traces and spreading of covid-19. medRxiv (2020).
4. UITP. Management of Covid-19 guidelines for public transport operators (2020).
5. Gkiotsalitis, K. & Cats, O. Optimal frequency setting of metro services in the age of covid-19 distancing measures. arXiv preprint [arXiv:2006.05688](https://arxiv.org/abs/2006.05688) (2020).
6. Tirachini, A. & Cats, O. Covid-19 and public transportation: current assessment, prospects, and research needs. *J. Public Transp.* **22**, 1 (2020).
7. Guerriero, C., Haines, A. & Pagano, M. Health and sustainability in post-pandemic economic policies. *Nat. Sustain.* 1–3 (2020).
8. Alonso-Mora, J., Samaranayake, S., Wallar, A., Frazzoli, E. & Rus, D. On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. *Proc. Natl. Acad. Sci.* **114**, 462–467 (2017).
9. Kucharski, R. & Cats, O. Exact matching of attractive shared rides (exmas) for system-wide strategic evaluations. *Transp. Res. Part B Methodol.* **139**, 285–310. <https://doi.org/10.1016/j.trb.2020.06.006> (2020).
10. Yang, B. *et al.* Phase transition in taxi dynamics and impact of ridesharing. *Transp. Sci.* **54**, 250–273. <https://doi.org/10.1287/trsc.2019.0915> (2020) (arXiv:1801.00462).
11. Riascos, A. P. & Mateos, J. L. Networks and long-range mobility in cities: a study of more than one billion taxi trips in New York City. *Sci. Rep.* **10**, 1–14. <https://doi.org/10.1038/s41598-020-60875-w> (2020).
12. Tachet, R. *et al.* Scaling law of urban ride sharing. *Sci. Rep.* **7**, 42868. <https://doi.org/10.1038/srep42868> (2017).
13. Chen, X. M., Chen, C., Ni, L. & Li, L. Spatial visitation prediction of on-demand ride services using the scaling law. *Phys. A Stat. Mech. Appl.* **508**, 84–94. <https://doi.org/10.1016/j.physa.2018.05.005> (2018).
14. Santi, P. *et al.* Quantifying the benefits of vehicle pooling with shareability networks. *Proc. Natl. Acad. Sci.* **111**, 13290–13294 (2014).
15. Liu, Y., Eggo, R. M. & Kucharski, A. J. Secondary attack rate and superspreading events for sars-cov-2. *The Lancet* **395**, e47 (2020).
16. Della Rossa, F. *et al.* A network model of Italy shows that intermittent regional strategies can alleviate the COVID-19 epidemic. *Nat. Commun.* **11**, 5106. <https://doi.org/10.1038/s41467-020-18827-5> (2020).
17. Liu, Z. *et al.* Near-real-time monitoring of global CO<sub>2</sub> emissions reveals the effects of the COVID-19 pandemic. *Nat. Commun.* **11**, 5172. <https://doi.org/10.1038/s41467-020-18922-7> (2020).
18. Developing infectious disease surveillance systems. *Nat. Commun.* **11**, 4962. <https://doi.org/10.1038/s41467-020-18798-7> (2020).
19. Pastor-Satorras, R., Castellano, C., Van Mieghem, P. & Vespignani, A. Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**, 925 (2015).
20. Wang, Z. *et al.* Statistical physics of vaccination. *Phys. Rep.* **664**, 1–113. <https://doi.org/10.1016/j.physrep.2016.10.006> (2016).
21. Sienkiewicz, J. & Holy, J. A. Statistical analysis of 22 public transport networks in Poland. *Phys. Rev. E* **72**, 046127. <https://doi.org/10.1103/PhysRevE.72.046127> (2005).
22. Barthélemy, M. Spatial networks. *Phys. Rep.* **499**, 1–101. <https://doi.org/10.1016/j.physrep.2010.11.002> (2011).
23. Gallotti, R. & Barthélemy, M. The multilayer temporal network of public transport in Great Britain. *Sci. Data* **2**, 140056. <https://doi.org/10.1038/sdata.2014.56> (2015).
24. Aleta, A., Hu, Q., Ye, J., Ji, P. & Moreno, Y. A data-driven assessment of early travel restrictions related to the spreading of the novel covid-19 within mainland China. *Chaos Solitons Fractals* **139**, 110068. <https://doi.org/10.1016/j.chaos.2020.110068> (2020).
25. Chinazzi, M. *et al.* The effect of travel restrictions on the spread of the 2019 novel coronavirus (covid-19) outbreak. *Science* **368**, 395–400. <https://doi.org/10.1126/science.aba9757> (2020).

26. Kraemer, M. U. G. *et al.* The effect of human mobility and control measures on the Covid-19 epidemic in china. *Science* **368**, 493–497. <https://doi.org/10.1126/science.abb4218> (2020).
27. Azizi, A., Montalvo, C., Espinoza, B., Kang, Y. & Castillo-Chavez, C. Epidemics on networks: reducing disease transmission using health emergency declarations and peer communication. *Infect. Dis. Model.* **5**, 12–22. <https://doi.org/10.1016/j.idm.2019.11.002> (2020).
28. Xue, L. *et al.* A data-driven network model for the emerging covid-19 epidemics in Wuhan, Toronto and Italy. *Math. Biosci.* **326**, 108391. <https://doi.org/10.1016/j.mbs.2020.108391> (2020).
29. Ciaperoni, M. *et al.* Relevance of temporal cores for epidemic spread in temporal networks. *Sci. Rep.* **10**, 12529. <https://doi.org/10.1038/s41598-020-69464-3> (2020).
30. 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).
31. Volz, E. & Meyers, L. A. Epidemic thresholds in dynamic contact networks. *J. R. Soc. Interface* **6**, 233–241 (2009).
32. Feng, Z. & Thieme, H. R. Recurrent outbreaks of childhood diseases revisited: the impact of isolation. *Math. Biosci.* **128**, 93–130. [https://doi.org/10.1016/0025-5564\(94\)00069-C](https://doi.org/10.1016/0025-5564(94)00069-C) (1995).
33. Molkenhain, N., Schröder, M. & Timme, M. Scaling laws of collective ride-sharing dynamics. *Phys. Rev. Lett.* **125**, 248302. <https://doi.org/10.1103/PhysRevLett.125.248302> (2020).
34. Aslak, U. & Maier, B. F. Netwulf: interactive visualization of networks in python. *J. Open Source Softw.* **4**, 1425 (2019).
35. Arentze, T. A. & Timmermans, H. J. A learning-based transportation oriented simulation system. *Transp. Res. Part B: Methodol.* **38**, 613–633. <https://doi.org/10.1016/j.trb.2002.10.001> (2004).
36. Alonso-González, M. J. *et al.* What are the determinants of the willingness to share rides in pooled on-demand services?. *Transportation* **1–33**, (2020).
37. Alonso-González, M. J., van Oort, N., Cats, O., Hoogendoorn-Lanser, S. & Hoogendoorn, S. Value of time and reliability for urban pooled on-demand services. *Transp. Res. Part C: Emerg. Technol.* **115**, 102621 (2020).
38. Li, W., Pu, Z., Li, Y. & Ban, X. Characterization of ridesplitting based on observed data: a case study of Chengdu, China. *Transp. Res. Part C: Emerg. Technol.* **100**, 330–353. <https://doi.org/10.1016/j.trc.2019.01.030> (2019).
39. Veve, C. & Chiabaut, N. Estimation of the shared mobility demand based on the daily regularity of the urban mobility and the similarity of individual trips. *PLoS ONE* **15**, 1–15. <https://doi.org/10.1371/journal.pone.0238143> (2020).
40. Bettencourt, L. M. A., Lobo, J., Helbing, D., Kühnert, C. & West, G. B. Growth, innovation, scaling, and the pace of life in cities. *Proc. Natl. Acad. Sci.* **104**, 7301–7306. <https://doi.org/10.1073/pnas.0610172104> (2007).
41. Leitão, J. C., Miotto, J. M., Gerlach, M. & Altmann, E. G. Is this scaling nonlinear?. *R. Soc. Open Sci.* **3**, 150649. <https://doi.org/10.1098/rsos.150649> (2016).
42. Altmann, E. G. Spatial interactions in urban scaling laws. *PLoS ONE* **15**, 1–12. <https://doi.org/10.1371/journal.pone.0243390> (2020).
43. Raz, R. & Safra, S. A sub-constant error-probability low-degree test, and a sub-constant error-probability pcp characterization of np. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, STOC '97, 475–484. <https://doi.org/10.1145/258533.258641> (Association for Computing Machinery, New York, NY, USA, 1997).
44. Pedersen, M. G. & Meneghini, M. A simple method to quantify country-specific effects of covid-19 containment measures. *medRxiv* <https://doi.org/10.1101/2020.04.07.20057075> (2020). <https://www.medrxiv.org/content/early/2020/04/10/2020.04.07.20057075.full.pdf>.
45. Odagaki, T. Analysis of the outbreak of Covid-19 in Japan by SIQR model. *Infect. Dis. Model.* **5**, 691–698. <https://doi.org/10.1016/j.idm.2020.08.013> (2020).
46. Hethcote, H., Zhen, M. & Shengbing, L. Effects of quarantine in six endemic models for infectious diseases. *Math. Biosci.* **180**, 141–160. [https://doi.org/10.1016/S0025-5564\(02\)00111-6](https://doi.org/10.1016/S0025-5564(02)00111-6) (2002).
47. Cao, Z., Feng, W., Wen, X., Zu, L. & Cheng, M. Dynamics of a stochastic SIQR epidemic model with standard incidence. *Phys. A: Stat. Mech. Appl.* **527**, 121180. <https://doi.org/10.1016/j.physa.2019.121180> (2019).
48. Lauer, S. A. *et al.* The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: estimation and application. *Ann. Intern. Med.* **172**, 577–582 (2020).

## Acknowledgements

This research was supported by the CriticalMaaS project (Grant Number 804469), which is financed by the European Research Council and Amsterdam Institute for Advanced Metropolitan Solutions and the Transport Institute of TU Delft. This research was funded by National Science Centre in Poland program OPUS 19 (Grant Number 2020/37/B/HS4/01847) and by IDUB against COVID-19 project granted by Warsaw University of Technology under the program Excellence Initiative: Research University (IDUB).

## Author contributions

R.K. and O.C. conceived the experiment(s), R.K. conducted the experiment(s), R.K O.C and J.S. analysed the results. All authors reviewed the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to R.K.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://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