Submitted to the Annals of Applied Statistics

LEVERAGING CELLPHONE-DERIVED MOBILITY NETWORKS TO ASSESS COVID-19 TRAVEL RISK

BY JUSTIN J. SLATER[1,a], PATRICK E. BROWN[2,b]
JEFFREY S. ROSENTHAL[2,c] AND JORGE MATEU[3,d]

[1]*Department of Mathematics and Statistics, University of Guelph ,* [a]*jslate04@uoguelph.ca*

[2]*Department of Statistical Sciences, University of Toronto ,* [b]*patrick.brown@utoronto.ca;* [c]*jeff@math.toronto.edu*

[3]*Department of Mathematics, University Jaume I of Castellon ,* [d]*mateu@mat.uji.es*

Since the beginning of the Covid-19 pandemic, public health authorities across the globe have implemented policies, such as lockdowns, in an attempt to reduce population mobility, and consequently, person-to-person contacts. It is well known that lockdowns reduce mobility, but to what extent does this reduction in mobility lead to lower infection rates? In this paper, we extend the endemic-epidemic modeling framework in a principled manner, incorporating temporally changing mobility network data and quantifying the risk associated with travelling throughout the first year of the pandemic in two Spanish Communities.

1. Introduction. The relationship between mobility and Covid-19 is of utmost importance, as mobility-reducing policies such as lockdowns and travel restrictions are often used to thwart the spread of such infectious diseases. Countless studies have attempted to quantify the effectiveness of mobility reductions by using a variety of data sources and statistical methods. Cellphone-derived mobility data is well-suited for this purpose, as we can use it to quantify the severity of a lockdown as well as relate it to case counts via a statistical model such as a generalized linear model or infectious disease model. Slater et al. (2021) showed that mobility data better captures spatial heterogeneity in Covid-19 case counts than spatial proximity in Bayesian spatial models. Furthermore, cellphone-derived mobility data can capture changes in the temporal dimension, a limitation in existing mobility models for infectious disease surveillance data (Meyer and Held, 2014). However, the temporal relationship between mobility and case counts poses great modeling challenges, as the correlation between the two changes in each wave, with high correlation in the first wave, and little/negative correlation in subsequent waves (Gottumukkala et al., 2021). We argue that since mobility affects the reproduction rate of infectious diseases (as opposed to the absolute counts), we can indeed infer the impact of mobility on case counts using a spatio-temporal infectious disease model.

In the last two decades, a class of infectious disease models known as *endemic - epidemic* models have gained popularity due to their simplicity and forecasting ability (Held, Höhle and Hofmann, 2005). A simple version of these models can be written as:

$$Y_t|Y_{t-1} \sim \text{Pois}(\lambda_t)$$
$$\lambda_t = \omega + \alpha Y_{t-1}$$

where Y_t is the number of cases at time t, ω is the "endemic" component which describes new cases that are not explained by previous cases, and αY_{t-1} is the "epidemic component" which describes new cases that are directly attributable to previous cases. These models have

Keywords and phrases: cellphone mobility data, Covid-19, infectious diseases, endemic-epidemic models, poisson network autoregressions, spatio-temporal statistics.

since been extended to include temporally changing α (Held et al., 2006), multiple diseases (Paul, Held and Toschke, 2008), random effects (Paul and Held, 2011), seasonal effects (Held and Paul, 2012), serial interval distributions of disease (Bracher and Held, 2020) and more. Endemic-epidemic models overcome the computational difficulties of fitting classic compartmental (SIR) models, and are an attractive alternative when an abundance of data is available (Wakefield, Dong and Minin, 2019).

An example of a multivariate, or in the context of this paper, multi-region endemic-epidemic model is

$$\begin{aligned} Y_{it}|\boldsymbol{Y}_{t-1} &\sim \text{Pois}(\lambda_{it}) \\ \lambda_{it} &= \omega_{it} + \alpha \sum_j v_{ji} Y_{j,t-1} \end{aligned} \tag{1}$$

where Y_{it} is the number of cases in region i at time t, i and j are region indicators, and v_{ji}'s represent (potentially asymmetric) weights between regions j and i. Typically these weights are row-normalized (sum to 1) but this is not necessary. Some common weights are functions of physical distance or proximity such as those suggested in Paul, Held and Toschke (2008)

$$v_{ji} = \frac{1}{\text{Ne}(j)}$$

or those suggested in Meyer and Held (2014)

$$v_{ji} = (o_{ji} + 1)^{-\rho}$$

where $\text{Ne}(j)$ is the number of regions sharing a border (neighbors) with region j, o_{ji} is the minimum number of region borders you would have to cross to get from region j to i, and ρ is a parameter to be estimated . These weights tend to work well because they are good proxy for the number of people moving between regions, and resultingly, contact rates between infectious and susceptible people. More interestingly, these weights have been combined with or replaced by other data sources to more accurately estimate the contact rates between individuals of different regions. For instance, Schrödle, Held and Rue (2012) used assymetric mobility weights to model the spread of Coxiellosis in Swiss cows. Geilhufe et al. (2014) used mobility data to estimate the relationship between distance and mobility, and define their weights based on this relationship. Meyer and Held (2017) estimate contact rates between age groups using external data and combine these data with spatial proximity weights and used this as an estimate for contact rates between age groups across various regions. Fritz and Kauermann (2022) build weights based on estimated social connectedness via social media data. Grimée et al. (2022) created temporally changing weights by combining border closure, proximity, and mobility data to assess the effectiveness of lockdowns during the Covid-19 pandemic, and estimate case counts under counterfactual scenarios. Celani and Giudici (2022) incorporated mobility weights to assess the effectiveness of containment measures in Italy. Each of these works show that proximity weights can be supplemented or replaced with external data to improve forecasting or inference.

Much of the methodological progress surrounding endemic-epidemic models aims to improve forecasting ability based on the framework presented in Gneiting and Raftery (2007). Consequently, the applications of these models tend to lack interpretability. When the goal is learning about the biological properties of an infectious disease, we must make every effort to ensure our model parameters have clear meanings, and that our results are biologically plausible. Covariates introduced should be done so carefully, and should effect model parameters in a way that are consistent with infectious disease dynamics.

In this paper, we derive a mobility extended spatio-temporal endemic-epidemic model where contact rates are a temporally changing function of mobility. In doing so, we ensure

interpretability of our important parameters, and carefully specify the functional form of the reproduction number via data exploration methods. We use this model to infer the risk associated with travelling during the first 12-15 months of the Covid-19 pandemic in two Spanish Autonomous Communities using high resolution areal mobility networks derived from cellphone GPS signals.

This paper is structured as follows. We introduce the data that motivated this work in Section 2, and present our model and methods in Section 3. In Section 4, we apply our model to two Spanish Communities, inferring the risk associated with travelling in both. We end with a discussion of our model results, limitations, and future work.

2. Data. This paper focusses on Madrid and Castilla-Leon, two Communities in Spain. Madrid, with a population of approximately 6.8 million, is home to Madrid City, the capital of Spain. Castilla-Leon is geographically the largest Community in Spain, with a population of 2.5 million, and is thus much more rural than the Community of Madrid. Each community is divided into smaller subregions (Madrid has 179 subregions, Castilla-Leon has 245), which are depicted in Figure 4. We obtain weekly mobility network data for the trips between and within each of these subregions, alongside Covid-19 cases. The mobility data used in this paper is a temporal extension of that used in Slater et al. (2021), where it was shown that static mobility networks improved the fit of classic spatial models for Covid-19 in two Spanish communities. These data can be downloaded from Ministerio de Transportes Movilidad Y Agends Urbana, Gobierno de España (2022), and are described in detail in Ponce-de Leon et al. (2021). Although daily mobility data is available, we aggregated it by week to match the resolution of the case data, avoiding the well-known day-of-the-week effect of Covid-19 case reporting (Slater, Brown and Rosenthal, 2021).

For Castilla-Leon, the weekly case data from March 1, 2020, to March 7, 2021 was obtained from the open data portal of Castilla-Leon (General Directorate of Information Systems, Quality and Pharmaceutical Provision, 2022). For Madrid, case data from March 1, 2020, to May 9, 2021 was obtained from Epidemiological Surveillance Network of Madrid (Epidemiological Surveillance Network of Madrid, 2022). For both Communities, cases were identified using PCR tests up until October 7, 2020, after which antigen tests were also utilized. Note that nearing the end of our Madrid case data, vaccines were being administered to the public, and thus should be accounted for. Country level vaccine data was obtained from Ministerio de Sanidad, Gobierno De España (2022a), where by the end of our study period, 28.7% of the population had received at least one dose (Pfizer, Moderna, AstraZeneca, Janssen), while 11.5% were fully vaccinated (received two doses). Although we don't expect this to have a substantial impact on our results, the effect of vaccines should at least be explored.

Daily testing data was acquired from the Government of Spain Ministry of Health website (Ministerio de Sanidad, Gobierno De España, 2022b). Testing data was only available at the community level, which we aggregated by week.

Aggregated mobility, case count, and testing data are shown in Figure 1. Mean daily mobility for the study period is depicted in Figure 2.

3. Methodology. In this section, we start by introducing a single-region version of a mobility-extended endemic-epidemic model, a derivation inspired by Bauer and Wakefield (2018). We then extend this model to a multi-region model, and describe reasonable assumptions that make implementing this model computationally feasible. We then describe the careful processes of accounting for delayed reporting, serial intervals, and underreporting, while retaining interpretability of our model. We conclude the section with an explanation of the summary statistics used in this paper, followed by our inference methodology.

3.1. *Single region model.* In epidemiology, the force of infection at time t, λ_t, is defined as the rate at which susceptible individuals become infected. Mathematically, we write it as (Halloran, Longini and Struchiner, 2010)

$$\lambda_t = \mathcal{C}_{t-1} \times \mathcal{P}_{t-1} \times \frac{I_{t-1}}{N}$$

where $\mathcal{C}_{t-1}$ is the rate of contacts between an infectious person and susceptibles individuals at time $t-1$, $\mathcal{P}_{t-1}$ is the probability of infection given a contact between an infectious and susceptible individual, I_{t-1} is the number of infectious individuals at time $t-1$, meaning that $\frac{I_{t-1}}{N}$ is the prevalence at time $t-1$ (where N is the population). For simplicity, we will assume $I_{t-1} = Y_{t-1}$, that the number of infectious individuals equals the number of cases, but will later relax this assumption. Typically, the number of contacts is assumed to be constant (*frequency dependent*) or proportional to the population size N (*density dependent*). Distinguishing between these is inconsequential in our models as we will see later on. However, in this paper, we assume C_t is a function of mobility, w, which is the number of trips as described in Section 2. That is, we assume that the contacts function takes the form

$$\mathcal{C}_{t-1} = c_0 + \sum_{d=1}^{D} c_d w_{t-d}$$

where the $c_0, ..., c_D$ are unknown constants and D is some small integer (i.e 1,2 or 3) chosen by the analyst. If we assume a serial interval of one week (relaxed in later sections), the cases this week, Y_t, were caused by cases in the previous week, Y_{t-1}. However, the reason for including higher lags of mobility is because it is possible that a case that appeared in the quantity Y_t should have appeared in Y_{t-1} due to a case at Y_{t-2}, but was delayed because they didn't immediately produce a positive test. We can't differentiate between a reporting delay and a longer serial interval so we hope that inclusion of Y_{t-2} in our model will account for both. Mobility data included in this way has been shown to improve forecasting ability of univariate endemic-epidemic models (Douwes-Schultz et al., 2022). If we assume that the per-contact probability of infection is time constant $\mathcal{P}_{t-1} = p$ (an assumption that could be relaxed based on available data), then our force of infection is

$$\begin{aligned}
\lambda_t &= \left(c_0 + \sum_{d=1}^{D} c_d w_{t-d}\right) \times p \times \frac{Y_{t-1}}{N} \\
&= c_0 p \frac{Y_{t-1}}{N} + \left(\sum_{d=1}^{D} c_d p \cdot w_{t-d}\right) \frac{Y_{t-1}}{N} \\
&= \alpha^{AR} \frac{Y_{t-1}}{N} + \left(\sum_{d=1}^{D} \alpha_d^{\text{mob}} w_{t-d}\right) \frac{Y_{t-1}}{N}
\end{aligned}$$

where $\alpha^{\text{AR}} = c_0 p$ and $\alpha_d^{\text{mob}} = c_d p$ are parameters representing the "autoregressive" and "mobility" components of the model, and will be estimated from data. Note that we are estimating the product $\alpha = pc$, so assuming time constant p and c leads to time constant α's. Given that the first major variants didn't start to appear until early 2021 (Alpha/Beta/Gamma, and later Omnicron), a time constant p seems reasonable, as disease characteristics likely didn't change much in this time period. If this model were extended well into 2021, we could allow the α's to change as variants are introduced. If we make the additional assumption that infected people are equally likely to move as the rest of the population, then α_d^{mob} can be interpreted as the number of infections at time t per infected trip at time $t-d$, and α^{AR} is the number of new infections at time t per infection at time $t-1$, but not related to mobility. This assumption

may not be as problematic as it may sound, as people can be infectious several days prior to showing any symptoms (He et al., 2020), and thus likely not to change their behavior in this time.

Bauer and Wakefield (2018) show that when the number of susceptibles $\approx N$, the number of infections at time t is approximately Poisson distributed

$$Y_t|Y_{t-1} \sim \text{Pois}(\lambda_t).$$

Furthermore, it is common, and mathematically convenient, to assume that there is some number infections, α^{EX}, that come from outside the region, not related to the previous cases Y_{t-1}. In doing so, we arrive at an extension of the univariate endemic-epidemic model (Held, Höhle and Hofmann, 2005)

$$\begin{aligned} Y_t|Y_{t-1} &\sim \text{Pois}(\lambda_t^\dagger) \\ \lambda_t^\dagger &= \alpha^{\text{EX}} + \alpha^{\text{AR}} Y_{t-1} + \Big(\sum_{d=1}^{D} \alpha_d^{\text{mob}} w_{t-d}\Big) Y_{t-1} \end{aligned} \tag{2}$$

This model can be thought of as a branching process with immigration, with reproduction number, $\alpha^{\text{AR}} + \sum_{d=1}^{D} \alpha_d^{\text{mob}} w_{t-d}$, that linearly depends on mobility, and an immigration of α^{EX}. This implies that mobility only effects the reproduction rate of the disease, and does not relate directly to the case counts. This is an attractive property of this model, as mobility can only cause infections in the next generation if infectious people from the previous generation move around. This is because our model assumes that such movements are directly related to the number of contacts between susceptible and infectious individuals, where each contact will lead to an infection with probability p. If the effect of mobility is small, then the reproduction number will be almost entirely described by the constant α^{AR}. In other words, α^{AR} can be thought of as an autoregressive term that relates previous cases to current cases, or it can be thought of as the intercept in the line relating the reproduction number to mobility.

It should be noted that the models in this paper assume that the reproduction number is independent of observed cases, given mobility. We justify this decision as follows. There are two main ways that observed cases may impact the reproduction number. The first way is through behaviour change, in that people will likely change their behaviour if they hear that the case count is rising. However, we expect that this behaviour change will (at least somewhat) be captured by changes in mobility. For example, companies encouraging employees to work from home will likely lower the reproduction number, but this is because it is reducing the number of person-to-person contacts, which should be reflected in a reduction in mobility. The second way that the observed case counts can impact reproduction numbers is through heard immunity, in that some fraction of the population will no longer be susceptible for some period of time. However, during the duration of our study, this fraction is small (0 to 10%) and thus including additional parameters to account for current prevalence in the reproduction number would likely do very little for improving model fit.

α^{EX} represents an influx of cases caused by infectious people outside our dataset infecting susceptibles in our region. When regions are large, this number should be relatively small. Including α^{EX} serves to prevent our branching process from dying out, which will be especially helpful in the multi-region case when there are small subregions with low amounts of mobility. In some applications, this component is appropriately referred to as an "endemic" component, as it may describe predictable yearly fluctations/periodicities in disease incidence. However, even cases that arise in an "endemic" are often still attributable previous cases, but with a more predictable/periodic pattern, and can be thought of as the "background rate of disease" (Gordis, 2013). This interpretation is common in endemic-epidemic models and models that predated them (Knorr-Held and Richardson, 2003). Covid-19 had not

yet reached endemic status, thus estimating the background rate of infection is challenging. Thus we believe the term "exogenous" is more appropriate for our application, and should be viewed as factors influencing the absolute number of cases in a region as opposed to the infectiousness of the disease.

Note that is is common in the endemic-epidemic literature to use a negative binomial likelihood to account for conditional overdispersion. However, this is done purely for practical reasons, as this distribution does not arise in the derivation from a discrete-time SIR model (Wakefield, Dong and Minin, 2019). We consider negative binomial models briefly in the application section of this paper.

An alternative way to view model (2) is using the competing risks framework as in Bauer and Wakefield (2018). That is, we can view the exogenous, autoregressive, and movement terms as their own individual Poisson process, and the total force of infection, indicated by †, is the sum of three Poisson random variables with mean

$$\lambda_t^\dagger = \lambda_t^{\text{EX}} + \lambda_t^{\text{AR}} + \lambda_t^{\text{mob}}.$$

In other words, a susceptible can be infected in one of three ways, all with some positive probability. Furthermore, we can compute the proportion of cases attributable to movement (PCAtM) at time t as $\frac{\lambda_t^{\text{mob}}}{\lambda_t^\dagger}$. We will use this measure and its associated uncertainty to assess the association between mobility and infection. We will now extend our model to the multi-region case.

3.2. *Multi-region model.* Now that we are dealing with more than one geographic region (245 regions for Castilla-Leon, 179 for Madrid), we must define a region-wise force of infection. The force of infection, λ_{jit} is defined as the rate at which infectious individuals in region j, infect susceptible individuals in region i, at time t. Similar to the univariate case, we can write this mathematically as

$$\begin{aligned}
\lambda_{jit} &= \mathcal{C}_{ji,t-1} p_{ji,t-1} \frac{y_{j,t-1}}{N_j} \\
&= (c_{ji,t-1} p_{ji,t-1} + \sum_{d=1}^{D} c_{ji,t-d}^{\text{mob}} p_{ji,t-1} w_{ji,t-d}) \frac{y_{j,t-1}}{N_j} \\
&= (\alpha_{jit} + \sum_{d=1}^{D} \alpha_{ji,t-d}^{\text{mob}} w_{ji,t-d}) \frac{y_{j,t-1}}{N_j},
\end{aligned}$$

where $\alpha_{jit} = c_{ji,t-1} p_{ji,t-1}$ is number of cases in region i attributed to a single case in region j that is not accounted for by mobility, p_{jit} is the (potentially time-varying) probability that an infectious individual in region j infects an individual in region i, and $w_{ji,t-d}$ is the number of trips from region j to region i at time $t-d$. $\alpha_{jit}^{\text{mob}} = c_{jit}^{\text{mob}} p_{jit}$ is the of number cases in region i caused by an infected trip from region j to i. As is, the number of model parameters grow at a rate of $O(I^2 \times T)$ where I is the number of regions and T is the number of time points. Given that we will be dealing with hundreds of subregions, we simplify the problem by making the following assumptions:

- We assume that α_{jit} is temporally constant, and is equal to the sum of an autoregressive term and a spatial term: $\alpha_{jit} = \alpha_i^{\text{AR}} + \alpha_i^{\text{spat}} \sum_{j \sim i} v_{ji}$, where $v_{ji} = \frac{1}{|\text{Ne}(j)|}$, with $j \sim i$ representing all the neighbours of i, and $|\text{Ne}(j)|$ being the number regions sharing a border (neighbors) with region j.
- We assume that $\alpha_{jit}^{\text{mob}}$ is temporally constant, and does not depend on the origin j, but only on the destination i: $\alpha_{ji,t-d}^{\text{mob}} = \alpha_{i,d}^{\text{mob}}$. Note that not making these assumptions pertaining to temporal constancy may lead to identifiability issues.

- For every i, t there are j independent Poisson processes (with mean λ_{jit}) competing to infect susceptibles in region i. Since the sum of Poisson processes is Poisson, we arrive at $\lambda_{it} = \sum_j \lambda_{jit}$.

The number of parameters to be estimated is now $O(I)$, which is much more computationally feasible. Adding an exogenous component, α_i^{EX}, for each region, leads us to an extension of the multi-region endemic-epidemic model

$$Y_{it}|\boldsymbol{Y}_{t-1} \sim \text{Pois}(\lambda_{it}^{\dagger})$$

$$\lambda_{it}^{\dagger} = \underbrace{\alpha_i^{\text{EX}}}_{\lambda_{it}^{\text{EX}}} + \underbrace{\alpha_i^{\text{AR}} \frac{Y_{i,t-1}}{N_i}}_{\lambda_{it}^{\text{AR}}} + \underbrace{\alpha_i^{\text{spat}} \sum_{j \sim i} v_{ji} \frac{Y_{j,t-1}}{N_j}}_{\lambda_{it}^{\text{spat}}} + \underbrace{\sum_{d=1}^{D} \alpha_{id}^{\text{mob}} \sum_j w_{ji,t-d} \frac{Y_{j,t-1}}{N_j}}_{\lambda_{it}^{\text{mob}}} \tag{3}$$

If the α_d^{mob}'s are found to be close to 0, then this model reduces to a typical endemic-epidemic model as seen frequently in the literature. Note that we also consider a negative binomial likelihood with mean $\lambda_{it}^{\dagger}$ and overdispersion parameters ψ_i.

3.3. *Delayed reporting, serial intervals, and incubation periods.* The modeling challenges caused by delayed reporting of cases is closely tied with the serial interval of infection and to the incubation period. The serial interval for Covid-19 has been estimated to be between 4 and 7 days, while the incubation period is between 4 and 9 days (Alene et al., 2021). These quantities can vary between individuals, and can be hard to measure due to delayed reporting/testing. García-García et al. (2021) showed that in Spain, cases may have peaked several days before the observed peak in cases, but the delay varied across Spanish provinces. Although we don't attempt to estimate any of these factors individually, we may be able to account for their combination by including additional time lags in our model. Bracher and Held (2020) showed that including cases from several time units in the past improved forecasting ability of endemic-epidemic models in the presence of random serial intervals. Following their guidance, we assume that the number of cases at time t is a weighted average of cases at s time points in the past. Our force of infection is now:

$$\begin{aligned} \lambda_{jit} &= \left(\alpha_i^{\text{AR}} + \alpha_i^{\text{spat}} \sum_{j \sim i} v_{ji} + \sum_{d=1}^{D} \alpha_{i,d}^{\text{mob}} w_{ji,t-d}\right) \sum_{s=1}^{S} \rho_s \frac{Y_{j,t-s}}{N_j} \\ &= \left(\alpha_i^{\text{AR}} + \alpha_i^{\text{spat}} \sum_{j \sim i} v_{ji}\right) \sum_{s=1}^{S} \rho_s \frac{Y_{j,t-s}}{N_j} + \sum_{d=1}^{D} \alpha_{i,d} w_{ji,t-d} \sum_{1 \le s < d} \rho_s \frac{Y_{j,t-s}}{N_j} \end{aligned} \tag{4}$$

where $\sum_{s=1}^{S} \rho_s = 1$. Note that in the second term, we exclude terms where the mobility lag is higher than the cases lag (e.g Y_{t-2}, w_{t-1}) as we don't suspect any reporting delay with our mobility data. If someone tests positive at $t-2$, when they move at time $t-1$, they should no longer be infectious, thus their mobility won't contribute to new cases.

It remains to specify D and S. In determining D, we first consider a univariate model for case counts: $Y_t|Y_{t-1} \sim \text{Poisson}(\lambda_t)$ with $\lambda_t = \phi_t Y_{t-1}$ where ϕ_t is the effective reproduction number at time t. If we solve for ϕ_t, and replace λ_t with Y_t, then we arrive at a crude estimate of R_t (Crude R_t) $\phi_t \approx \frac{Y_t}{Y_{t-1}}$. To determine how many mobility lags to include in our model, we visually examine the relationship of w_{t-h} with $\frac{Y_t}{Y_{t-1}}$ for various lags $h > 0$ using scatterplots. If w_{t-h} has a strong relationship with $\frac{Y_t}{Y_{t-1}}$, then we include this in our model.

Determining S is more challenging, but we can be fairly confident that $S \le 2$, as it is fairly unlikely that a case would take 3 or more weeks from primary case to cause a secondary

case, given that the serial interval and incubation period are likely less than 7 and 9 days respectively (Alene et al., 2021). Thus we will investigate values of $S = 1$ and $S = 2$.

3.4. *Underreporting.* Epidemic curves are well-known to suffer from underreporting and under-ascertainment (since our model can't distinguish between the two, we will simply use the term underreporting to encompass both). The number of cases is usually an underestimate of the number of infections, because of testing capacity limitations and asymptomatic or minimally symptomatic cases going undetected. This is troublesome when conducting inference, as our estimate of the α's will be affected by the reporting probability. There are several methods that account for underreporting, such as those in Wakefield, Dong and Minin (2019) or Bracher and Held (2021), but this is still an open problem in endemic-epidemic modelling, and we don't provide a perfect solution here.

A computationally simple method to account for underreporting is to adjust the intensity parameter for changes in testing via a log-linear model:

$$Y_{it}|\boldsymbol{Y}_{t-1} \sim \text{Pois}(\gamma_t \lambda_{it}^{\dagger}) \tag{5}$$

where $\gamma_t \in [0, 1], \log(\gamma_t) = \beta_{\text{test}} \log\left(\frac{x_t}{\max_t(x_t)}\right)$, where x_t is the number of tests performed at time t. This method assumes that the effect of testing is time constant. In Madrid, this assumption seemed reasonable and hence this method was employed. We present our estimated γ_t in Appendix E.

In Castilla-Leon, however, a method that accounts for time-varying reporting effects is required. Although Bracher and Held (2021) present methods for estimating endemic-epidemic model parameters in the presence of binomially thinned (underreported) case counts, it is unclear how these methods would generalize to the multivariate, mobility augmented case. We therefore present a novel approach (presented in Appendix E) which will be used as a sensitivity analysis in Castilla-Leon.

3.5. *Summary Statistics.* Although the parameters in our model are interpretable themselves, it is really their combinations that allow us to answer meaningful epidemiological questions about infectious diseases. Examples of such questions are: What was the effective reproduction number of Covid-19 in the first year of the pandemic? How many infected trips does it take to lead to a new infection? What proportion of cases are attributable to mobility? These are all examples of questions that can be answered by combining estimates of our model parameters.

The basic reproduction number, R_0, is a succinct way to describe the infectiousness of a disease, and is defined as the average number of secondary cases caused by an index primary case in a fully susceptible population (Diekmann, Heesterbeek and Britton, 2013). When dealing with more than one region (or some other strata), measuring R_0 is nontrivial. Rewriting (1) in matrix form, we obtain

$$\boldsymbol{\lambda}_t = \boldsymbol{\omega} + \boldsymbol{\alpha}\boldsymbol{Y}_{t-1}$$

with $\boldsymbol{\alpha}_{ij} = \alpha v_{ij}$. Diekmann, Heesterbeek and Metz (1990) use a limit argument to show that after a large number of generations, the typical number of primary cases given secondary cases is well described by the dominant eigenvalue of $\boldsymbol{\alpha}$ (assuming $\boldsymbol{\alpha}$ is irreducible and aperiodic). They thus define R_0 to be this dominant eigenvalue. We don't believe that this argument extends well to the case when $\boldsymbol{\alpha}$ is temporally changing, since $\boldsymbol{\alpha}$ only 'acts' on $\boldsymbol{Y}$ for ≈ 1 generation (this is assuming that the generation time is one time unit). However, in keeping with the endemic-epidemic modelling literature, we will present dominant eigenvalues over time where possible, as it is likely a good representation of infectiousness, but we

suspect it is biased and is more noisy than R_0 should be. Furthermore, since we are estimating infectiousness over time in a population with a changing proportion of susceptibles, we consider dominant eigenvalues to be estimates of the effective reproduction number, R_t.

Where possible, we compute the dominant eigenvalue of the matrix with entries

$$\alpha_i^{\text{AR}} I_{\{i=j\}} + \alpha_i^{\text{spat}} v_{ji} + \sum_{d=1}^{D} \left(\alpha_{i,d}^{\text{mob}} w_{ji,t-d}\right) \tag{6}$$

where I is the indicator function. We will plot this over time t.

Reproduction numbers measure the number of new cases stemming from old cases, but we also want to quantify the number of new cases stemming from the mobility of infectious people. We summarize the number of new infections per infected trip over time as

$$\frac{\sum_i (\sum_{d=1}^{D} \alpha_{i,d}^{\text{mob}} \sum_j w_{ji,t-d} \frac{Y_{j,t-1}}{N_j})}{\sum_i \sum_{d=1}^{D} \sum_j w_{ji,t-d} \frac{Y_{j,t-1}}{N_j}}$$

Although this may look cumbersone, it is simply a weighted average of the $\alpha_{i,d}^{\text{mob}}$'s over time. Furthermore, we can look at the number of infections per infected trip at the region level by summing over t instead of i. This formula can be easily modified in the presence of a serial interval.

3.6. *Inference.* All model parameters were estimated using Bayesian Markov chain Monte Carlo. In particular, we used the No-U-Turn sampler readily available in Stan (Carpenter et al., 2017) and its associated R package (Stan Development Team, 2021). Weakly informative/vague priors were used for inference (see Table B2 for complete table of priors). Four chains with 1000 interations, with the first half being warmup were used for each model. Trace plots were used to visually assess convergence of Markov chains. $\hat{R}$ values with a threshold of 1.01, as described in Vehtari et al. (2021), were also used to determine an appropriate level of mixing. Note that a handful of parameters in the model for underreporting presented in Appendix E did not meet this threshold, and is hence should be interpreted with caution. Since each of our summary statistics is a function of the model parameters, we can easily obtain credible intervals for each statistic by using draws from the joint posterior.

For model comparison, we approximate leave-one-out cross-validation using Pareto smoothed importance sampling (PSIS) implemented in the *loo* R-package (Vehtari et al., 2024).

4. Application. In this section, we apply our model to two Spanish Communities separately. In Section 4.1.1, we treat all of Castilla-Leon as a single region, which is mainly used as an exploratory analysis to inform our multi-region (spatial) model. In 4.1.2, we apply our multi-region model to the 245 subregions of Castilla-Leon and quantify the risk associated with travelling during the pandemic. We then apply our model to the 179 subregions of Madrid in Section 4.2.

4.1. *Assessing the risk associated with travelling in Castilla-Leon.*

4.1.1. *Castilla-Leon - aggregated model.* A plot of the case, test, and mobility data for all of Castilla-Leon is shown in Figure 2a. As noted by other authors, there is often a large time lag between a peak in mobility and the subsequent peak in cases, and this effect appears to change over time (Gottumukkala et al., 2021). However, mobility should only affect the relative change in the number of infections, as mobility can only affect cases through current infectious individuals coming into contact with susceptibles. To examine the relationship

between mobility and infectiousness, we compute the Crude R_t over time and look at the cross correlation between it and mobility. We found that mobility at time $t-2$ and $t-1$ show strong correlation with the Crude R_t at time t, followed by a sharp drop in correlation (0.27 to 0.08) when mobility is lagged by 3 or more time units. For this reason, we will consider the following mean as a starting point:

$$\lambda_t = \alpha^{\text{EX}} + \alpha^{\text{AR}} y_{t-1} + \alpha_1^{\text{mob}} w_{t-1} y_{t-1} + \alpha_2^{\text{mob}} w_{t-2} y_{t-1}$$

The fitted values of this model are shown in Figure A1. This plot suggests that mobility is explaining the majority of the case counts. However, this seems too large and warrants investigation. If we plot the Crude R_t versus the mobility (Figure A3), we can see that there are two extremely high leverage points with high mobility and Crude R_t. These two points correspond to the first weeks of March 2020, prior to lockdowns, and there were no mask mandates or policies enacted to slow the spread. As a result of these high leverage points, the effect of mobility (slope of the solid line in Figure A3) is too high. Although this plot is an oversimplification of exactly what our model is doing when estimating the effect of mobility, they warrant our attention. If we remove these points, the least squares line becomes much shallower and fits the Crude R_t estimates much better, as seen in Figure A3.

We now fit the model without the first three weeks of March, with the results shown in Figure A2, where the autoregressive component is much more substantial relative to the mobility component. For this reason, we will exclude these first three weeks of data when extending our model to multiple regions.

4.1.2. *Castilla-Leon – Multi-region Model.* The results of fitting the multi-region model is summarized in Figure 3. The movement component appears to be the strongest, followed by the autoregressive component and the spatial component. In some regions the movement component was very small, while it dominated the infections in others. Given that our results can be sensitive to one or two time points, we suspect that the region-level mobility effects are noisy. However, the aggregation of them is more likely to produce a clear signal.

When adjusting for testing using the method described by Equation 5, we found that the estimated γ_t was very close to 1. This may be due to a time-dependent effect of testing on the the number of reported cases. Another explanation is that that we may need region-level testing data to tease out the potentially spatially heterogenous effect. For this reasons, we did not control for testing in Castilla-Leon, acknowledge this as a limitation, and explore it further in Appendix E. This may cause underestimation of PCAtM, dominant eigenvalues, and infections per infected trip.

The proportion of cases attributable to movement (PCAtM) is presented in Table 1 for four different models

1. No serial interval, two mobility lags, and no testing adjustment
2. No serial interval, two mobility lags, and testing adjustment
3. Serial interval of 2 weeks, 3 mobility lags, and no testing adjustment
4. No serial interval, 3 mobility lags, and no testing adjustment

Adjusting for testing had little effect on the PCAtM. Similarly, using a serial interval of 2 weeks (as opposed to 1 week) had little effect on the PCAtM, but the additional mobility lag seems to be accounting for additional cases. However, the $\text{elpd}_{\text{loo-cv}}$ is lowest for model 1, and we focus our results on this model.

The proportion of cases attributable to movement (PCAtM) and the trips per infection for each region is shown in Figures 4a and 4b. Both the PCAtM and the trips per infection show a high amount of heterogeneity between regions. The temporal variation in trips per infection, averaged across Castilla-Leon are shown in Figure 5a. Based on this model, it takes roughly 70 infected trips to see a new infection.

The temporally changing dominant eigenvalues computed from (6) are shown in Figure 6. The dominant eigenvalue exceeding one seems to correspond with increases in case counts in Castilla-Leon, with the exception of the third viral wave. This may be due to properties of the virus at this time (such as a new variant), the drastic increase in testing that we had trouble accounting for, or some other confounding factors.

For completeness, we refit Model 1 using a negative binomial likelihood. We found that the PCAtM was larger than the one presented in Table 1. However, the uncertainty was larger in each model component, and the model fit was deemed inappropriate based on visual inspection and PSIS. Thus, the additional flexibility in the model did not lead to an improved model fit.

4.2. *Assessing the risk associated with travelling in the Community of Madrid.* For completeness, we present the results for Madrid with two major caveats: 1) a single large region (Madrid City) contains 49.6% of the Community of Madrid's population and 51% of the Covid-19 cases and 2) the intra-regional mobility in Madrid City (10.2% of the Community of Madrid's mobility) shows a highly different pattern (see Figure C1) than the rest of the mobility in the Community, with a peak during the first lockdown. Since the trend in case counts is roughly the same as the rest of the region, but the mobility is highly different, we do not believe that our model accurately captures the relationship between mobility and infectiousness in Madrid City.

Figure 2b displays time series of weekly trips, tests, and cases aggregated across the community of Madrid. After removing the first three weeks of data (as with Castilla-Leon) and correcting for changes in testing, we find that our assumption regarding the reproduction number being a linear function of mobility is reasonable (see Figure C2). Furthermore, we adjusted the per-contact-probability of infection for vaccinations (as described in Appendix D), but found no substantial difference in our results.

The proportion of cases attributable to movement (PCAtM) is presented in Table 1 for three different models:

1. No serial interval, three mobility lags, and testing adjustement
2. Serial interval, three mobility lags, and testing adjustment
3. Serial interval, excluding Madrid City, and testing adjustment

Noting that Model 3 is simply a subset of Model 2, the fit of Model 3 is shown in Figure 3. Mobility accounts for a substantial proportion of the cases, but the autoregressive term explains the most. In the model with the serial interval, ρ_1 was very close to zero, indicating that the model with the serial interval is more appropriate. However, this may have occurred due to the lag one mobility effect being very small, and our model is avoiding including that term.

The spatial distribution of the PCAtM is shown in Figure 4c. Note that the regions with a low PCAtM tend to be very close to Madrid City, while the regions with high PCAtM don't show a spatial pattern. This may seem unintuitive, as we may expect a higher PCAtM around Madrid City due to gravity. However, it is likely that people stopped commuting in and out of Madrid City, causing within-Madrid City trips to replace incoming/outgoing trips. Lower amounts of mobility in the regions surrounding Madrid City will ultimately lead to lower PCAtMs in those regions.

The number of infections per infected trip is shown both spatially and temporally in Figures 4d and 5b. Figure 4d suggests that the trips required for a new infection are spatially correlated, indicated by the clusters of regions of the same colour. Figure 5b suggests that, excluding Madrid City, roughly 140 infected trips are required for a new infection to arise.

5. Discussion. In this paper we developed an infectious disease model where the number of contacts between people is a linear function of trips between regions. We showed that this model is an extension of endemic-epidemic models frequently found in the literature. We applied this model to two Spanish Communities with the intention of quantifying the risk associated with travelling in each Community. In Castilla-Leon, we found that we could relate just under half of the trips to our cellphone mobility data, while this was much lower in Madrid. The potential reason for the disparity in results is that Madrid is much more population dense, and has a city-centre with over half the Covid-19 cases of the entire province. This causes potential underestimation of the PCAtM and trips per infection, as our Madrid model is likely underestimating the importance of mobility. Our model appears to work better when regions are small, as our cellphone mobility data is more informative.

We found that this class of models is sensitive to large changes in case counts that occurred early in the pandemic, as well as rapid changes in testing capacity. Although we took great care in specifying each model component, developing robust methods for modeling the infectiousness of the disease when using this class of models should be researched further. We stress the importance that exploratory and diagnostic plots can greatly improve inference and interpretation when using endemic-epidemic, or any infectious disease model.

One strength of this work is that we utilize rich mobility data and spatial data to model disease spread through a carefully parametrized infectious disease model. In doing so we were able to assign a number to the risk associated with travelling during a pandemic. This framework can be easily extended to include other spatial/temporal covariates such as mask usage, if such data are available. We explored adjusting the α's for the stringency index, which is an aggregate measure of how strict policies are in specific regions Hale et al. (2021). However, this index was extremely highly correlated with mobility, and thus was not appropriate to include in our model, as they are largely measuring the same thing.

A further strength of this work is that it was done during a time period prior to mass vaccinations and the introduction of the major Covid-19 variants, which could have confounded our analysis. This could also be viewed as a limitation, as we could have allowed the α's to change when the major variants (e.g Omnicron) arose, and could easily account for higher vaccination rates using the methodology from this paper. In our analysis of the Community of Madrid, too few people had been vaccinated for it to make any major difference in our results. Ideally, we would have mobility data over the course of the entire pandemic, so that we could see how the risk associated with travelling changes with new variants and increasing levels of immunity in the population.

A limitation of our work that we must emphasize is that we cannot associate individual trips to individual infections, and thus cannot infer causality. Although we are confident that mobility is required for Covid-19 to spread, we cannot be sure that the trips recorded in our data are causing cases according to our model specification, as there may be confounding factors associated with between-region mobility and case counts.

This work opens the door for many avenues of future research. Firstly, robust methods for modeling infectiousness as a function of mobility (or any covariate) would be extremely useful. For instance, a method utilizing quantiles would be insensitive to rapid changes in the observed cases. Furthermore, although this paper presents a novel method for modelling reproduction numbers based on mobility, we need to further theoretically examine how we define reproduction numbers (i.e the dominant eigenvalue of the next-generation matrix) from this class of models. Furthermore, we need to rethink how to estimate temporally changing reproduction numbers from this class of models, especially as the model becomes more complex.

Although this study has focussed on Covid-19, we want to emphasize that the model and associated principles can be extended to a wide variety of infectious diseases, and various forms of network data. Extensions and simplifications should be made on a case-by-case basis, and should be guided by careful data exploration.

6. Data Availability. Code and data are available at https://github.com/JustinJSlater/mobilityCovid.

7. Acknowledgements. JJS was supported by the Natural Sciences and Engineering Research Council of Canada (PGSD3-559264-2021). JSR is supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant (RGPIN-2019-04142). PEB is supported by the Natural Sciences and Engineering Research Council of Canada (RGPIN-2022-05164).

REFERENCES

ALENE, M., YISMAW, L., ASSEMIE, M. A., KETEMA, D. B., GIETANEH, W. and BIRHAN, T. Y. (2021). Serial interval and incubation period of COVID-19: a systematic review and meta-analysis. *BMC Infectious Diseases* **21** 1–9.

BAUER, C. and WAKEFIELD, J. (2018). Stratified space–time infectious disease modelling, with an application to hand, foot and mouth disease in China. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **67** 1379–1398.

BRACHER, J. and HELD, L. (2020). Endemic-epidemic models with discrete-time serial interval distributions for infectious disease prediction. *International Journal of Forecasting* **38** 1221–1233.

BRACHER, J. and HELD, L. (2021). A marginal moment matching approach for fitting endemic-epidemic models to underreported disease surveillance counts. *Biometrics* **77** 1202–1214.

CARPENTER, B., GELMAN, A., HOFFMAN, M. D., LEE, D., GOODRICH, B., BETANCOURT, M., BRUBAKER, M., GUO, J., LI, P. and RIDDELL, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software* **76** 1–32.

CELANI, A. and GIUDICI, P. (2022). Endemic–epidemic models to understand COVID-19 spatio-temporal evolution. *Spatial Statistics* **49** 100528.

DIEKMANN, O., HEESTERBEEK, J. A. P. and METZ, J. A. (1990). On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology* **28** 365–382.

DIEKMANN, O., HEESTERBEEK, H. and BRITTON, T. (2013). *Mathematical tools for understanding infectious disease dynamics* **7**. Princeton University Press.

DOUWES-SCHULTZ, D., SUN, S., SCHMIDT, A. M. and MOODIE, E. E. (2022). Extended Bayesian endemic–epidemic models to incorporate mobility data into COVID-19 forecasting. *Canadian Journal of Statistics* **50** 713–733.

FRITZ, C. and KAUERMANN, G. (2022). On the interplay of regional mobility, social connectedness and the spread of COVID-19 in Germany. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **185** 400.

GARCÍA-GARCÍA, D., VIGO, M. I., FONFRÍA, E. S., HERRADOR, Z., NAVARRO, M. and BORDEHORE, C. (2021). Retrospective methodology to estimate daily infections from deaths (REMEDID) in COVID-19: the Spain case study. *Scientific Reports* **11** 1–15.

GEILHUFE, M., HELD, L., SKRØVSETH, S. O., SIMONSEN, G. S. and GODTLIEBSEN, F. (2014). Power law approximations of movement network data for modeling infectious disease spread. *Biometrical Journal* **56** 363–382.

GENERAL DIRECTORATE OF INFORMATION SYSTEMS, QUALITY AND PHARMACEUTICAL PROVISION (2022). Open Data of Castile and Leon. https://datosabiertos.jcyl.es/web/es/datos-abiertos-castilla-leon.html. Accessed 2022-03-10.

GNEITING, T. and RAFTERY, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association* **102** 359–378.

GORDIS, L. (2013). *Epidemiology*, 5th ed. Elsevier Health Sciences.

GOTTUMUKKALA, R., KATRAGADDA, S., BHUPATIRAJU, R. T., KAMAL, A. M., RAGHAVAN, V., CHU, H., KOLLURU, R. and ASHKAR, Z. (2021). Exploring the relationship between mobility and COVID-19 infection rates for the second peak in the United States using phase-wise association. *BMC Public Health* **21** 1–14.

GRIMÉE, M., BEKKER-NIELSEN DUNBAR, M., HOFMANN, F., HELD, L. et al. (2022). Modelling the effect of a border closure between Switzerland and Italy on the spatiotemporal spread of COVID-19 in Switzerland. *Spatial Statistics* 100552.

HALE, T., ANGRIST, N., GOLDSZMIDT, R., KIRA, B., PETHERICK, A., PHILLIPS, T., WEBSTER, S., CAMERON-BLAKE, E., HALLAS, L., MAJUMDAR, S. et al. (2021). A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behaviour* **5** 529–538.

HALLORAN, M. E., LONGINI, I. M. and STRUCHINER, C. J. (2010). R_0 and Deterministic Models. In *Design and Analysis of Vaccine Studies* 85–102. Springer.

HE, X., LAU, E. H., WU, P., DENG, X., WANG, J., HAO, X., LAU, Y. C., WONG, J. Y., GUAN, Y., TAN, X. et al. (2020). Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine* **26** 672–675.

HELD, L., HÖHLE, M. and HOFMANN, M. (2005). A statistical framework for the analysis of multivariate infectious disease surveillance counts. *Statistical Modelling* **5** 187–199.

HELD, L. and PAUL, M. (2012). Modeling seasonality in space-time infectious disease surveillance data. *Biometrical Journal* **54** 824–843.

HELD, L., HOFMANN, M., HÖHLE, M. and SCHMID, V. (2006). A two-component model for counts of infectious diseases. *Biostatistics* **7** 422–437.

HERZOG, S., PAUL, M. and HELD, L. (2011). Heterogeneity in vaccination coverage explains the size and occurrence of measles epidemics in German surveillance data. *Epidemiology & Infection* **139** 505–515.

KNORR-HELD, L. and RICHARDSON, S. (2003). A hierarchical model for space–time surveillance data on meningococcal disease incidence. *Journal of the Royal Statistical Society Series C: Applied Statistics* **52** 169–183.

MEYER, S. and HELD, L. (2014). Power-law models for infectious disease spread. *Annals of Applied Statistics* **8** 1612–1639.

MEYER, S. and HELD, L. (2017). Incorporating social contact data in spatio-temporal models for infectious disease spread. *Biostatistics* **18** 338–351.

MINISTERIO DE SANIDAD, GOBIERNO DE ESPAÑA (2022a). Estrategia de Vacunacion COVID-19 EN España. https://www.sanidad.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov/vacunaCovid19.htm. Accessed 2022-05-17.

MINISTERIO DE SANIDAD, GOBIERNO DE ESPAÑA (2022b). Datos abiertos de pruebas realizadas. https://www.sanidad.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov/pruebasRealizadas.htm. Accessed 2022-05-17.

MINISTERIO DE TRANSPORTES MOVILIDAD Y AGENDS URBANA, GOBIERNO DE ESPAÑA (2022). Evolución de la movilidad diara. www.mitma.gob.es/ministerio/covid-19/evolucion-movilidad-big-data. Accessed 2022-10-26.

EPIDEMIOLOGICAL SURVEILLANCE NETWORK OF MADRID (2022). COVID 19-TIA by Municipalities and Districts of Madrid. https://datos.gob.es/en/catalogo/a13002908-covid-19-tia-por-municipios-y-distritos-de-madrid1. Accessed 2022-05-17.

PAUL, M., HELD, L. and TOSCHKE, A. M. (2008). Multivariate modelling of infectious disease surveillance data. *Statistics in Medicine* **27** 6250–6267.

PAUL, M. and HELD, L. (2011). Predictive assessment of a non-linear random effects model for multivariate time series of infectious disease counts. *Statistics in Medicine* **30** 1118–1136.

PONCE-DE LEON, M., DEL VALLE, J., FERNANDEZ, J. M., BERNARDO, M., CIRILLO, D., SANCHEZ-VALLE, J., SMITH, M., CAPELLA-GUTIERREZ, S., GULLÓN, T. and VALENCIA, A. (2021). COVID-19 Flow-Maps an open geographic information system on COVID-19 and human mobility for Spain. *Scientific Data* **8** 1–16.

SCHRÖDLE, B., HELD, L. and RUE, H. (2012). Assessing the impact of a movement network on the spatiotemporal spread of infectious diseases. *Biometrics* **68** 736–744.

SLATER, J. J., BROWN, P. E. and ROSENTHAL, J. S. (2021). Forecasting subnational COVID-19 mortality using a day-of-the-week adjusted Bayesian hierarchical model. *Stat* **10** e328.

SLATER, J. J., BROWN, P. E., ROSENTHAL, J. S. and MATEU, J. (2021). Capturing spatial dependence of COVID-19 case counts with cellphone mobility data. *Spatial Statistics* 100540.

STAN DEVELOPMENT TEAM (2021). RStan: the R interface to Stan. R package version 2.21.3.

VEHTARI, A., GELMAN, A., SIMPSON, D., CARPENTER, B. and BÜRKNER, P.-C. (2021). Rank-normalization, folding, and localization: An improved R-hat for assessing convergence of MCMC (with Discussion). *Bayesian Analysis* **16** 667–718.

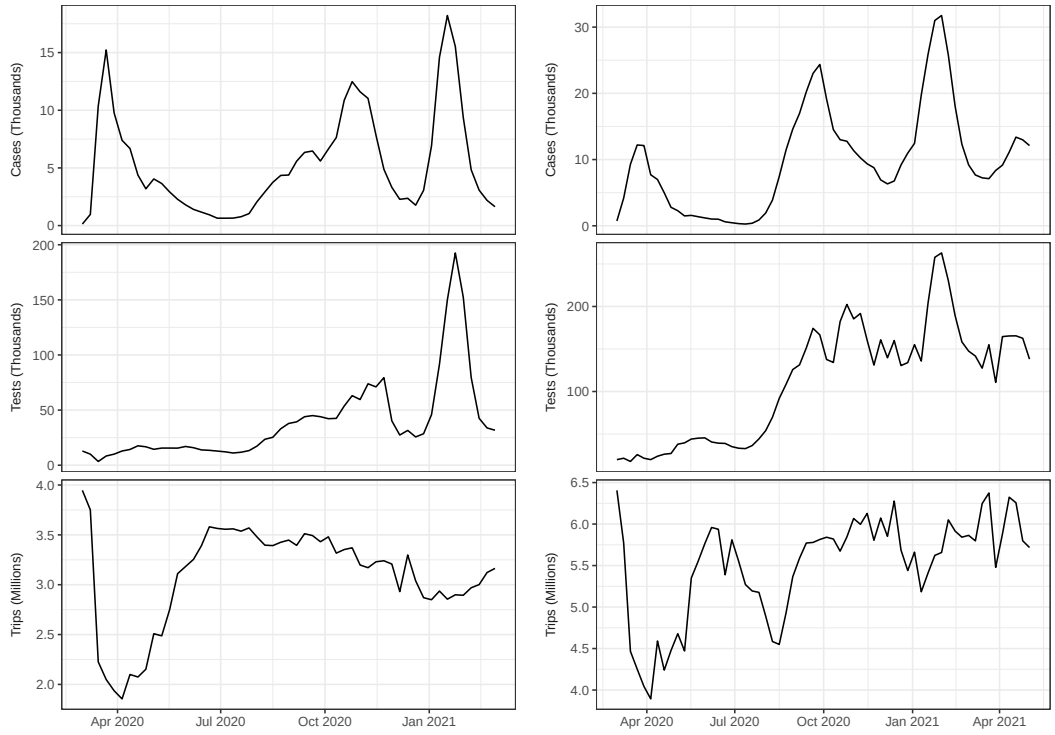
VEHTARI, A., GABRY, J., MAGNUSSON, M., YAO, Y., BÜRKNER, P.-C., PAANANEN, T. and GELMAN, A. (2024). loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models. R package version 2.7.0.

WAKEFIELD, J., DONG, T. Q. and MININ, V. N. (2019). Spatio-temporal analysis of surveillance data. In *Handbook of Infectious Disease Data Analysis* 455–475. Chapman and Hall/CRC.

		PCAtM (95% CrI)	ρ_1	$\widehat{\text{elpd}}_{\text{loo-cv}}$
Castilla	1. No SI, No testing	44.96 (43.75, 46.31)	-	134352.3
	2. No SI, testing	43.79 (42.55, 44.88)	-	134459.0
	3. SI, no testing, additional lag	56.99 (55.95, 58.03)	0.99996	135044.2
	4. No SI, no testing, additional lag	57.01 (55.94, 57.98)	-	134996.7
Madrid	1. No SI, with testing	17.00 (16.11, 18.03)	-	86181.6
	2. SI, with testing	14.00 (13.12, 15.02)	0.00001	85857.4
	3. SI, with testing (no Madrid City)	28.54 (26.76, 30.68)		

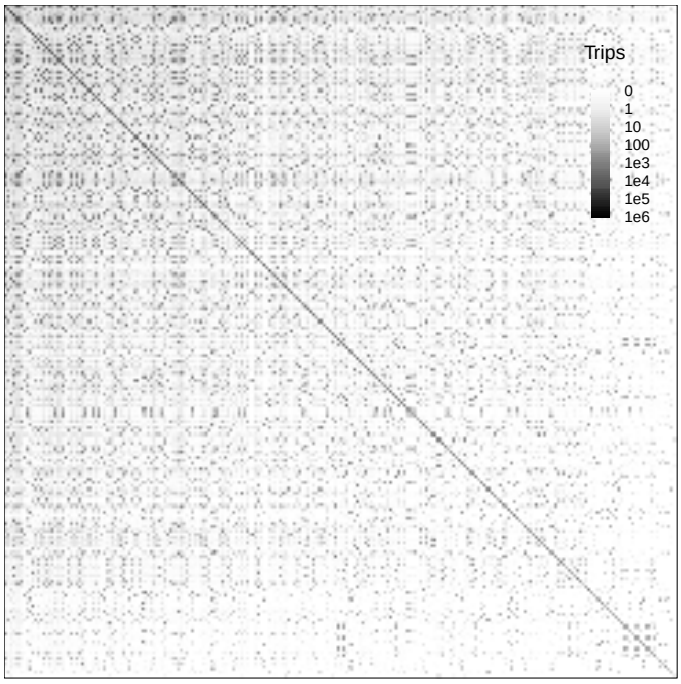
TABLE 1

Percentage of cases attributable to movement (PCAtM) for various models fit to Castilla-Leon and Madrid data. In models with a serial interval (SI), ρ_1 is presented. Posterior medians and 95% credible intervals (CrI) are presented for PCAtM, while posterior medians are presented for ρ_1. Expected log predictive density, estimated using leave-one-out cross-validation is presented. Note that roughly 3% of Pareto k were above 0.7 for the Castilla models, while roughly 1.5% were above 0.7 for the Madrid models, indicating potential reliability issues with PSIS-LOOCV.

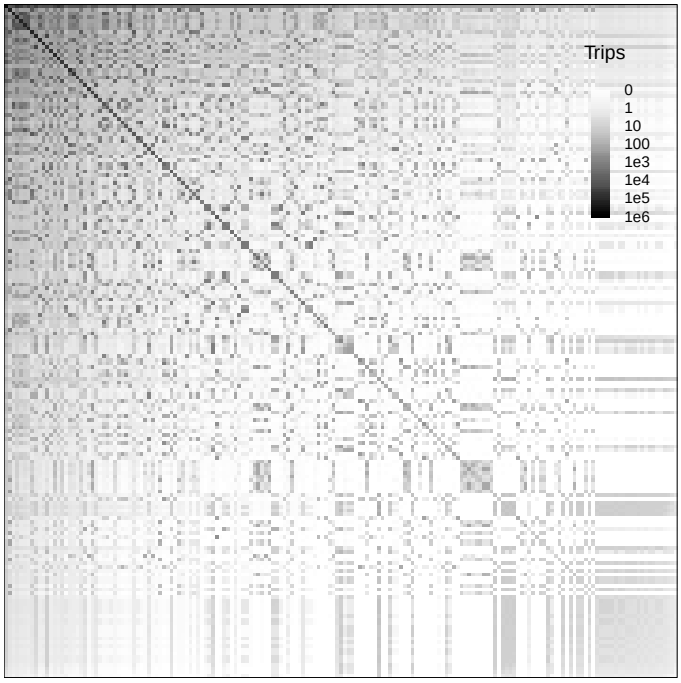


(a) Castilla-Leon (b) Madrid

Fig 1: Time series of cases, trips, and tests between March 2020 and March 2021 (Castilla-Leon), and March 2020 and May 2021 (Madrid).

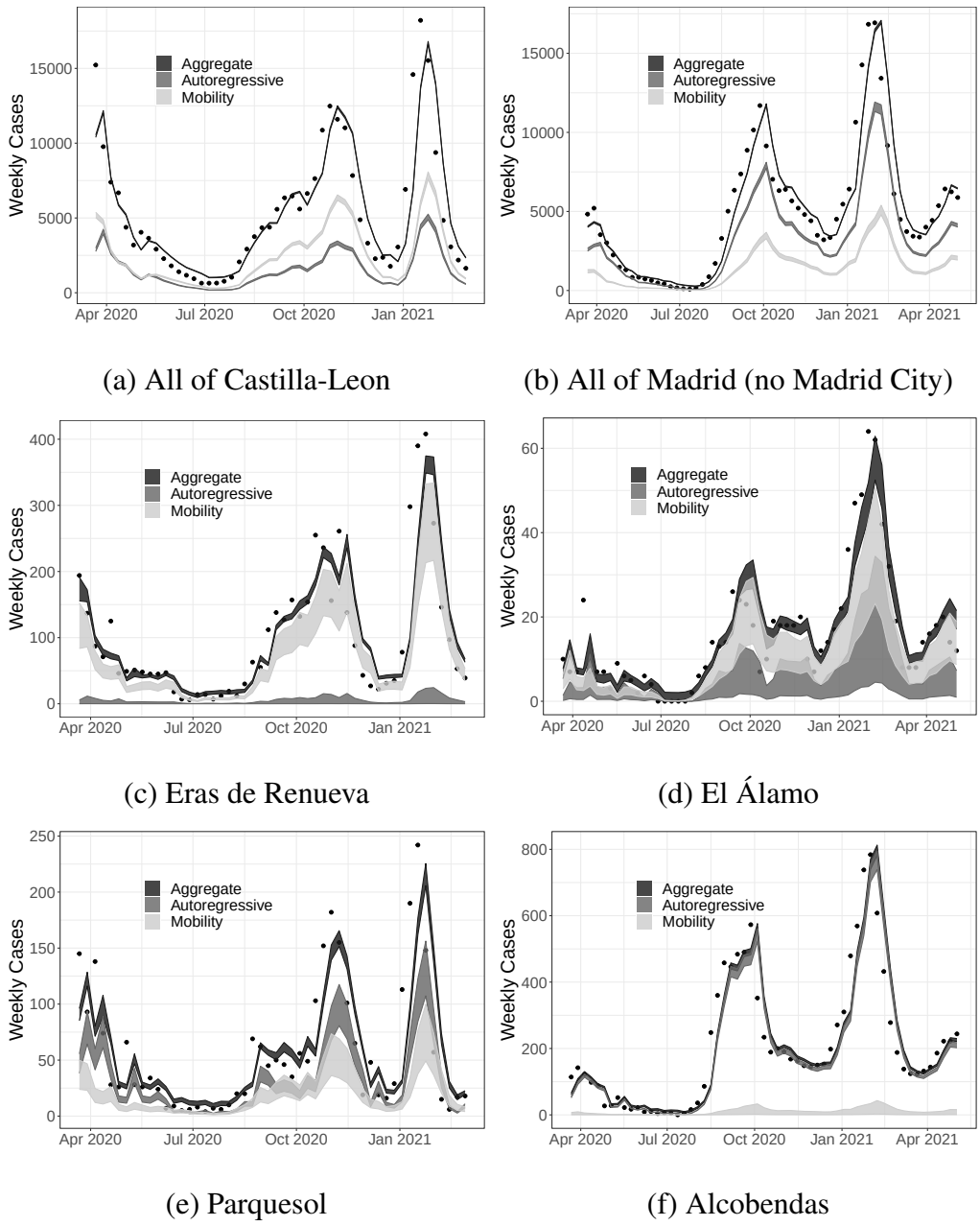


(a) Castilla-Leon



(b) Madrid

Fig 2: Mean daily trips between regions, arranged by total rowwise mobility for improved visual clarity. The darkness of the pixel indicates higher amounts of mobility. The diagonal line indicates that there is generally more mobility within regions than there is between other pairs of regions.



(a) All of Castilla-Leon

(b) All of Madrid (no Madrid City)

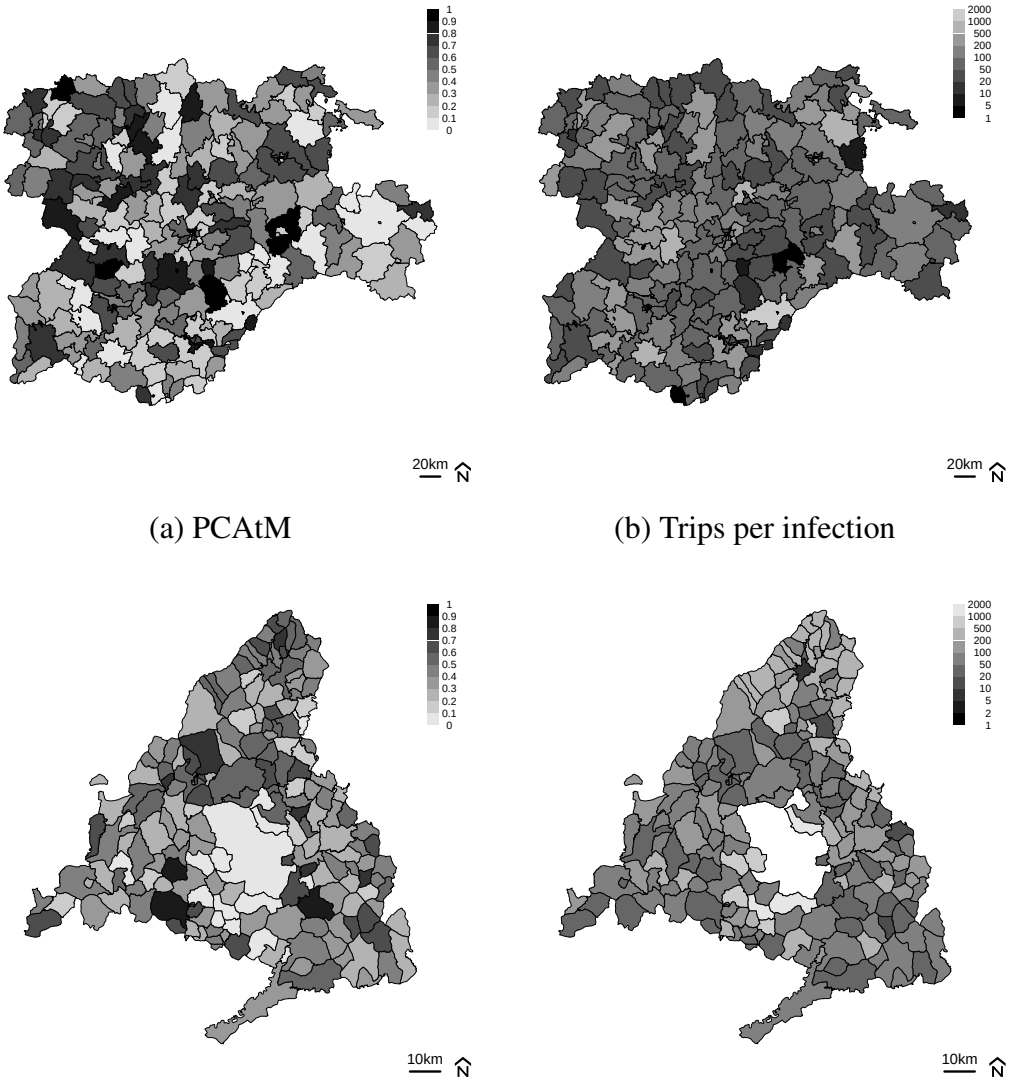
(c) Eras de Renueva

(d) El Álamo

(e) Parquesol

(f) Alcobendas

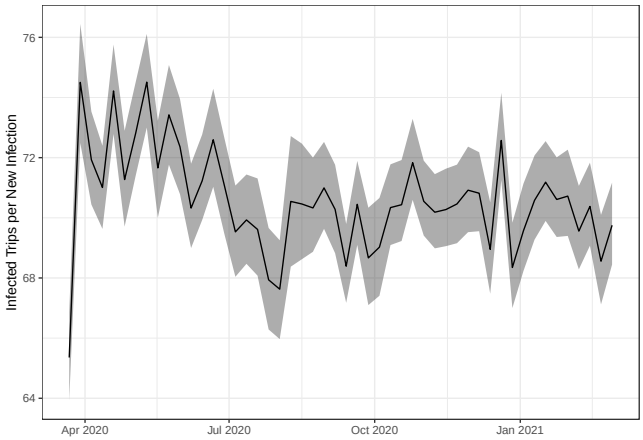
Fig 3: Results of our multi-region mobility extended endemic-epidemic model. For both Castilla-Leon (left) and Madrid (right), we present the results for the entire region, alongside a region that showed a strong mobility effect (Eras de Renueva and El Álamo), and a region showing a weaker mobility effect (Parquesol and Alcobendas). The 95% credible interval for each model component is presented, alongside their aggregation ($\lambda_t^{\dagger}$). Observed case counts are shown as black points. The locations of the regions depicted in c)-f) are shown in Appendix F. Note that the endemic and spatial components were neglible, and thus were omitted for improved visual clarity.



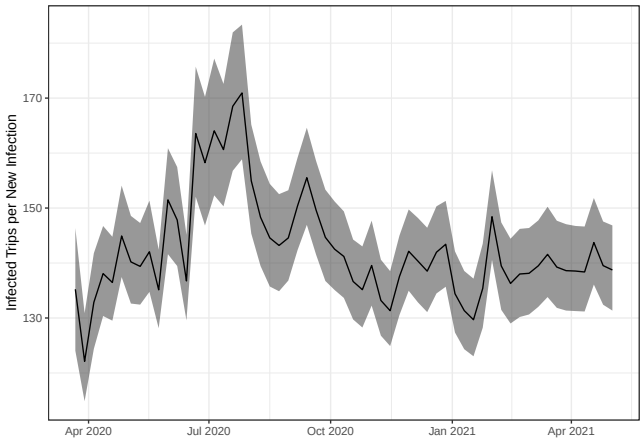
(a) PCAtM (b) Trips per infection

(c) PCAtM (d) Trips per infection

Fig 4: Spatial distribution of proportion of cases attributable to movement (PCAtM) and the number of trips associated with one new infection. The trips per infection in Madrid City (white region in d) was calculated to be 3753.



(a) Castilla-Leon



(b) Madrid

Fig 5: Temporal variation of number of trips associated with one new infection. Madrid City was excluded from this analysis, as the data quality issues caused this number to be implausibly high. The posterior median, alongside 95% credible intervals are presented.

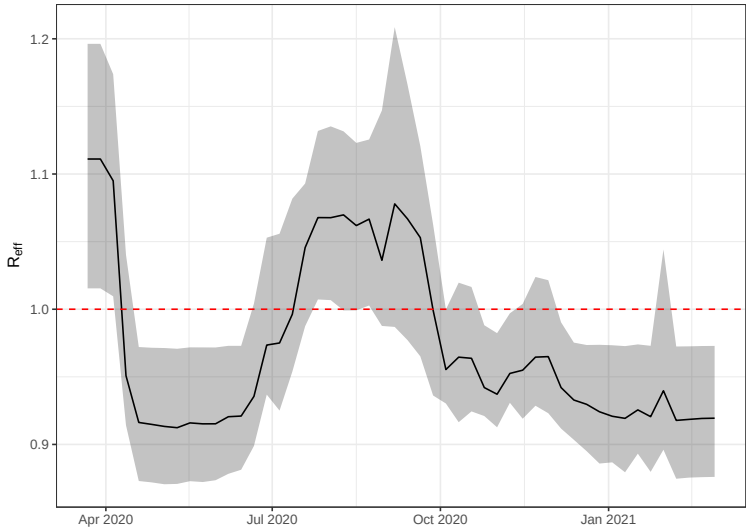


Fig 6: Posterior median and 95% credible interval of R_t in Castilla-Leon. An $R_t > 1$ will generally lead to an increase in cases.

Appendix A Treating Castilla-Leon as a single region. These figures were part of our exploratory analysis that guided model selection.

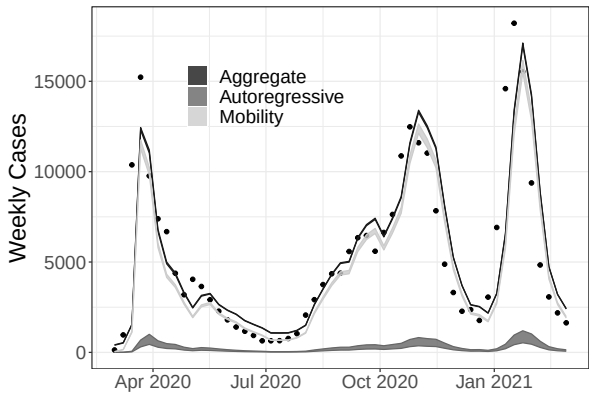


Fig A1: Single region, mobility-extended endemic-epidemic model fit to aggregate Castilla-Leon data, separated into components.

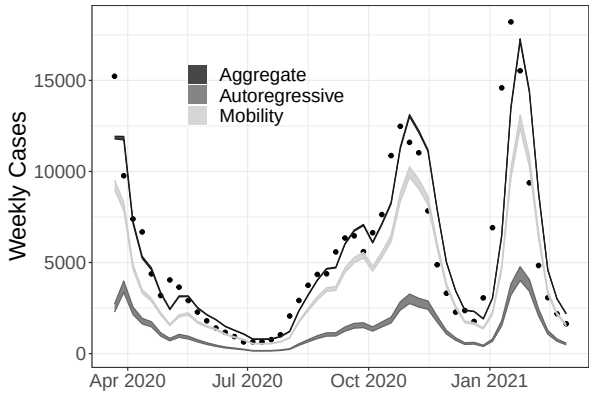


Fig A2: Single region, mobility-extended endemic-epidemic model fit to aggregate Castilla-Leon data with the first three weeks of data removed.

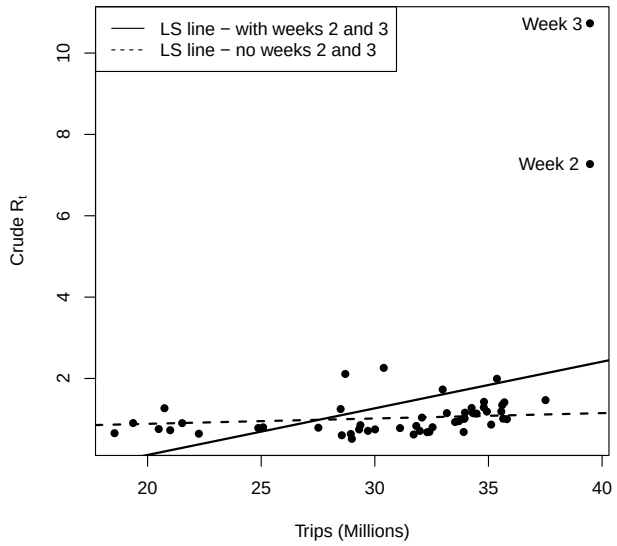


Fig A3: Crude R_t vs number of trips. There are two high leverage points which correspond to the first three weeks of the pandemic. These have a strong influence on the effect of mobility and cause the solid line to be much steeper than it should be. The dotted line is the least squares (LS) line with the two influential points removed, and visually fits the data much better.

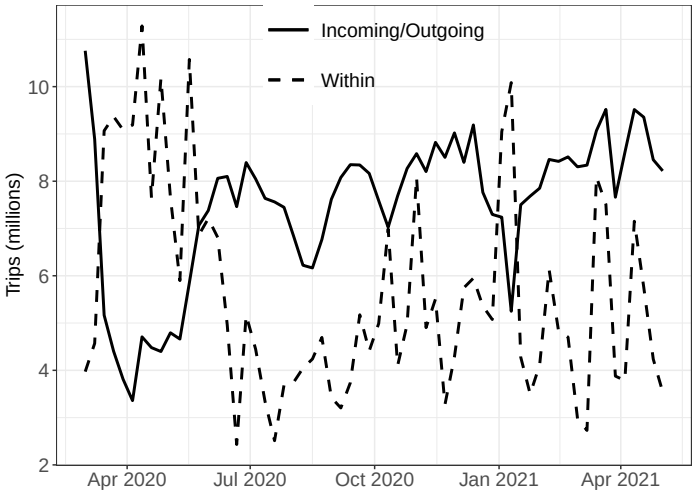
Appendix B Priors.

Parameter	description	Prior
$\alpha^{\text{mob}}_{1,2,3}$	Mobility effects	$N(0,.1)$
α^{AR}	Autoregressive effect	$N(0,2)$
α^{spat}	Spatial effect	$N(0,500)$
β_{test}	Test effect	$N(0,1)$
$\boldsymbol{\rho}$	Serial interval distribution	Dirichlet(1,1)

TABLE B2

Prior distributions used for models used in the main manuscript. Not all parameters are present in all models. Also note that these parameters vary for each region, with the subscript i removed for clarity.

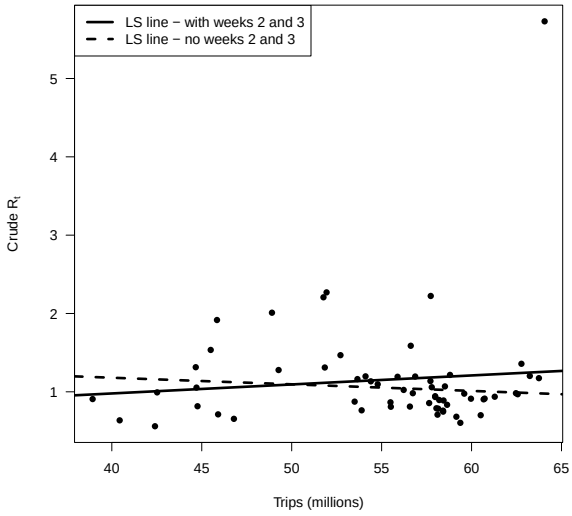
Appendix C Madrid: Supplementary plots.



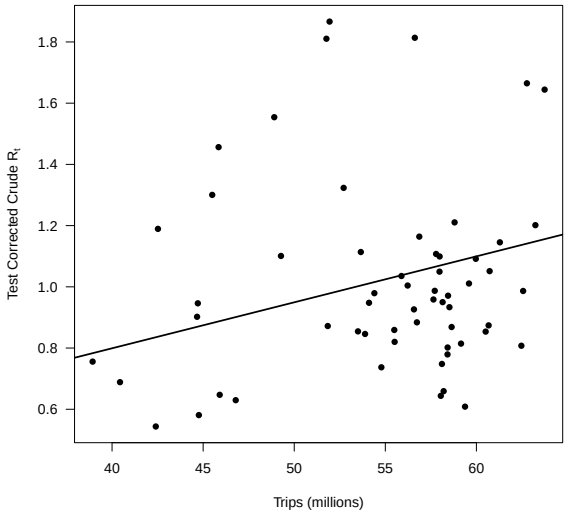
(a) Comparing the number of trips within Madrid City with the incoming/outgoing mobility of Madrid City

(b) Madrid - all mobility not involving Madrid City

Fig C1



(a) Community of Madrid - Trend lines for Crude R_t vs. lag-one mobility including (solid-line) and excluding (dotted line) the first 3 weeks of data



(b) Community of Madrid - Trend lines for test corrected Crude R_t vs. lag-one mobility, excluding the first three weeks of data.

Fig C2

Appendix D Accounting for Vaccinations in the Community of Madrid. In Spain, vaccines started to reach the general public in January of 2021, with about 30% of the public having at least one dose of the vaccine by May 2021. Thus, this may impact our Madrid results and should be explored. Vaccine data has been incorporated into endemic-epidemic models in previous work (Herzog, Paul and Held, 2011), and we follow suit here. The immunity induced by vaccines should reduce the per-contact probability of infection. In other words, the per-contact probability of infection will change over time as the proportion of people vaccinated increases. Using the univariate model with $D = 1$ as an example, our force of infection is now

$$\lambda_t^{\text{vacc}} = \mathcal{C}_{t-1} \times p(u_{t-1}) \times \frac{Y_{t-1}}{N}$$

where

$$g[p(u_{t-1})] = p_0 - \tau u_{t-1}$$

where p_0 is the per-contact probability of infection in an unvaccinated population, τ is a reduction in infection probability due to vaccination, g is a link function, and u_t is the proportion of the population that is vaccinated at time t. The force of infection becomes

$$\lambda_t^{\text{vacc}} = (c_0 + c^{\text{mob}} w_{t-1}) \cdot g^{-1}(p_0 - \tau u_{t-1}) \frac{Y_{t-1}}{N}.$$

The identity link would lead to

$$\lambda_t^{\text{vacc}} = \left(\alpha^{\text{AR}} + \alpha^{\text{mob}} w_{t-1} - (c_0 + c^{\text{mob}} w_{t-1}) \tau u_{t-1}\right) \frac{Y_{t-1}}{N}$$

which allows for potentially negative values of λ_t without some numerically unstable constraints. Furthermore, this would assume a linear relationship between proportion vaccinated and infection probability, which seems unrealistic. Instead, we used a log link leading to

$$\lambda_t^{\text{vacc}} = \left(\alpha^{\text{AR}} + \alpha^{\text{mob}} w_{t-1}) e^{-\tau u_{t-1}}\right) \frac{Y_{t-1}}{N}.$$

Appendix E Accounting for underreporting via a hierarchical model.

In Madrid, we accounted for underreporting my correcting the intensity by a factor γ_t that is a function of testing. The plot of γ_t is shown in Figure E1.

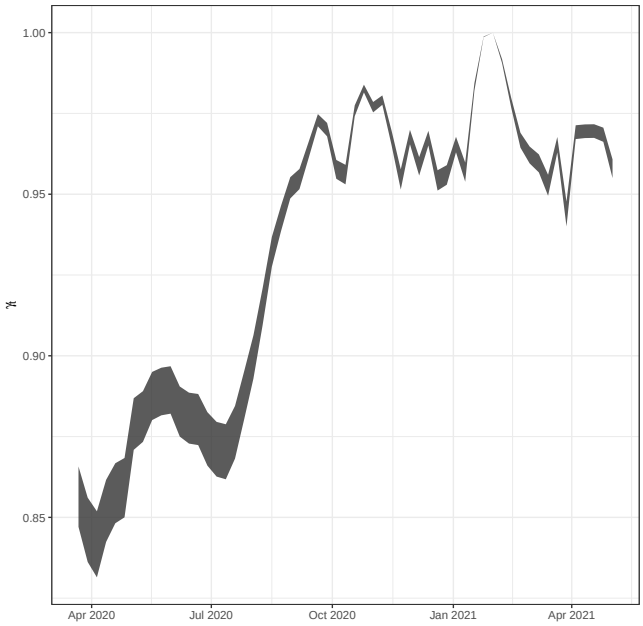


Fig E1

For Castilla-Leon, we require a more complex method, as the relationship between the infections and testing was more complex. Let the true case counts in region i at time t be Z_{it}. Then we assume:

$$Y_{it}|Z_{it} \sim \text{Bin}(Z_{it}, \pi_t)$$
$$Z_{it}|\boldsymbol{Z}_{t-1} \sim \text{Pois}(\lambda^Z_{i,t})$$

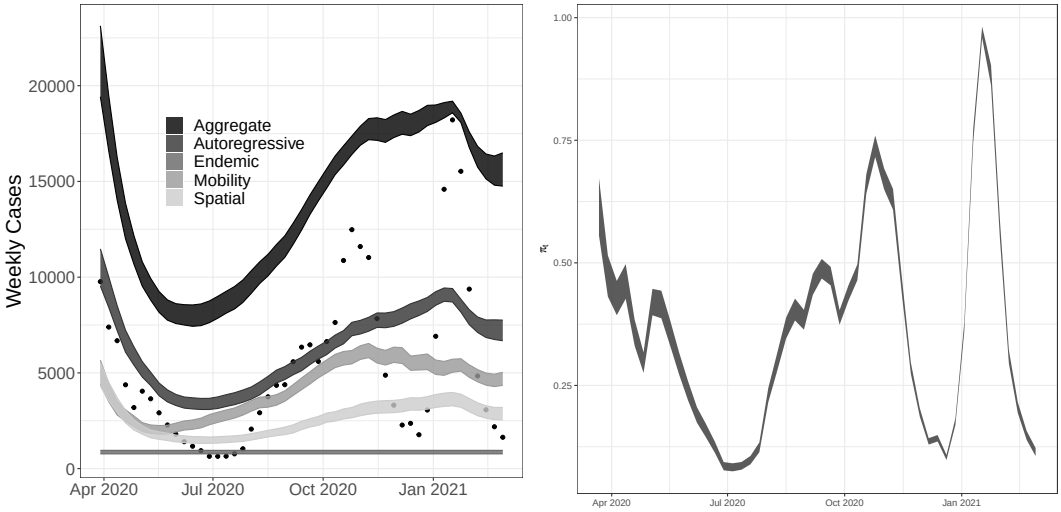
where λ^Z_{it} is the same as $\lambda^\dagger_{it}$ in Equation 3 but with Z's instead of Y's, and π_t is the reporting probabiliity at time t. π_t is assumed to be the same across regions for computational and identifiability purposes. Conducting inference on this model is challenging, as it involves $I \times T$ latent discrete variables. The unbounded, discrete nature of these parameters makes modern Bayesian computation with Hamiltonian Monte Carlo (HMC) or Integrated Nested Laplace Approximations (INLA) infeasible. We thus use normal approximations in both layers of the hierarchical model:

$$Y_{it}|Z_{it} \dot{\sim} N(Z_{it}\gamma_t, Z_{it}\pi_t(1-\pi_t))$$
$$Z_{it}|\boldsymbol{Z}_{t-1} \dot{\sim} N(\lambda^Z_{it}, \lambda^Z_{it})$$

where $\dot{\sim}$ means 'approximately distributed as'. The larger the counts in each region, the better this approximation will be.

We fit this model to the Castilla-Leon data with no serial interval and two mobility lags. We chose this simpler model for computational purposes. We also note that our $\hat{R}$ values were relatively high for the first 10 time points (~ 1.15) or so, so these results should be interpretted with caution.

Figure E2a shows the 95% credible interval of what we estimate to be the true epidemic curve based on the hierarchical model above. We can see that our model is consistent with the idea that the third wave was an artifact of a surge in testing. Furthermore, the autoregressive portion of the model seems to be explaining a majority of the cases. The 95% credible interval of the reporting probability is shown in Figure E2b. Although this model did not use testing data, the estimated reporting probability still closely follows the pattern in the testing data, with the exception of the first three months, when testing was not readily available to the public.



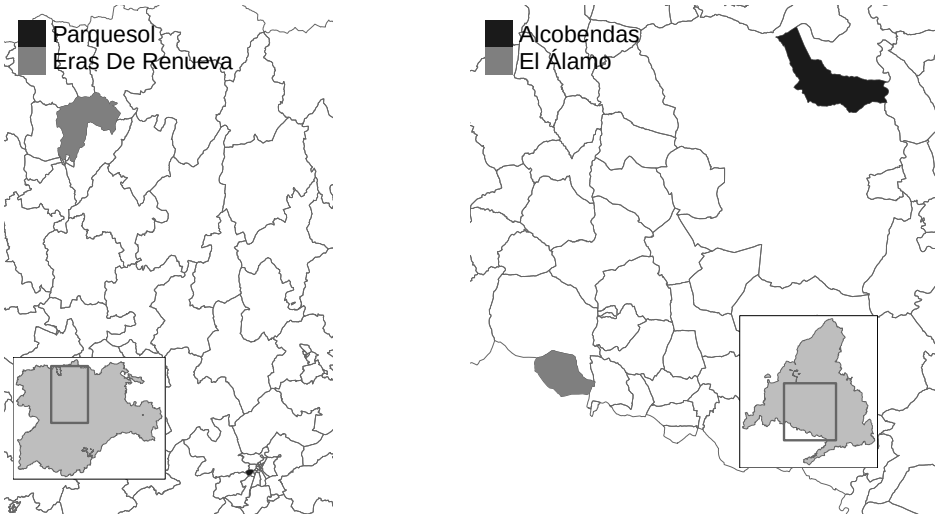
(a) 95% credible intervals for each model component. Points are observed cases.

(b) 95% credible intervals for π_t estimated from hierarchical model

Fig E2: Model results for Castilla-Leon with 1 week serial interval, two mobility lags, and a hierarchical component to account for underreporting.

Although the mobility component does not appear dominant in this case, it is still larger than it was in the model where we simply corrected for changes in testing. This indicates that underreporting is likely causing us to underestimate the true number of cases that are associated with mobility.

Appendix F Locations of regions depicted in Figure 2.



(a) Castilla-Leon (b) Madrid

Fig F1: Locations of regions depicted in Figure 3