Epidemiological modeling of the 2005 French riots: a spreading wave and the role of contagion

Laurent Bonnasse-Gahot[1], Henri Berestycki[1] , Marie-Aude Depuiset[2,3,4], Mirta B. Gordon[4], Sebastian Roché[2], Nancy Rodriguez[5], Jean-Pierre Nadal[1,6,*]

[1]École des Hautes Études en Sciences Sociales, PSL Research University, CNRS, Centre d'Analyse et de Mathématique Sociales, Paris, France
[2]Univ. Grenoble Alpes, Institut d'Etudes Politiques de Grenoble, CNRS, PACTE, Grenoble, France
[3]Université de Lille, CNRS, UMR 8019 - CLERSÉ - Centre Lillois d'Études et de Recherches sociologiques et Économiques, Lille, France
[4]Univ. Grenoble Alpes, CNRS, LIG, Grenoble France
[5]The University of North Carolina at Chapel Hill, Department of Mathematics, Chapel Hill, USA
[6]Laboratoire de Physique Statistique, École Normale Supérieure, PSL Research University; Université Paris Diderot, Sorbonne Paris-Cité; Sorbonne Universités, UPMC – Univ. Paris 06; CNRS; Paris, France.
*corresponding author. Email: jpnadal@ehess.fr

(Dated: January 25th, 2017)

Abstract

As a large-scale instance of dramatic collective behavior, the 2005 French riots started in a poor suburb of Paris, then spread in all of France, lasting about three weeks. Remarkably, although there were no displacements of rioters, the riot activity did traveled. Daily national police data to which we had access have allowed us to take advantage of this natural experiment to explore the dynamics of riot propagation. Here we show that an epidemic-like model, with less than 10 free parameters and a single sociological variable characterizing neighborhood deprivation, accounts quantitatively for the full spatio-temporal dynamics of the riots. This is the first time that such data-driven modeling involving contagion both within and between cities (through geographic proximity or media) at the scale of a country is performed. Moreover, we give a precise mathematical characterization to the expression "wave of riots", and provide a visualization of the propagation around Paris, exhibiting the wave in a way not described before. The remarkable agreement between model and data demonstrates that geographic proximity played a major role in the riot propagation, even though information was readily available everywhere through media. Finally, we argue that our approach gives a general framework for the modeling of spontaneous collective uprisings.

Introduction

During the Autumn of 2005, France experienced the longest and most geographically extended riot of the contemporary history of Europe[1, 2]. Without any political claims nor leadership, localization was mainly limited to the "banlieues" (suburbs of large metropolitan cities), where minority groups are largely confined. Contrary to the London "shopping riots" of 2011, rioting in France essentially consisted in car destruction and confrontations with the police. The triggering event took place in a deprived municipality at the north-east of Paris: on October 27, 2005, two youths died when intruding into a power substation while trying to escape a police patrol. Inhabitants spontaneously gathered on the streets with anger. The extension in time and space of these riots – three weeks, more than 800 municipalities hit across all of France – produces a *natural experiment* amenable to analysis thanks to detailed police data available[3]. The present work aims at analyzing these data through a mathematical model that sheds new light on qualitative features of the riots as instances of collective human behavior[4, 5].

Several works[6, 7, 8, 9, 10, 11, 12, 13, 14, 15] have developed mathematical approaches to rioting dynamics, and their sociological implications have also been discussed[16]. The 1978 article of Burbeck *et al.* [10] pioneered quantitative epidemiological modeling to study the dynamics of riots. Very few works followed the same route, but similar ideas have been applied to other social phenomena such as the spreading of ideas or rumors[17] and the viral propagation of *memes* on the Internet[18]. This original epidemiological modeling was however limited to the analysis within single cities, without spatial extension. Few studies aim at quantitatively describing the spatial spread of riots coupled with contagion. Notable exceptions are an analysis of the US ethnic riots[13], correlating the occurrences of riots in different cities, and the studies of the 2011 London riots[7, 11], which describe the displacements of rioters. In contradistinction with the London case, media reports and case studies[3, 19] show that the 2005 French rioters remained localized in a particular neighborhood of each municipality. However, the riot itself did travel. Our data give a unique opportunity to study this contagion from city to city at the level of a whole country.

We introduce here a compartmental epidemic model of the Susceptible-Infected-Recovered (SIR) type[20, 21, 22]. Infection takes place through contacts within cities as well as through other short- and long-range interactions arising from either social networks or media coverage. These influence interactions are the key to riots spreading over the discrete set of French municipalities. In particular, diffusion based on geographic proximity played a major role in generating a kind of riot wave around Paris which we exhibit here. This is substantiated by the remarkable agreement between the data and the model at various geographic scales. Indeed, one of our main findings is that less than ten free parameters together with only one sociological variable (the size of the population of poorly educated young males) are enough to accurately describe the complete spatio-temporal dynamics of the riots.

The qualitative features taken into account by our model – the role of a single triggering effect, a 'social tension' buildup, a somewhat slower and rather smooth relaxation, and local as well as global spreading –, are common to many riots. This suggests that our approach gives a general framework for the modeling of spontaneous collective uprising.

Results

The 2005 French riots dataset

We base our analysis here on the daily crime reports[3] of all incidents recorded by the French police at the municipalities (corresponding to the French "communes") under police authority, which cover municipalities with a population of at least 20,000 inhabitants. Such data, on the detailed time course of riots at the scale of hours or days, and/or involving a large number of cities, are rare. In addition, as an output of a centralized national recording procedure applied in all national police units operating at the local level, the data are homogeneous in nature. These qualities endow these data with a unique scientific value. We adopt a simple methodology for quantifying the rioting activity: we define as a single event any rioting-like fact, leaving aside its nature and its apparent intensity. Thus, each one of "5 burnt cars", "police officers attacked with stones" or "stoning of firemen", is labeled as a single event. We thus get a dataset composed of the number of riot-like events for each municipality, every day from October 26 to December 8, 2005, a period of 44 days which covers the three weeks of riots and extends over two weeks after.

The top of Fig. 1a shows two typical examples of the time course of the number of events for municipalities (see also SI Fig. S1, same plots for the 12 most active Île-de-France municipalities). A striking observation is that there is a similar up-and-down dynamics at every location, showing no rebound, or, if any, hardly distinguishable from the obvious stochasticity in the data. This pattern is similar to the one observed for the US ethnic riots[10]. In addition, we observe the same pattern across different spatial scales (see Fig. 1 and SI Fig. S3). Moreover, this pattern shows up clearly despite the difference in amplitudes (see also Fig. 4a and 4b). This fractal-like property suggests an underlying mechanism for which geographical proximity matters. Finally, the rioting activity appears to be on top of a background level: as can be seen on Fig. 1, the number of events relaxes towards the very same level that it had at the outset of the period. Actually, in the police data, one cannot always discriminate rioting facts from ordinary criminal ones, such as the burning of cars unrelated to collective uprising. For each location, we assume that the stationary background activity corresponds to this "normal" criminal activity.

Modeling framework

We now introduce our modeling approach. Section Materials and Methods provides the full model and numerical details. The model features presented below are based on the analysis at the scale of municipalities. However, since aggregated data at the scale of *départements* present a pattern similar to the data of municipalities, we also fit the model at the département scale, as if the model assumptions were correct at the

scale of each département. A 'site', below, is a municipality or a département depending on the scale considered.
As the rioting activities are described by a discrete set of events, we assume an underlying point process[23] characterized by its mean value. Assuming statistical independence between the rioting and the criminal activities, the expected number of events at each site k ($k = 1, ..., K$, K being the number of sites), is the sum of the mean (time independent) background activity λ_{bk}, and of the (time dependent) rioting activity, $\lambda_k(t)$. In fitting the model to the data, we take the background activity λ_{bk} as the average number of events at the considered site over the last two weeks of our dataset. Assuming Poisson statistics (which appears to be in good agreement with the data, see Materials and Methods), the means $\lambda_k(t)$ fully characterize the rioting activities. We make the assumption that this number of events $\lambda_k(t)$ is proportional to the local number of rioters, $I_k(t)$:

$$\lambda_k(t) = \alpha I_k(t).$$

We model the coupled dynamics of the set of $2 \times K$ variables, the numbers $I_k(t)$ of rioters (*infected* individuals in the terminology of the SIR model) and the numbers $S_k(t)$ of individuals *susceptible* to join the riot, by writing an epidemic SIR model[21, 22] in a form suited for the present study. This gives the coupled dynamics of the $\lambda_k(t)$ and of the associated variables, $\sigma_k(t) \equiv \alpha S_k(t)$. We fit the model to the data by considering a discrete time version of the equations (events are reported on a daily basis), and by optimizing the choice of the model free parameters with a maximum likelihood method.

We now sketch the proposed SIR model. We assume homogeneous interactions *within* each municipality (a hypothesis justified by the coarse-grained nature of the data, and by the absence of displacements of rioters), and influences between sites. The model thus belongs to the category of metapopulation epidemic models [24]. Motivated by the relative smoothness of the time course of events, we make the strong assumption that, at each site, there is a constant rate at which rioters leave the riot. This parameter aggregates the effects of different factors – arrests, stringent policing, other sources of deterrence, fear, fatigue, etc. –, none of them being here modeled explicitly. In addition, since there are almost no rebounds of rioting activity, we assume that there is no flux from *recovered* (those who left the riot) to *susceptible* (and thus we do not have to keep track of the number of recovered individuals).

In the epidemic of an infectious disease, contagion typically occurs by dyadic interactions, so that the probability for a susceptible individual to be infected is proportional to the *fraction* of infected individuals – leading to equations written in terms of the fractions of infected and susceptible individuals. In the present context, contagion results from a bandwagon effect[25, 12, 4]. The probability of becoming a rioter is thus a function of the *number* of rioters, hence of the number of events given the above hypothesis. This function is nonlinear since, being a probability, it must saturate at some value (at most 1) for large rioting activities.

Single site epidemic modeling

As a first step, following Burbeck *et al.*, we ignore interactions between sites, and thus specify the SIR model for each site *separately.* We consider here one single site (and omit the site index k in the equations). Before a triggering event occurs at some time t_0, there is a certain number $S_0 > 0$ of susceptible individuals but no rioters. At t_0 there is an exogenous shock leading to a sudden increase in the I population, hence in λ, yielding an initial condition $\lambda(t_0) = A > 0$. From then on, the rioting activity at a single (isolated) site evolves according to:

$$\begin{cases} \dfrac{d\lambda(t)}{dt} = -\omega\, \lambda(t) + \beta\, \sigma(t)\, \lambda(t), \\ \dfrac{d\sigma(t)}{dt} = -\beta\, \sigma(t)\, \lambda(t), \end{cases}$$

where β is a susceptibility parameter. Here we work within a linear approximation of the probability to become infected, which appears to provide good results for the single site modeling. The condition for the riot to start after the shock is that the reproduction number[20] $R_0 = \beta\sigma(t_0)/\omega$ is greater than 1. In such a case, from $t = t_0$ onward, the number of infected individuals increases, passes through a maximum and relaxes back towards zero.

We obtain the initial condition $\sigma_0 = \sigma(t_0) = \alpha S_0$ from the fitting procedure. Thus for each site, we are left with five free parameters to fit in order to best approximate the time course of the rioting events: ω, β, t_0, A and σ_0.

By showing examples at different scales, Fig. 1 (b, red curves) illustrate the remarkable quality of the resulting fits (see also SI Fig. S1). The obvious limitation is that fitting all the 853 municipalities present in the dataset amounts to determining $853 \times 5 = 4265$ free parameters. The fit is very good but meaningless (overfitting) for sites with only one or two events. In addition, these single site fits cannot explain why the riot started on some particular date at each location. Fitting the single site model requires one to assume that there is one exogenous specific shock at a specific time at each location, whereas the triggering of the local riot actually results from the riot events that occurred before elsewhere. Nevertheless, we see that everywhere the patterns are compatible with an epidemic dynamics and that through the use of the model it is possible to fill in missing data and to smooth the data (filtering out the noise). As a result of this filtering, the global pattern of propagation becomes more apparent. Indeed, looking at the Paris area, one observes a kind of wave starting at Clichy-sous-Bois municipality, diffusing to nearby locations, spreading around Paris, and eventually dying out in the more wealthy south-west areas (see Fig. 3a and SI Video 1).

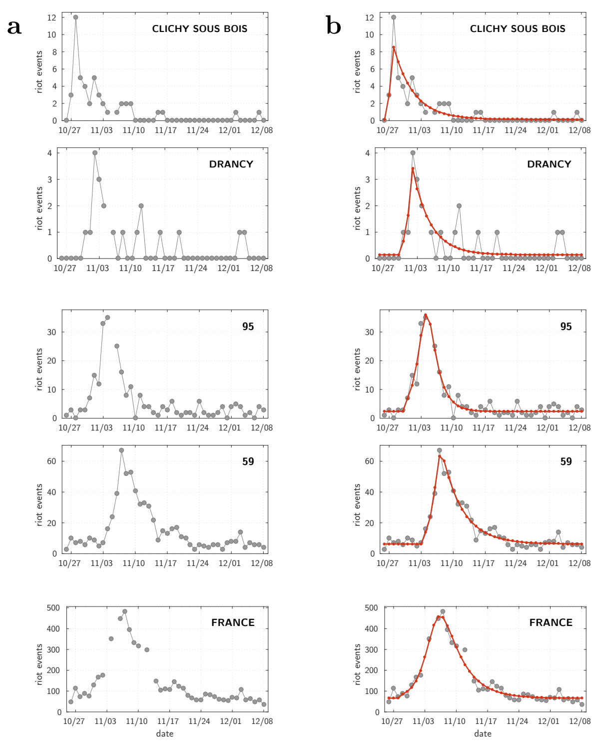


Figure 1: Data and single site fits at different scales. Top: municipalities; Middle: départements; Bottom: all of France (see Materials and Methods for a description of these administrative divisions). (a) Raw data. (b) Same data (gray dots) along with the calibrated model (red curve). Here and in all the other figures involving time, the thin dotted lines divide the time axis into one week periods, starting from the date of the shock, October 27, 2005.

Modeling the riot wave

We now take into account the interactions between sites, specifying the global metapopulation SIR model. Among the K sites under consideration, only one site k_0, the municipality of Clichy-sous-Bois (département 93 when working at département scale), undergoes a shock at a time t_0, 27 October 2005. To avoid a number of parameters which would scale with the number of sites, we choose here all free parameters to be site-independent (in Materials and Methods we give a more general presentation of the model). The resulting system of $2 \times K$ coupled equations writes as follows: for $t > t_0$, for $k = 1, ..., K$,

$$\begin{cases} \dfrac{d\lambda_k(t)}{dt} = -\omega\, \lambda_k(t) + \sigma_k(t)\, \Psi(\Lambda_k(t)), \\ \dfrac{d\sigma_k(t)}{dt} = -\sigma_k(t)\, \Psi(\Lambda_k(t)). \end{cases}$$

Here ω is the site-independent value for the recovering rate. For the interaction term we consider that at any site k the probability to join the riot is a function Ψ of a quantity $\Lambda_k(t)$, the global activity as 'seen' from site k. This represents how, on average, susceptible individuals feel concerned by rioting events occurring either locally, in neighboring cities, or anywhere else in France. We make the hypothesis that the closer the events, the stronger their influence. We thus write that $\Lambda_k(t)$ is a weighted sum of the rioting activities occurring in all sites,

$$\Lambda_k(t) = \sum_j W_{kj}\, \lambda_j(t),$$

where the weights W_{kj} are given by a decreasing function of the distance between sites k and j. We tested several ways of choosing the weights and obtained the best results for two types of parametrizations. One is a power law decay with the distance, motivated by several empirical observations[26, 27, 28]. The second option is the sum of an exponential decay and of a constant term. Both involve two parameters, a proximity scale d_0 and, respectively, the exponent δ and the strength ξ of the constant term. The constant term may correspond to proximity perceived in terms of cultural, socio-economic characteristics.
For the (site independent) function $\Psi(.)$, we consider either its linear approximation, writing

$$\Psi(\Lambda_k(t)) \; = \; \beta \, \Lambda_k(t) \; = \; \beta \sum_j W_{kj} \, \lambda_j(t),$$

with the susceptibility β as a site-independent free parameter, or various non-linear cases, involving up to four parameters.

Finally, we have to make the crucial choice of the initial values $\sigma_{k,0} = \sigma_k(t_0)$, specific to each site. By definition, they must be proportional to the size of the initial susceptible population. We make the hypothesis that the latter scales with the size of a population defined by a sociological specification. Thus we assume

$$\sigma_{k,0} = \zeta_0 \; N_k,$$

where ζ_0 is a site-independent free parameter, and N_k is the size of a reference population provided by INSEE, the French national statistics and economic studies institute. The results we present here take as reference the population of males aged between 16 and 24 out-of-school with no diploma – see Materials and Methods for a discussion on the choice of the reference population.

For the whole dynamics, we are left with only six free parameters in the linear case: ω, A, ζ_0, d_0, δ or ξ, and β, and up to nine free parameters in the non linear case: five as for the linear case, ω, A, ζ_0, d_0, δ or ξ, and, in place of β, up to four parameters characterizing the function Ψ.

The above model, in the case of the linear approximation, makes links to the classical spatially continuous, non-local, SIR model[29] (see section *Links to the original spatially continuous SIR model* in Materials and Methods, and SI Videos 3 and 4). In dimension one, when the space is homogeneous, we know[30] that traveling waves can propagate, quite similar to the way the riot spread around Paris as exhibited in the previous section. The new class of models we have introduced is however somewhat different and more general, and raises several open mathematical questions. The next section shows the wave generated by our global model and the fit to the data.

Fitting the data: the wave around Paris

We first focus on the contagion around Paris, characterized by a continuous dense urban fabric with deprived neighborhoods. There are 1280 municipalities in Île-de-France. Among the ones under police authority (a total of 462 municipalities, for all of which we have data), 287 are mentioned for at least one riot-like event. For all the other municipalities, which are under "gendarmerie" authority (a military status force with policing duties), we have no data. Since their population size is small, we expect the associated numbers of riot events to be very small if not absent, so that these sites have little influence on the whole dynamics. We choose the free parameters with the maximum likelihood method, making use of the available data, i.e. the 462 municipalities. However, the model simulations take into account all the 1280 municipalities. Results are presented for a power law decrease of the weights and a non linear function Ψ characterized by 3 parameters. Thus, we have here a total of 8 free parameters: $\omega, A, \zeta_0, d_0, \delta$, in addition to three for the non-linear function.

Figures 2, 3b and the SI Video 2 illustrate the main results. Figure 2 compares the model and the data on four aspects: time course in each département (a), amplitude of the events (b), date at which the number of events is maximum (c), and spatial distribution of the riots (d). The global model with a single shock correctly reproduces the up-and-down pattern at each location, as illustrated on Fig. 2 at département scale. One can note the preservation of the smooth relaxation at each site, despite the influence of other (still active) sites. This can be understood from the SIR dynamics: at a given location, the relaxation term $(-\omega\lambda_k)$ dominates when there is no more enough susceptible individuals, so that the local dynamics becomes essentially independent of what is occurring elsewhere. Quite importantly, these local patterns occur at the correct times. One sees that the date of maximum activity spreads over several days and varies across locations, which reflects the propagation of the riot.

On the SI Video 2 one can see the wave generated by the model. Figure 3b shows a sketch of this wave, with one image every 4 days – which corresponds to the timescale found by the parameter optimization, $1/\omega \sim 4$ days. On the same figure, this timeline has to be compared with the one obtained from the single site

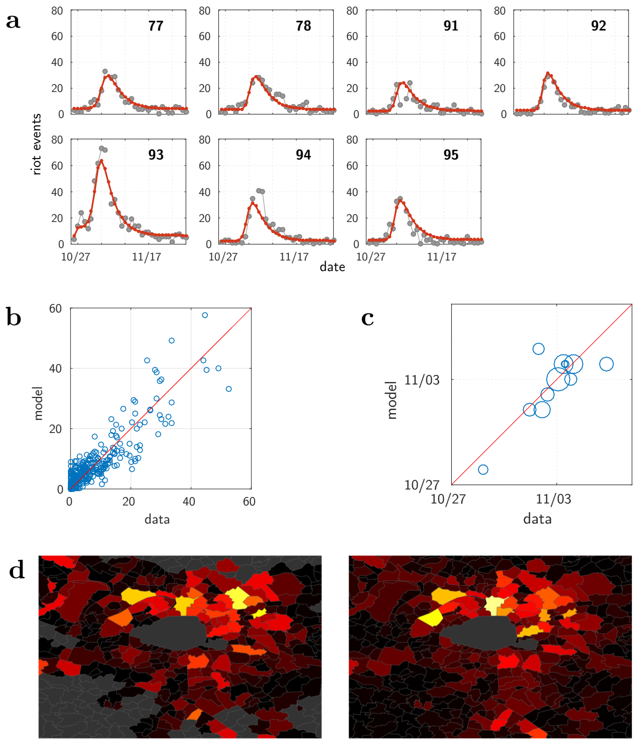


Figure 2: Results: Île-de-France region, model calibration at the scale of municipalities. (a) Time course of the riots: data (dots) and model (continuous curves) – results presented here are aggregated by département. (b) Total number of events, model vs. data. Each dot represents one municipality. In order to compute the total sum of events in the data, missing values were filled using linear interpolation. (c) Temporal unfolding, model vs. data. Date (unit=day) of the maximum rioting activity, shown for the 12 most active municipalities (those with more than 30 events). Each circle has a diameter proportional to the size of the reference population of the corresponding municipality. The red lines depict the identity diagonal line. (d) Geographic map of the total rioting activity. Data (left) vs Model (right), shown for the inner suburb of Paris (the "petite couronne", départements 92, 93 and 94). For each municipality, the colour codes the total number of events (the warmer the larger, same scale for both panels; grey areas: data not available).

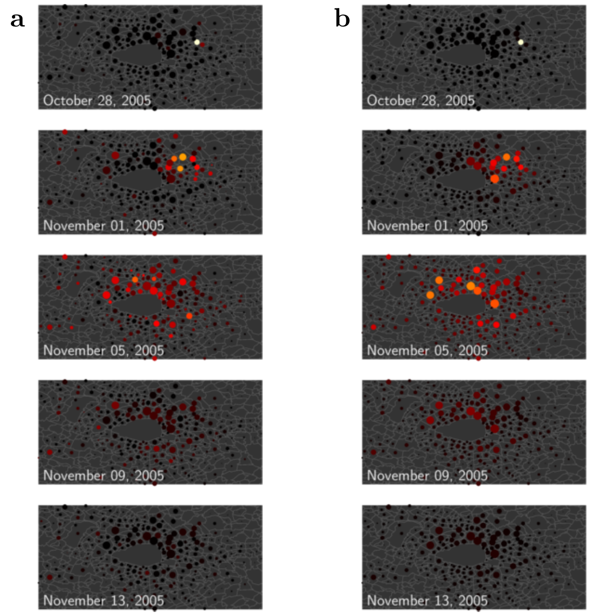


Figure 3: Timeline of the riots in Paris area (Île-de-France). The rioting activity is shown every 4 days, starting on the day following the triggering event (top). (a) data smoothed by making use of the fitted values given by the single site models (see SI Video 1 for the full dynamics). (b) dynamics generated by the global model (see SI Video 2). The map shows the municipality boundaries, with Paris at the center. For each municipality under police authority, a circle is drawn with an area proportional to the size of the corresponding reference population. The color represents the intensity of the rioting activity: the warmer the color, the higher the activity. Figure best viewed magnified in the electronic version.

fits. One can see the good agreement, except for few locations where the actual rioting activity occurs earlier than predicted by the global model. A most visible exception is Argenteuil municipality (north-west of Paris on the map, see Fig. 3, second images from the top), where the Minister of Interior made a speech (October 25) perceived as provocative by the banlieues residents. This could potentially explain the faster response to the triggering event.

The calibrated model gives the time course of the expected number of events in all municipalities, including those under 'gendarmerie' authority for which we do not have any data. For all of the later, the model predicts a value remaining very small during all the studied period, except for one, the municipality of Fleury-Mérogis (see Fig. 2d, South of the map). Remarkably, searching in the media coverage, we found that a kindergarten has been burnt in that municipality at that period of time (Nov. 6).

Fitting the data: the wave across the whole country

We now show that the same model reproduces the full dynamics across the whole country. We apply our global model considering each one of the départements of metropolitan France (except Corsica and Paris, hence 93 départements) as one homogeneous site – computing at municipality scale would be too demanding (more than 36,000 municipalities). The best model options lead to 9 free parameters: $\omega, A, \zeta_0, d_0, \xi$, the susceptibility β (using the linear approximation), and different susceptibility values at three specific locations. In this case too, the resulting fit is very good, as illustrated on Fig. 4 (see also SI Fig. S3). Figures 4a and 4b show the results for the 12 most active départements. Figure 4c compares model and data on the total number of events, and Fig. 4d on the date of the maximum activity. For the latter, the data for the Île-de-France municipalities (Fig. 2c) are reported. One sees that the wave indeed spread over all France, with the dynamics in Paris area essentially preceding the one elsewhere. Remarkably, as shown on Fig. 5, the model also predicts correctly the time of arrival and the weak amplitude of the riot wave for sites where very few rioting events have been recorded.

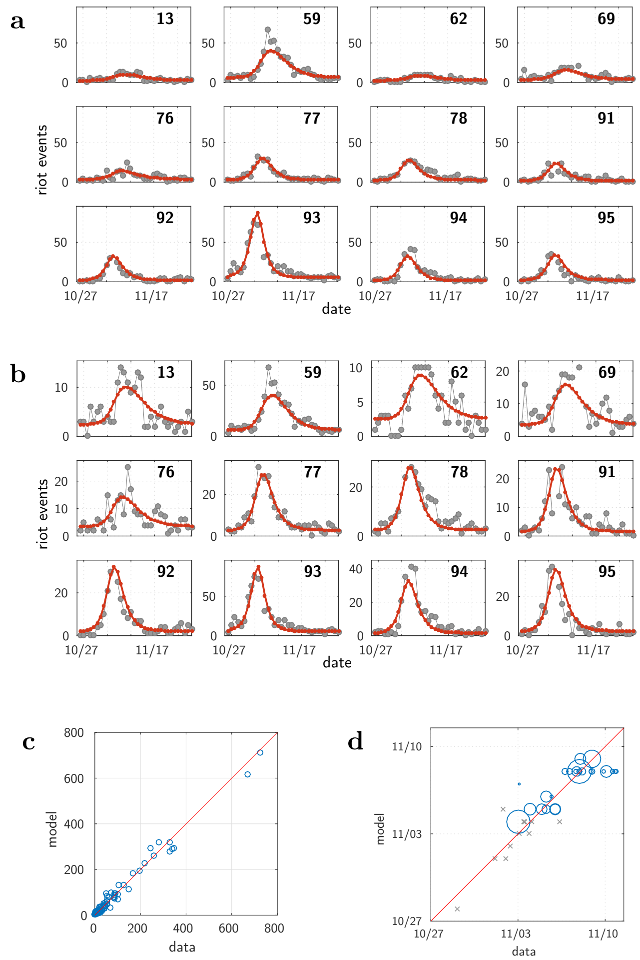


Figure 4: Results: All of France, model calibration at the scale of the départements. (a) Time course of the riots in France: data (dots) and model (continuous curves). Only the 12 most active départements are shown. The plots share a common scale for the number of events. (b) Same as (a), but with relative scales. (c) Total number of events. (d) Temporal unfolding (date when the number of riot events reaches its maximum value), shown for the départements having more than 60 events. Each blue circle has a diameter proportional to the reference population of the corresponding départements. The grey crosses remind the plot locations of the major municipalities of Île-de-France shown in Fig. 2c. The red lines depict the identity diagonal line.

Discussion

Studying the dynamics of riot propagation, a dramatic example of large-scale social contagion, is difficult due to the scarcity of data. The present work takes advantage of the access to detailed national police data on the 2005 French riots that offer both the timescale of the day over a period of 3 weeks, and the geographic extension over the country. These data exhibit remarkable features that warrant a modeling approach. We have shown that a simple epidemic-like model with less than 10 free parameters, combining contagion both within and between cities, allows one to reproduce the daily time course of events, revealing the wave of contagion. A crucial model ingredient is the choice of a single sociological variable, taken from the census statistics as a proxy for calibrating the size of the susceptible population. It shows that the wave propagates in an excitable medium of deprived neighborhoods. Since our analysis shows that diffusion by geographic proximity is a key

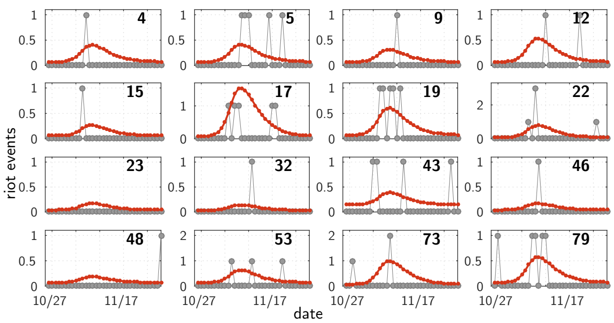


Figure 5: Results, calibration at the scale of the départements: minor sites. Even where the number of events was very small, the model predicts the sites to be hit by the wave, with a small amplitude and at the correct period of time.

underlying mechanism, having the outbreak location surrounded by a dense continuum of deprived neighborhoods made the large-scale contagion possible.

What lesson on human behavior can we draw from our analysis? Firstly, 'geography matters'[26]: despite the modern communication media, physical proximity is still a major feature in the circulation of ideas, here of riotous ideas. Secondly, strong interpersonal ties are at stake for dragging people into actions that confront social order. The underlying interpretation is that interpersonal networks are relevant for understanding riot participation. Human behavior is a consequence not only of individuals' attributes but also of the strength of the relation they hold with other individuals[31]. Strong interpersonal connections to others who are already mobilized draws new participants into particular forms of collective action such as protest, and identity (ethnic or religious based) movements[32, 33]. Thirdly, concentration of socio-economic disadvantage facilitates formation of a sizable group and therefore involvement in destruction: *numbers* of rioters in the model (rather than proportions) can be interpreted as an indirect indication of risk assessment before participating to a confrontation with the police[34, 5]. Rioters seem to adopt a rational behavior and only engage in such event when their number is sufficient.

The question of parsimony is of the essence in our modeling approach: an outstanding question was to understand whether a limited number of parameters might account for the observed phenomena at various scales and in various locations. We answer this question positively here, thus revealing the existence of a general mechanism at work. The similar up-and-down pattern observed for the US ethnic riots in different cities suggests that this process is indeed common to a larger class of spontaneous riots. The 2005 riot propagation from place to place is also reminiscent of the spreading of other riots, such as the one of the food riots in the late eighteen century in the UK[35], and of the ethnic riots in US cities – also coined as a wave[13]. We thus believe that the modeling approach introduced here provides a general framework, different from the statistical/econometric approach, that may be adapted to the detailed description of the propagation of spontaneous collective uprisings – notably, the interaction term in our SIR model can be modified to include time delays (time for the information to travel) or to take into account time integration of past events.

Materials and Methods

French administrative divisions

The three main French administrative divisions are: the "commune", which we refer to as municipality in the paper (more than $36,000$ communes in France); at a mid-level scale the 'département', somewhat analogous to the English district (96 départements in Metropolitan France, labeled from 1 to 95, with $2A$ and $2B$ for Corsica); the 'région' aggregating several neighboring départements (12 in metropolitan France, as of 2016). At a given level, geographic and demographic characteristics are heterogeneous. However the typical diameter of a département is ~ 100 km, and the one of a region, ~ 250 km.

There are two national police forces, the 'police' and the 'gendarmerie' (a civilian like police force reporting to the ministry of Interior whose agents have a military status in charge of policing the rural parts of the country). Most urbanized areas (covering all municipalities with a population superior to 20000) are under police authority. The more rural ones are under gendarmerie authority. The available data for the present study only concern the municipalities under police authority, except Paris, for which we lack data (but was not much affected by the riots).

The full list of the municipalities is available on the French government website,
https://www.data.gouv.fr/fr/datasets/competence-territoriale-gendarmerie-et-police-nationales/.

Dataset

The present analysis is based on the daily crime reports of all incidents of civil unrest reported by the French police at the municipalities under police authority[3] (see above). The daily reports are written in natural language, and have been encoded to allow for statistical treatment. From the reports we selected only facts related to urban violence. Some facts are reported more than once (a first time when the fact was discovered, and then one or two days later e.g. if the perpetrators have been identified). We carefully tried to detect and suppress double counting, but some cases may have been missed. There are also a few missing or incomplete data – notably in the nights when rioting was at its maximum, as the police was overwhelmed and reported only aggregated facts, instead of details city per city.

In the police data, incidents in relation with the riots are mostly cases of vehicles set on fire (about 70%), but also burning of public transportation vehicles, public buildings, of waste bins, damages to buses and bus shelters, confrontations between rioters and police, etc. Facts have been encoded with the maximum precision: day and time of the fact, who or what was the target, the type of damage, the number and kind of damaged objects and the number and quality of persons involved, whenever these details are mentioned in the report. In the present work, as explained in the main text, from these details, we compute a daily number of events per municipality. We generated two different datasets, one for the events (with a total of 6877 entries concerning 853 municipalities) and the other for the arrests (2563 entries), that in forthcoming work will serve to characterize the rioters as well as the severity of the penalties. In future work we also plan to explore the rioting events beyond the sole number of events as studied here. One can expect to see what has been for instance the role of curfews and other deterrent effects (in space and time).

In addition, we have also built a specific dataset of media coverage, from both local newspapers and national TV and radio broadcasts. In ongoing work we extend our modeling framework by considering the coupling between the dynamics of riot events and the media coverage.

Background activity

The rioting activity appears to be above a constant level which most likely corresponds to criminal activities (an average of an order of 100 vehicles are burnt every day in France, essentially due to criminal acts not related to collective uprising). Since this background activity has the same level before and after the riot, we assumed that the riot and the criminal activities are statistically independent. In addition, we also considered alternative models where the background activity is an equilibrium state, the riot being a transient excited state. Such models would predict an undershooting of the activity just after the end of the riot – more exactly a relaxation with damped oscillations –, but the data do not exhibit such behavior.

For the fits, the background activity λ_b is taken as the mean activity over the last two weeks in our dataset (November 25 to December 8), period that we can consider as the tail of the data for which there is no longer any riot activity (see Fig. 1). For the sites with a non zero number of events in the tail, we observed that this base rate is proportional to the size of the reference population chosen for calibrating the size of the susceptible population. For the sites where the number of events in the tail is either zero or unknown (which is the case for a large number of small municipalities, in particular the ones under gendarmerie authority), one needs to give a non zero value of the corresponding base rate in order to apply the maximum likelihood method (see below). We estimated it from the size of the reference population (set as 1 when it is 0), using the latter proportionality coefficient (with a maximum value of λ_b set to one over the length of the tail, *ie* 1/14).

Poisson noise assumption

In order to calibrate the model to the data, we assumed Poisson statistics, although we do not claim that the underlying process is exactly of Poisson nature. However, it is a convenient working hypothesis for numerical reasons (see below *Free parameters: numerical optimization*). From a theoretical point of view, this choice is a priori appropriate as we deal with discrete values (often very small). In addition, we have seen that the data suggests that a same kind of model is relevant at different scales: this points towards infinitely divisible distributions, such as the Poisson distribution – the sum of several independent Poisson processes still being a Poisson process. We wish to emphasize that we show that the statistics of the data are compatible with the Poisson hypothesis. Under such a hypothesis, the variance is equal to the mean. Figure 6a shows that, for each département, if we look at the last two weeks exhibiting a stationary behavior, the variance/mean relationship is indeed in good agreement with a Poisson hypothesis. Figure 6b shows Poissonness plot[36] for each of the 12 regions, which gives additional support to the Poisson noise property.

For completeness, we recall the meaning of a Poissonness plot. One has a total number of observations n. Each particular value x is observed a certain number of

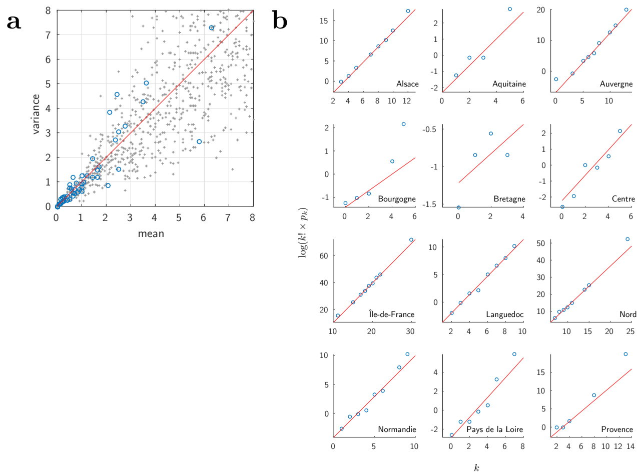


Figure 6: Test of Poisson statistics assumption. (a) Mean and variance computed on the tail of the data (taken as the last two available weeks considered as background noise, without rioting activity). Each circle corresponds to one département. For comparison, we generated fake Poisson data mimicking the typical values observed in the dataset: the gray crosses give the sample mean and variance for 14 realizations of 1000 Poisson processes, with true means randomly generated between 0 and 10. (b) For each region, Poissonness plot computed for the tail of the data.

times n_x, hence an empirical frequency of occurrence n_x/n. If the underlying process is Poisson with mean λ, then one must have $\log(x!n_x/n) = -\lambda + x\log(\lambda)$. Thus, in that case, the plot of the quantity $\log(x!n_x/n)$ (the blue circles in Fig. 6b) as a function of x should fall along a straight line with slope $\log(\lambda)$ and intercept $-\lambda$ (the red lines on Fig. 6b).

Single site epidemic modeling

We detail here the compartmental SIR model when applied to each site *separately* (each municipality, or, after aggregating the data, for each département). Let us consider a particular site (we omit here the site index k in the equations). At each time t there is a number $S(t)$ of individual *susceptible* to join the riot, and $I(t)$ of *infected* individuals (rioters). Those who leave the riots become *recovered* individuals. Since we assume that there is no flux from *recovered* to *susceptible*, we do not have to keep track of the number of recovered individuals. Initially, before a triggering event at some time t_0 occurs, there is a certain number $S_0 > 0$ of susceptible individuals but no rioters, that is, $I(t) = 0$ for all $t \leq t_0$. At t_0 there is an exogenous shock and the number of rioters becomes positive $I(t_0) = I_0 > 0$. From there on, neglecting fluctuations, the numbers of rioters and of susceptible individuals evolve according to the following set of equations:

$$\begin{cases} \dfrac{dI(t)}{dt} = -\,\omega\, I(t) + S(t)\, P(s \to i, t) & \text{(1a)} \\ \dfrac{dS(t)}{dt} = -\, S(t)\, P(s \to i, t) & \text{(1b)} \end{cases}$$

Let us now explain this system of equations. In Eq. (1a), ω is the constant rate at which rioters leave the riot. The second term in the right hand side of Eq. (1a) gives the flux from susceptible to infected as the product of the number of susceptible individuals, times the probability $P(s \to i, t)$ for a susceptible individual to become infected. The second equation, Eq. (1b), simply states that those who join the riot leave the subpopulation of susceptible individuals.

We now specify the probability to join the riot, $P(s \to i, t)$ (to become infected when in the susceptible state). In line with accounts of other collective uprising phenomena[5], testimonies from participants in the 2005 riots suggest a bandwagon effect: individuals join the riot when seeing a group of rioters in action. Threshold decision models[25, 12] describe this herding behavior assuming that each individual has a threshold. When the herd size is larger than this threshold the individual joins the herd. Granovetter[12] has specifically applied such a model to riot formation, the threshold being then the number of rioters beyond which the individual decides to join the riot. Here we make the simpler hypothesis that the probability to join the riot does not depend on idiosyncratic factors, and is only an increasing function of the *total* number of rioters at the location (site) under consideration. It is worth emphasizing that this herding behavior is in contrast with the epidemic of an infectious disease, where contagion typically occurs from dyadic interactions, in which case the probability is proportional to the *fraction* of infected individuals, $I(t)/S_0$. Being a probability, $P(s \to i, t)$ must saturate at some value (at most 1) for large I, and is thus a nonlinear function of I. Nevertheless, we will first assume that conditions are such that we can approximate $P(s \to i, t)$ by its linear behavior: $P(s \to i, t) \sim \kappa I(t)$ (but note that κ does not scale with $1/S_0$) and discuss later a different specification for this term.

Given the assumption $\lambda(t) = \alpha I(t)$, it is convenient to define

$$\sigma(t) = \alpha S(t) \tag{2}$$

so that the riot dynamics at a single (isolated) site is described by:

$$\begin{cases} \dfrac{d\lambda(t)}{dt} = -\omega\,\lambda(t) + \beta\,\sigma(t)\,\lambda(t) & \text{(3a)} \\ \dfrac{d\sigma(t)}{dt} = -\beta\,\sigma(t)\,\lambda(t) & \text{(3b)} \end{cases}$$

where $\beta \equiv \frac{\kappa}{\alpha}$. Initially $\lambda = 0$, which is a fixed point of this system of equations. With $\sigma(t_0) = \sigma_0 > 0$, the riot starts after the shock if the reproduction number[20] $R_0 \equiv \beta\sigma_0/\omega = \kappa S_0/\omega$ is greater than 1. In such a case, from $t = t_0$ onward, the number of infected individuals first increases, then goes through a maximum and eventually relaxes back towards zero. Because κ is *not* of order $1/S_0$, this condition seems too easy to satisfy: at any time, any perturbation would initiate a riot. One may assume that the particular parameter values allowing one to fit the data describe the state of the system at that particular period. Previous months and days of escalation of tension may have led to an increase in the susceptibility κ, or in the number of susceptible individuals S_0.

Non local contagion model

We give here the details on the global SIR model, with interactions between sites. We have a discrete number K of sites, with homogeneous mixing within each site, and interactions between sites. At each site k, there is a number S_k of 'susceptible' individuals, I_k of 'infected' (rioters), and R_k of 'recovered' individuals. As above, there is no flux from recovered to susceptible (hence we can ignore the variables R_k), and individuals at site k leave the riot at a constant rate ω_k. Assuming homogeneous mixing in each site, the dynamics is given by the following set of equations:

$$\begin{cases} \dfrac{dI_k(t)}{dt} = -\,\omega_k\,I_k(t) + S_k(t)\,P_k(s \to i, t) & \text{(4a)} \\ \dfrac{dS_k(t)}{dt} = -\,S_k(t)\,P_k(s \to i, t) & \text{(4b)} \end{cases}$$

with the initial conditions $t < t_0$ $I_k(t) = 0$, $S_k(t) = S_{k0} > 0$, and at $t = t_0$, a shock occurs at a single location k_0, $I_k(t_0) = I_0 > 0$. In the above equations, ω_k is the local recovering rate, and $P_k(s \to i, t)$ is the probability for a s-individual at location k to become a rioter at time t.

We now write the resulting equations for the λ_k. We assume the rioting activity to be proportional to the number of rioters:

$$\lambda_k(t) = \alpha I_k(t) \tag{5}$$

Note that different hypothesis on the dependency of λ_k on I_k could be considered. For instance we tested $\lambda_k \sim (I_k)^q$ with some exponent q coming as an additional free parameter. In that case, the optimization actually gives that q is close to 1.

Multiplying each side of (4) by α, one gets

$$\begin{cases} \dfrac{d\lambda_k(t)}{dt} = -\,\omega_k\,\lambda_k(t) + \sigma_k(t)\,P_k(s \to i, t) & \text{(6a)} \\ \dfrac{d\sigma_k(t)}{dt} = -\,\sigma_k(t)\,P_k(s \to i, t) & \text{(6b)} \end{cases}$$

where as before we introduce $\sigma_k(t) = \alpha S_k(t)$. Taking into account the hypothesis on the linear dependency of the number of event in the number of rioters, (5), we write $P(s \to i, t)$ directly in term of the λs:

$$P_k(s \to i, t) = \Psi_k(\Lambda_k(t)) \tag{7}$$

where $\Lambda_k(t)$ is the activity 'seen' from site k (see main text):

$$\Lambda_k(t) \equiv \sum_j W_{kj}\lambda_j(t) \tag{8}$$

where the weights W_{kj} are given by a decreasing function of the distance $\text{dist}(k,j)$ between sites k and j: $W_{kj} = W(\text{dist}(k,j))$ (see below). The single site case is recovered for $W_{kj} = \delta_{k,j}$.

In the linear approximation,

$$\Psi_k(\Lambda) = \beta_k\,\Lambda, \tag{9}$$

in which case one gets the set of equations

$$\begin{cases} \dfrac{d\lambda_k(t)}{dt} = -\,\omega\,\lambda_k(t) + \,\beta\,\sigma_k(t)\,\displaystyle\sum_j W_{kj}\lambda_j(t) & \text{(10a)} \\ \dfrac{d\sigma_k(t)}{dt} = -\,\beta\,\sigma_k(t)\,\displaystyle\sum_j W_{kj}\lambda_j(t). & \text{(10b)} \end{cases}$$

The form of these equations is analogous to the ones of the original distributed contacts continuous spatial SIR model [30] (see below) but here with a discrete set of spatial locations.
In the non-linear case, we choose parameters for $\Psi_k(\Lambda)$ in order to have a function (i) being zero when there is no rioting activity; (ii) which saturates at a value (smaller or equal to 1) at large argument; (iii) with a monotonous increasing behavior giving a more or less pronounced threshold effect (e.g. a sigmoidal shape). This has to be done looking for the best compromise between quality of fit and number of parameters (as small as possible). We tested several sigmoidal functions. For the fit of the Paris area at the scale of the

municipalities, we made use of a variant with a strict threshold:

$$\begin{cases} \Lambda < \Lambda_{ck},\ \Psi_k(\Lambda) = 0 & \text{(11a)} \\ \Lambda > \Lambda_{ck},\ \Psi_k(\Lambda) = \eta_k\ (1 - \exp -\gamma_k\ (\Lambda - \Lambda_{ck})) & \text{(11b)} \end{cases}$$

The fit being done with site-independent free parameters, this function thus contributes to three free parameters, Λ_c, η and γ.

Choice of the weights

The best results are obtained for two options. One is a power law decay with the distance:

$$W_{kj} = (1 + \mathrm{dist}(k,j)/d_0)^{-\delta} \tag{12}$$

where $\mathrm{dist}(k,j)$ is the distance between site k and site j (see below for its computation). The second option is the sum of an exponential decay and of a constant term

$$W_{kj} = \xi\ +\ (1-\xi)\exp\left(-\mathrm{dist}(k,j)/d_0\right) \tag{13}$$

In both cases we normalize the weights so that for every site k, $W_{kk} = 1$. Taking site-independent free parameters, both cases give two free parameters, d_0 and δ for the choice (12), d_0 and ξ for the choice (13).

Links to the original spatially continuous SIR model

In the case of the linear approximation, the metapopulation SIR model that we have introduced leads to the set of equations (10) of a type similar to the space-continuous non local (distributed contact) SIR model. With a view to describe the spreading of infections in spatially distributed populations, Kendall[29] introduced in 1957 this non-local version of the Kermack-McKendrick SIR model in the form of space-dependent integro-differential equations. Omitting the recovered population R, the system in the S, I variables reads:

$$\begin{cases} \dfrac{dI(x,t)}{dt} = -\omega\, I(x,t) + \beta\, S(x,t) \displaystyle\int K(x,y)\ I(y,t)\ dy & \text{(14a)} \\ \dfrac{dS(x,t)}{dt} = -\beta\, S(x,t) \displaystyle\int K(x,y)\ I(y,t)\ dy & \text{(14b)} \end{cases}$$

where $x \in \mathbb{R}^N$, with $N = 1, 2$, and here $I(x,t)$ and $S(x,t)$ are *densities* of immune and susceptible individuals. In the particular case of dimension $N = 1$, and the space is *homogeneous*, meaning here that $K(x,y)$ is of the form $K(x,y) = w(x-y)$, we know[30] that there exist traveling waves of any speed larger than or equal to some critical speed. Furthermore, this critical traveling wave speed also yields the asymptotic speed of spreading of the epidemic[37]. There have been many mathematical works on this system and on various extensions[38, 18]. Thus, at least in dimension $N = 1$ and for homogeneous space, this non-local system can generate traveling fronts for the density of susceptible individuals, hence the propagation of a 'spike' of infected individuals. Although no proof exists in dimension $N = 2$, numerical simulations show that the model can indeed generate waves[39, 40], as illustrated by the SI Videos 3 and 4, similar to the way the riot spread around Paris giving rise to the informal notion of a *riot wave*.

However, the model we introduce here is more general and differs from the Kendall model in certain aspects. Indeed, rather than continuous and homogeneous, the spatial structure is discrete with heterogeneous sites. Moreover, the set of equations here (14) corresponds to the linear approximation (10), whereas our general model involves a non-linear term. The understanding of *generalized* traveling waves and the speed of propagation in this general context are interesting open mathematical problems.

More work is needed to assess the mathematical properties of the specific family of non local contagion models introduced here, that is defined on a discrete network, with highly heterogeneous populations, and a non linear probability of becoming infected.

Date of the maximum

Figures 2c and 4b show how well the model accounts for the temporal unfolding of the riot activity, thanks to a comparison between model and data of the date when the riot activity peaks at each location. Given the noisy nature of the data, the empirical date of that maximum itself is not well defined. For each site, we estimated this date as the weighted average of the dates of the 3 greatest values, weighted by those values. We filled in missing data values by linear interpolation.

Non-free parameters: Reference populations

Choice of the reference population. As discussed in the main text, the riots started and developed in poor neighborhoods, a feature common to many urban riots[5]. Making use of French national statistics provided by INSEE (see below), we tested the use of various specific populations, considering cross-linked databases that involve age, sex, diploma or unemployment. We found the best log-likelihood when using the size of the population of males aged between 16 and 24 with no diploma, while not attending school (INSEE statistics of 2006). We thus calibrate the susceptible population in the model by assuming that, for each site, its size is proportional to the one of the corresponding reference population.

Populations at the scale of a département. When applying the model at the scale of départements, for each département the size of the reference population is computed as the sum of the sizes of the corresponding populations of all its municipalities that are under police authority.

Influence on the results. The choice of the reference population has a major influence on the results. We find that an improper choice cannot be compensated by the optimization of the free parameters. As an example, compare SI Fig. S2, for which the reference population is the total population, with Fig. 2a and 2b.

Source of national French statistics: INSEE, the French national institute carrying the national census (`http://www.insee.fr/`). For the period under consideration, in most cases, we used relevant data from 2006 when data from 2005 were not available.

Non-free parameters: distance between sites

The geographic data are taken from the collaborative project Open Street Map (`http://osm13.openstreetmap.fr/~cquest/openfla/export/`).

The distance $\mathrm{dist}(k, j)$ is taken as the one (in km) between the centroid of each site. In the case of the municipalities, the centroid is taken as the geographic centroid computed with QGIS[41]. In the case of départements, the centroid is computed as the weighted centroid (weighted by the size of the reference population) of all its municipalities that are under police authority.

Free parameters: numerical optimization

The data fit makes use of the maximum likelihood approach[42]. Let us call $X = \{x_{k,i}, k = 1 \ldots K, i = 1 \ldots 44\}$ the data, where each $x_{k,i} \in \mathbb{N}$ corresponds to the number of events for the site k at day i ($i = 1$ corresponding to October 26, 2005), and let θ denote the set of free parameters (e.g. $\theta = \{\omega, A, \zeta_0, d_0, \delta, \beta\}$ in the multi-sites linear case). Assuming conditional independence, we have:

$$p(X|\theta) = \prod_k \prod_i p(x_{k,i}|\theta) \tag{15}$$

Under the Poisson noise hypothesis, the $x_{k,i}$ are Poisson probabilistic realizations with mean $(\lambda_{k,i}(\theta) + \lambda_{bk})$:

$$p(x_{k,i}|\theta) = \frac{(\lambda_{k,i}(\theta) + \lambda_{bk})^{x_{k,i}}}{x_{k,i}!} \exp\big(- (\lambda_{k,i}(\theta) + \lambda_{bk})\big) \tag{16}$$

The log-likelihood, computed over all the sites under consideration and over the whole period (44 days long) for which we have data, thus writes:

$$\begin{aligned} \ell(\theta|X) &= \log p(X|\theta) \\ &= \sum_{k,i} \big(- \lambda_{k,i}(\theta) - \lambda_{bk} + x_{k,i} \log(\lambda_{k,i}(\theta) + \lambda_{bk})\big) \\ &\qquad + \sum_{k,i} \log x_{k,i}! \end{aligned} \tag{17}$$

Note that the last term in the right hand side does not depend on the free parameters and we can thus ignore it.

We performed the numerical maximization of the log-likelihood using the interior point algorithm[43] implemented in the MATLAB[44] `fmincon` function.

The method developed here allows one to explore the possibility of predicting the future time course of events based on the observation of the events up to some date. Preliminary results indicate that, once the activity has reached its peak in the Paris area, the prediction in time and space of the riot dynamics for the rest of France becomes quite accurate.

Results, details: Paris area, municipality scale

For the results illustrated by the figures in the paper, we give here the free parameters numerical values obtained from the maximum likelihood method in the case of the fit at the scale of municipalities in Île-de-France. Note that this optimization is computationally demanding: it requires to generate a large number of times (of order of tens of thousands) the full dynamics (44 days) with 2560 (2×1280) coupled equations. For the choice of the function Ψ, we tested the linear case and several nonlinear choices. Results are presented for the non linear case, the function Ψ being given by (11). We find: $\omega = 0.26, A = 5.5$; for the power law decrease of the weights, $d_0 = 8.\ 10^{-3} km,\ \delta = 0.67$; $\zeta_0 = 7.7/N_{max}$, where $N_{max} = 1174$ is the maximum size of the reference populations, the max being taken over all Île-de-France municipalities; for the parameters of the non linear function: $\eta = 0.63, \gamma = 1.27, \Lambda_c = 0.06$.

Results, details: All of France, département scale

We detail here the model options and the numerical results for the global model, considering each one of the départements of metropolitan France (except Corsica and Paris, hence 93 départements) as one homogeneous site. We have thus 186 (2×93) coupled equations with 6 to 9 free parameters, depending on the choice of the function Ψ.

Outliers. Looking at the results for different versions, we observe some systematic discrepancy between data and model for three départements: 93, where the predicted activity is too low, and 13 and 62 where it is too high. Actually, if one looks at the empirical maximum number of events as a function of the size of the reference population used for calibrating the susceptible population, these three départements show up as outliers: the riot intensity is significantly different from what one would expect from the size of the poor population. Outliers are here defined as falling outside the mean ± 3 standard deviations range when looking at the residuals of the linear regression.

The cases of 93 and 13 are not surprising. Département 93 is the one where the riots started, and has the highest concentration of deprived neighborhoods. Inhabitants are aware of this particularity and refer to their common fate by putting forward their belonging to the 'neuf-trois' (nine-three, instead of ninety three). Events in département 13 are mainly those that occurred in the city of Marseille. Despite a high level of criminality, and large poor neighborhoods, the inhabitants consider that being "Marseillais" comes before being French, so that people might have felt less concerned. The case of 62 (notably when compared to 59) remains a puzzle for us.

Parametrization and results. Making use of the linear choice for Ψ, and the weights given by (13), we then introduced one more free parameter for each one of these sites, β_{93}, β_{62} andβ_{13}, allowing for a different value of the susceptibility than the one taken for the rest of France. Optimization is thus done over the choice of 9 free parameters: $\omega, A, \zeta_0, d_0, \xi, \beta, \beta_{13}, \beta_{62}$ and β_{93}.

The numerical values of the free parameters obtained after optimization are as follows: $\omega = 0.41, A = 2.6$; $\zeta_0 = 190/N_{max}$, where here $N_{max} = 15632$ is the maximum size of the reference populations of all metropolitan départements; here the function Ψ is taken linear, $\beta = 2.\,10^{-3}$, except for three départements as explained above. For these three départements with a specific susceptibility, one finds about twice the common value for the département 93 (where riots started), $\beta_{93}/\beta \sim 1.95$, and about half for the département 62 and 13, $\beta_{62}/\beta \sim 0.47$ and $\beta_{13}/\beta \sim 0.42$.

For the weights chosen with an exponential decrease plus a constant global value, Eq. (13), $d_0 = 36$ km, $\xi = 0.06$. When using instead the power law decrease, Eq. (12), one finds that the fit is almost as good. The exponent value is found to be $\delta = 0.80$, which is similar to the value $\delta = 0.67$ found for the fit at the scale of municipalities (restricted to Île-de-France region). Yet, these exponent values are much smaller than the ones, between 1. and 2., found in the literature on social interactions as a function of geographical distance[26, 27, 28]. A small value of δ means a very slow decrease, a hint to the need of keeping a non zero value at very large distance. This can be seen as another indication that the alternative choice with a long range part, Eq. (13), is more relevant, meaning that both geographic proximity and long range interactions matter.

Acknowledgements

This work received support from: the program SYSCOMM of the French National Research Agency, the ANR (project DyXi, grant ANR-08-SYSC-0008); the CNRS interdisciplinary programs, PEPS Humain and PEPS MoMIS; the European Research Council advanced grant ERC ReaDi (European Union's 7th Framework Programme FB/2007-13/, ERC Grant Agreement no 321186 held by H. Berestycki). N.R. was supported by the NSF Grant DMS-1516778.

Author contributions statement

S.R. collected the raw data and provided the sociological expertise; M.-A.D. and M.B.G. built the database; L.B-G, J.-P.N, H.B. and N.R. were involved in the mathematical modeling; L.B-G performed the numerical analyses and simulations; L.B-G, J.-P.N, H.B. and S.R. wrote the paper with input from all the authors. All authors reviewed the manuscript.

References

[1] S. Roché, “The nature of rioting. comparative reflections based on the French case study.,” in *Transnational Criminology Manual* (M. Herzog-Evans, ed.), (Nijmegen), pp. 155–170, Wolf legal publishers, 2010.

[2] D. Waddington, F. Jobard, and M. King, *Rioting in the UK and France: A Comparative Analysis*. Cullompton, Devon: Willan Publishing, 2009.

[3] C. Cazelles, B. Morel, and S. Roché, “Les violences urbaines de l’automne 2005: événements, acteurs: dynamiques et interactions. Essai de synthèse,” *Centre d’analyse stratégique*, 2007.

[4] R. M. Raafat, N. Chater, and C. Frith, “Herding in humans,” *Trends in cognitive sciences*, vol. 13, no. 10, pp. 420–428, 2009.

[5] M. Gross, “Why do people riot?,” *Current Biology*, vol. 21, no. 18, pp. R673–R676, 2011.

[6] P. Baudains, A. Braithwaite, and S. D. Johnson, “Target choice during extreme events: a discrete spatial choice model of the 2011 London riots,” *Criminology*, vol. 51, no. 2, pp. 251–285, 2013.

[7] P. Baudains, S. D. Johnson, and A. M. Braithwaite, “Geographic patterns of diffusion in the 2011 London riots,” *Applied Geography*, vol. 45, no. 0, pp. 211 – 219, 2013.

[8] H. Berestycki., J. Nadal, and N. Rodriguez, “A model of riots dynamics: shocks, diffusion and thresholds,” *Networs and Heterogeneous Media (NHM)*, vol. 10, no. 3, pp. 443–475, 2015.

[9] D. Braha, “Global civil unrest: contagion, self-organization, and prediction,” *PloS one*, vol. 7, no. 10, p. e48596, 2012.

[10] S. L. Burbeck, W. J. Raine, and M. A. Stark, “The dynamics of riot growth: An epidemiological approach,” *Journal of Mathematical Sociology*, vol. 6, no. 1, pp. 1–22, 1978.

[11] T. P. Davies, H. M. Fry, A. G. Wilson, and S. R. Bishop, “A mathematical model of the London riots and their policing,” *Scientific reports*, vol. 3, 2013.

[12] M. Granovetter, “Threshold models of collective behavior,” *American Journal of Sociology*, vol. 83, no. 6, pp. 1360–1380, 1978.

[13] D. J. Myers, “The diffusion of collective violence: Infectiousness, susceptibility, and mass media networks,” *American Journal of Sociology*, vol. 106, no. 1, pp. 173–208, 2000.

[14] B. L. Pitcher, R. L. Hamblin, and J. L. Miller, “The diffusion of collective violence,” *American Sociological Review*, pp. 23–35, 1978.

[15] M. J. A. Stark, W. J. Raine, S. L. Burbeck, and K. K. Davison, “Some empirical patterns in a riot process,” *American Sociological Review*, pp. 865–876, 1974.

[16] M. Salgado, A. Mascareño, G. Ruz, and J.-P. Nadal, “Models of contagion: Towards a theory of crises propagation,” *submitted*, 2016.

[17] K. Dietz, “Epidemics and rumours: A survey,” *Journal of the Royal Statistical Society. Series A (General)*, pp. 505–528, 1967.

[18] L. Wang and B. C. Wood, “An epidemiological approach to model the viral propagation of memes,” *Applied Mathematical Modelling*, vol. 35, no. 11, pp. 5442–5447, 2011.

[19] M. Mazars, “Les violences urbaines de l’automne 2005 vues du palais de justice. Etude de cas. Les procédures judiciaires engagées au tribunal de grande instance de Bobigny.,” *Centre d’analyse stratégique*, 2007.

[20] O. Diekmann and H. Heesterbeek, *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*. New York: Wiley, 2000.

[21] W. O. Kermack and A. G. McKendrick, “A contribution to the mathematical theory of epidemics,” in *Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences*, vol. 115, pp. 700–721, The Royal Society, 1927.

[22] H. W. Hethcote, “The mathematics of infectious diseases,” *SIAM review*, vol. 42, no. 4, pp. 599–653, 2000.

[23] A. Baddeley, *Spatial Point Processes and their Applications*, pp. 1–75. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007.

[24] M. Barthélemy, “Spatial networks,” *Physics Reports*, vol. 499, no. 1, pp. 1–101, 2011.

[25] T. C. Schelling, “Hockey helmets, concealed weapons, and daylight saving: A study of binary choices with externalities,” *The Journal of Conflict Resolution*, vol. XVII (3), 1973.

[26] J. Goldenberg and M. Levy, “Distance is not dead: Social interaction and geographical distance in the internet era,” *arXiv preprint arXiv:0906.3202*, 2009.

[27] R. Lambiotte, V. D. Blondel, C. de Kerchove, E. Huens, C. Prieur, Z. Smoreda, and P. Van Dooren, “Geographical dispersal of mobile communication networks,” *Physica A: Statistical Mechanics and its Applications*, vol. 387, no. 21, pp. 5317–5325, 2008.

[28] D. Liben-Nowell, J. Novak, R. Kumar, P. Raghavan, and A. Tomkins, “Geographic routing in social networks,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 33, pp. 11623–11628, 2005.

[29] D. G. Kendall, “Discussion of “Measles periodicity and community size” by M.S. Bartlett,” *Journal of the Royal Statistical Society Series A*, vol. 120, pp. 64–67, 1957.

[30] D. G. Kendall, “Mathematical models of the spread of infection,” *Mathematics and computer science in biology and medicine*, vol. 213, 1965.

[31] M. Emirbayer and J. Goodwin, "Network analysis, culture, and the problem of agency," *American journal of sociology*, pp. 1411–1454, 1994.

[32] R. V. Gould, "Why do networks matter? Rationalist and structuralist interpretations," in *Social movements and networks: Relational approaches to collective action* (M. Diani and D. McAdam, eds.), pp. 233–57, Oxford: Oxford University Press, 2003.

[33] S. Walgrave and R. Wouters, "The missing link in the diffusion of protest: Asking others," *American Journal of Sociology*, vol. 119, no. 6, pp. 1670–1709, 2014.

[34] G. L. Carter, "Local police force size and the severity of the 1960s black rioting," *Journal of Conflict Resolution*, vol. 31, no. 4, pp. 601–614, 1987.

[35] J. Bohstedt and D. E. Williams, "The diffusion of riots: the patterns of 1766, 1795, and 1801 in devonshire," *The Journal of Interdisciplinary History*, vol. 19, no. 1, pp. 1–24, 1988.

[36] D. C. Hoaglin, "A Poissonness plot," *The American Statistician*, vol. 34, no. 3, pp. 146–149, 1980.

[37] D. Aronson, "The asymptotic speed of propagation of a simple epidemic," in *Nonlinear diffusion*, vol. 14, pp. 1–23, Pitman London, 1977.

[38] S. Ruan, "Spatial-temporal dynamics in nonlocal epidemiological models," in *Mathematics for life science and medicine*, pp. 97–122, Springer, 2007.

[39] N. T. Bailey, "The simulation of stochastic epidemics in two dimensions," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 4, pp. 237–257, University of California Press Berkeley and Los Angeles, 1967.

[40] M. Rodríguez-Meza, "Spatial and temporal dynamics of infected populations: the Mexican epidemic," *Revista Mexicana de Física*, vol. 58, no. 1, pp. 58–62, 2012.

[41] QGIS Development Team, *QGIS Geographic Information System*. Open Source Geospatial Foundation, 2009.

[42] I. J. Myung, "Tutorial on maximum likelihood estimation," *Journal of mathematical Psychology*, vol. 47, no. 1, pp. 90–100, 2003.

[43] R. H. Byrd, M. E. Hribar, and J. Nocedal, "An interior point algorithm for large-scale nonlinear programming," *SIAM Journal on Optimization*, vol. 9, no. 4, pp. 877–900, 1999.

[44] MATLAB, *version 9.0.0 (R2016a)*. The MathWorks Inc., Natick, Massachusetts, 2016.

Supplementary Information

SI Figures

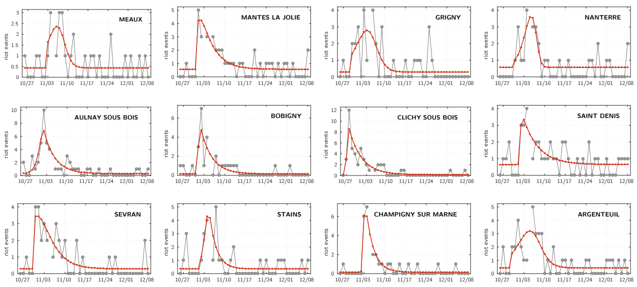


Figure S1: The 2005 French riots: data and single site fits for the 12 most active Île-de-France municipalities. Dots: number of events. Continuous curve: fit with the single site SIR model.

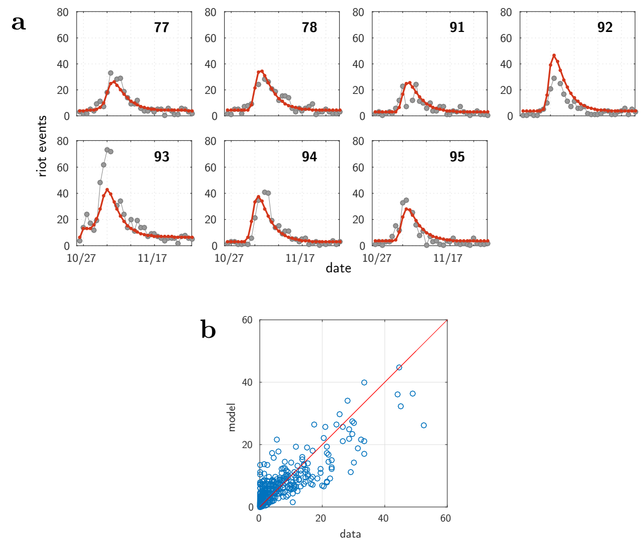


Figure S2: Control numerical experiment: fit with an inadequate reference population. Here the reference population is taken as the total population. (a) Île-de-France municipalities (aggregated by départements): data (dots), model (continuous curve) making use of the inadequate reference population. (b) Total number of events, model vs. data. Each dot represents one municipality. These results should be compared with the one on Fig. 2 of the main text – apart from the choice of a different reference population, the models options (choice of Ψ and of the weights) are the same.

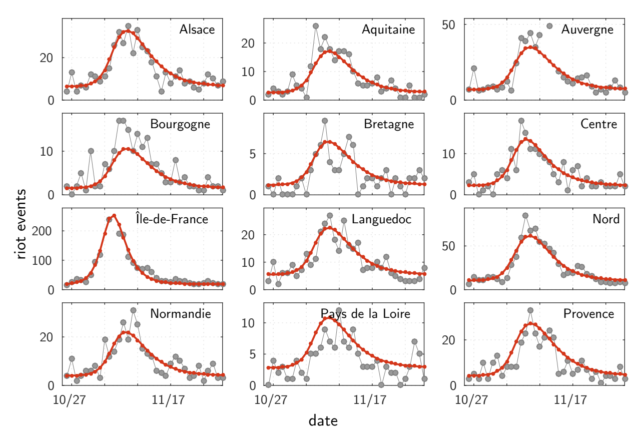
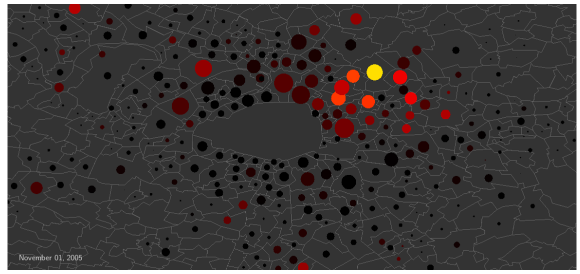


Figure S3: Results: all of France, spatial SIR model, calibration at the scale of the départements. Results aggregated by 'regions' (as of 2016), taking into account all the départements. See main text, section *Fitting the data: the wave across the whole country.*

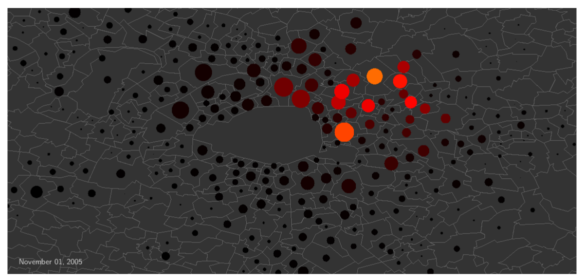
SI Videos

• SI Video 1. **Riot propagation around Paris: smoothed data.**
This video shows the riot propagation around Paris. The map shows the municipality boundaries, with Paris at the center. For each municipality for which data is available, a circle is drawn with an area proportional to the estimate of the size of the susceptible population (see main text, section Methods). Instead of making use of the raw data, for each municipality we replaced each day value by the one given by the fit with the single site epidemic model considered here as a tool for smoothing the data. The color represents the intensity of the rioting activity: the warmer the color, the higher the activity. The pace of the video corresponds to three days per second. In order to improve the fluidity of the video, we increased the number of frames per second by interpolating each day with 7 new frames, whose values are computed thanks to a piece-wise cubic interpolation of the original ones. The resulting frame rate is then 24 frames per second.
The movie is encoded with the open standard H.264.



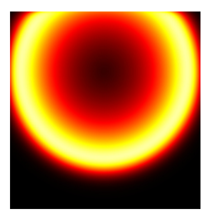
Still image from SI Video 1

• SI Video 2. **Riot propagation around Paris: model with non-local contagion.**
This video shows the riot activity as predicted by the data-driven global epidemic-like model. Same technical details as for the SI Video 1.
The movie is encoded with the open standard H.264.

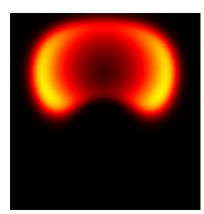


Still image from SI Video 2

• SI Video 3. **Spatial SIR: wave propagation in a homogeneous medium.**
We illustrate the formal continuous spatial SIR model with a video showing the propagation of a wave. The underlying medium is characterized by a uniform density of susceptible individuals. The weights $w(x-y)$ in the interaction term are given by a decreasing exponential function of the Euclidean distance $||x-y||$.
The movie is encoded with the open standard H.264.


Still image from SI Video 3

• SI Video 4. **Spatial SIR: wave propagation in a non-homogeneous medium.**
Same as SI Video 3, but with a heterogeneous density of susceptible individuals, characterized by (1) a decrease of the density towards 0 near the boundary of the image, so that the wave dies before leaving the frame, (2) a hole at the center, which is then bypassed by the wave, (3) a concentration of susceptible individuals that globally decreases on the y-axis, so that the wave dies while going downward.
The movie is encoded with the open standard H.264.


Still image from SI Video 4