SOCIAL HETEROGENEITY AND THE COVID-19 LOCKDOWN IN A MULTI-GROUP SEIR MODEL

JEAN DOLBEAULT

CEREMADE (CNRS UMR n° 7534), PSL university, Université Paris-Dauphine, Place de Lattre de Tassigny, 75775 Paris 16, France

GABRIEL TURINICI

CEREMADE (CNRS UMR n° 7534), PSL university, Université Paris-Dauphine, Place de Lattre de Tassigny, 75775 Paris 16, France

Abstract. The goal of the lockdown is to mitigate and if possible prevent the spread of an epidemic. It consists in reducing social interactions. This is taken into account by the introduction of a factor of reduction of social interactions q, and by decreasing the transmission coefficient of the disease accordingly. Evaluating q is a difficult question and one can ask if it makes sense to compute an average coefficient q for a given population, in order to make predictions on the basic reproduction rate $\mathcal{R}_0$, the dynamics of the epidemic or the fraction of the population that will have been infected by the end of the epidemic. On a very simple example, we show that the computation of $\mathcal{R}_0$ in a heterogeneous population is not reduced to the computation of an average q but rather to the direct computation of an average coefficient $\mathcal{R}_0$. Even more interesting is the fact that, in a range of data compatible with the Covid-19 outbreak, the size of the epidemic is deeply modified by social heterogeneity, as is the height of the epidemic peak, while the date at which it is reached mainly depends on the average $\mathcal{R}_0$ coefficient. This short note is an application and a summary of the technical results that can be found in [4].

E-mail addresses: dolbeaul@ceremade.dauphine.fr, Gabriel.Turinici@dauphine.fr.

2020 *Mathematics Subject Classification.* Primary: 92C60; Secondary: 92D30, 34C60.

Key words and phrases. Epidemic models; Disease control; Heterogeneous populations; Basic reproduction ratio; Equilibrium solutions .

We consider a compartmental model based on the SEIR equations (for *Susceptible*, *Exposed*, *Infected*, *Recovered*) with n categories of susceptible individuals

$$\frac{dS_k}{dt} = -\,\beta\, S_k\, \frac{I}{N}\,, \quad \frac{dE}{dt} = \sum_{k=1}^{n} \beta_k\, S_k\, \frac{I}{N} - \,\alpha\, E\,, \quad \frac{dI}{dt} = \alpha\, E - \,\gamma\, I\,, \quad \frac{dR}{dt} = \gamma\, I\,. \tag{1}$$

The *average incubation period* is $1/\alpha$, the parameter β_k is the product of the average number of contacts per person and per unit time by the probability of disease transmission in a contact between a susceptible individual in the group k and any infectious individual, γ is a *transition rate* so that $1/\gamma$ measures the duration of the infection of an individual (or actually how long he is infectious and able to contaminate other people before being isolated), and N is the total population size.

An individual in the group k is characterized by his *transmission rate* β_k. In simple models of lockdown, a single group is considered ($n = 1$) and it has been proposed for instance in [5] to introduce a *factor of reduction of social interactions* q so that the effect of the lockdown is to reduce the transmission rate β before lockdown to β/q after lockdown. Our goal is to study what happens if there are, after lockdown, $n \geq 2$ groups with different *factors of reduction of social interactions* q_k, so that

$$\beta_k = \frac{\beta}{q_k}\,.$$

We shall speak of *social groups* and *social heterogeneity* because each group has its own transmission rate β_k. This rate does not interfere with the dynamics of the disease once the corresponding individual is infected. In a period of lockdown, the reduction factor is not the same for a health worker, a supermarket cashier or an employee working from home by internet. If we focus on β_k, we can also take into account other characteristics, like age groups, which are not directly disease related, but apparently play a role in the risk of becoming infected.

Reduction of social interactions and basic reproduction ratio. Let us consider the probability distribution among the groups such that

$$S_k = p_k\, S \quad \text{where} \quad S = \sum_{k=1}^{n} S_k\,.$$

In a *disease free equilibrium* corresponding to $S = N$, it would be very natural consider an average *factor of reduction of social interactions*

$$q = \sum_{k=1}^{n} p_k\, q_k$$

and this is actually what is done when a single compartment of susceptible individuals is considered. In that case, the *basic reproduction ratio* computed for instance by the *next generation matrix method* (see [3] and references therein) is

$$\frac{\beta}{q\,\gamma}\,.$$

However, if we apply the next generation matrix method to (1), we find that the *basic reproduction ratio* is

$$\mathcal{R}_0 = \sum_{k=1}^{n} p_k\, \frac{\beta_k}{\gamma}\,. \tag{2}$$

Although this is a relatively elementary property, there are already important consequences, for instance in the case $\frac{\beta}{q\,\gamma} < 1 < \mathcal{R}_0$. While a model with a single compartment and an average factor of reduction of social interactions predicts the extinction of the epidemic, a small group with high transmission factor β_k, even if numerically not very large (that is, with p_k small), can spread the disease and trigger an outbreak. This is one of the possible explanations, probably among many other, why the statistical data were suggesting a pattern corresponding to a basic reproduction ratio larger than 1 in the initial stage of the lockdown rather than an exponential decay of the number of cases.

Limitations and choice of the numerical parameters. System (1) is an extremely simplified model, in which *Undetected* or asymptomatic cases are not taken into account, although this seems an important issue in the Covid-19 pandemic. It is also a model for short term (say of the order of three months) so that the evolution of the structure as well as natural birth and death rates are not taken into account. If Covid-19 becomes endemic or last a year or more, it will not be possible to keep ignoring such issues anymore. Feedback mechanisms and changes in the values of the parameters due to the evolution of the disease or the changes in social habits are certainly going to play a role. More important is the fact that the parameters are fitted on the basis of official data subject to revision and only for the very early stage of the disease.

System (1) is homogeneous so that we can simply consider the *fractions*

$$\mathsf{s}_k := \frac{S_k}{N}\,, \quad \mathsf{s} := \frac{S}{N}\,, \quad \mathsf{e} := \frac{E}{N}\,, \quad \mathsf{i} := \frac{I}{N}\,, \quad \mathsf{r} := \frac{R}{N}$$

of the *Susceptible*, *Exposed*, *Infected* and *Recovered* individuals among the whole population. Numerical values are taken from [2]: for the initial data, we consider a perturbation of the *disease free equilibrium* given by

$$\mathsf{s}(0) = 0.99981\,, \quad \mathsf{e}(0) = 8.81 \times 10^{-5}\,, \tag{3}$$

$$\mathsf{i}(0) = 1.88 \times 10^{-5}\,, \quad \mathsf{r}(0) = 8.04 \times 10^{-5}\,, \tag{4}$$

based on data available on March 15, 2020 in France from [6], and choose in (1) the parameters

$$\beta = 2.33\,, \quad \alpha = 0.25\,, \quad \gamma = 1\,, \tag{5}$$

which gives us a basic reproduction ratio (corresponding to a single group) of 2.33. The reader interested in further details and comparisons is invited to refer to [4], where the choice of the numerical parameters is discussed in greater details. This set of initial data and parameters is however not a key issue for our discussion. Our contribution is focused on understanding the theoretical implications of social heterogeneity and remains valid for other sets of numerical choices. Numerical choices (3)-(5) are given only for an illustrative purpose.

In [2], N. Bacaër is able to fit the curve of the number of infected individuals with a single group and $q_1 = 1.7$, at the beginning of the lockdown. In [4], it is shown that with two groups, $p_1 = 98\%$, $p_2 = 2\%$, $q_1 = 2.35$ and $q_2 = 0.117$, we can explain the reproduction ratio of 1.37 found in [2]. In order to fix ideas, we shall take

$$q_1 = 2.4$$

which corresponds to a reduction of the transmission rate of $1/q_1 \approx 60\%$ in the group $k = 1$. We emphasize that we have no empirical basis to determine this

parameter. The idea is to pick a value larger than 2.33 so that, with a single group, the epidemic would rapidly extinguish without spreading in the whole population. However, a second group with a moderate or high transmission coefficient β_2 is enough to produce an outbreak corresponding to a $\mathcal{R}_0 > 1$ and we study how various features of the epidemic curve depend on β_2. See Fig. (1).

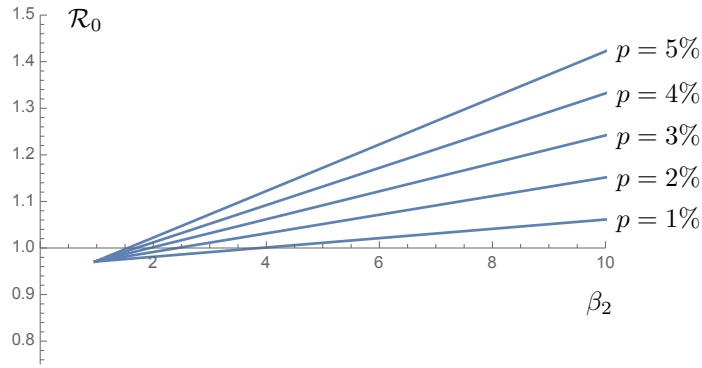

FIGURE 1. Numerical values of $\mathcal{R}_0$: we take p small but assume that the individuals in the group $k = 2$ may have a moderate or high transmission coefficient β_2.

From now on, we shall assume in all our numerical results that $n = 2$ (two groups) and vary $p = p_2$ (so that $p_1 = 1 - p$) at $t = 0$, which is the time of the beginning of the lockdown. Consistently with the idea that the initial stage of the outbreak is independent of the groups, we take as initial data $\mathsf{s}_1(0) = (1-p)\,\mathsf{s}(0)$ and $\mathsf{s}_2(0) = p\,\mathsf{s}(0)$. We keep p small and take it in the range 1 to 5% in our examples. It is not difficult to understand that the destabilization due to a small fraction p of the population (the $k = 2$ group) in a large group (the $k = 1$ group) in which the epidemic would be under control if it were isolated, is all the stronger when the basic reproduction ratio (inside the $k = 1$ group, considered as isolated) is close to 1: this is the reason why we arbitrarily choose $q_1 = 2.4$. The case of q_2 small, corresponding to individuals in the group $k = 2$ with high transmission rates, is of particular interest.

The epidemic size. The *epidemic size* ζ is defined as the fraction of individuals that are affected by the epidemic for large times, here $\mathsf{s}(0) - \mathsf{s}^\star$, where $\mathsf{s}^\star$ is the fraction which is not infected at the end of the epidemic. Under the condition that $\mathcal{R}_0 > 1$ and $1 - \mathsf{s}(0)$ is small, we find that $\mathsf{r}^\star = 1 - \mathsf{s}^\star$ is of the order of the solution r of

$$\sum_{k=1}^{n} e^{-\frac{\beta_k\, r}{\gamma}} + r = 1\,.$$

It is proved in [4] that ζ is decreased when n is increased form $n = 1$ to $n \geq 2$. See Figs. 2 and 3.

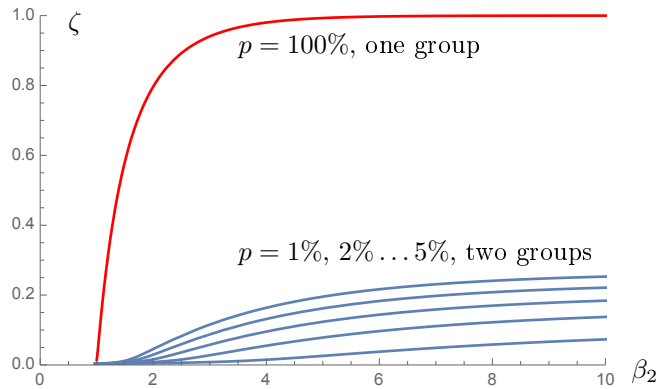

FIGURE 2. Epidemic size ζ as a function of β_2 for $p = 0.01$, $0.02, \ldots 0.05$ (blue; $n = 2$) and $p = 1$ (red, $n = 1$). This red curve corresponds to an outbreak with a basic reproduction ratio β_2/γ. This is however not so instructive as the $k = 2$ group is the one with highest transmission rate.

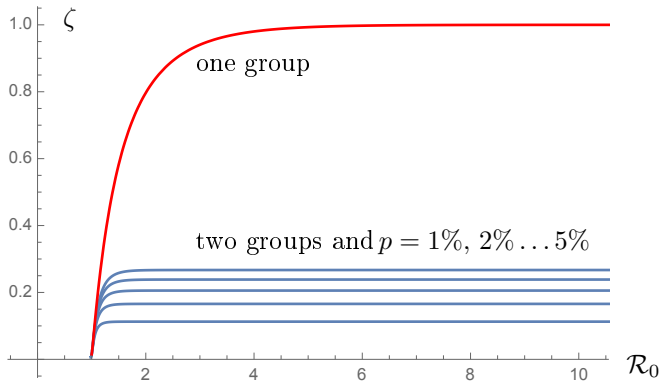

FIGURE 3. Epidemic size ζ as a function of $\mathcal{R}_0$ for $p = 0.01$, $0.02, \ldots 0.05$ (blue) and $p = 1$ (red), which corresponds to the case of a single group. With the same $\mathcal{R}_0$, we see that ζ is drastically decreased. For instance, $\mathcal{R}_0 = 1.37$ is achieved in a single group with $q \approx 1.70$ as in [2], which gives an epidemic size $\zeta \approx 49\%$. The same value $\mathcal{R}_0 = 1.37$, which is obtained in two groups with $p = 1\%$, $q_1 = 2.4$ (and $q_2 \approx 0.06$), gives an epidemic size $\zeta \approx 12\%$.

The epidemic peak. Not only the asymptotics of the epidemic are changed when replacing the model with $n = 1$ by a model with $n = 2$, but also the dynamical behavior of the solutions. The height of the epidemic peak (maximum of infected $\mathsf{e} + \mathsf{i}$ as function of time) is decreased when n is increased form $n = 1$ to $n \geq 2$: see [4] for a proof. The height and date of the epidemic peak strongly depend on p. See Figs. 4, 5 and 6.

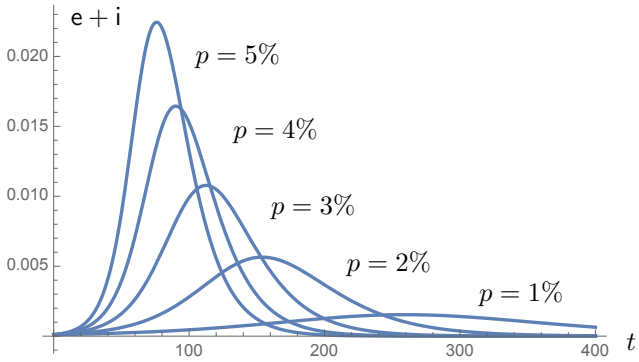

FIGURE 4. Fraction of exposed and infected individuals $\mathsf{e} + \mathsf{i}$ as a function of time (in days) for $p = 0.01$, $0.02, \ldots 0.05$, with $\beta_2 = 11.65$ corresponding to $q_2 = 0.2$. Our convention is that the epidemic peak is the maximum of $t \mapsto \mathsf{e}(t) + \mathsf{i}(t)$.

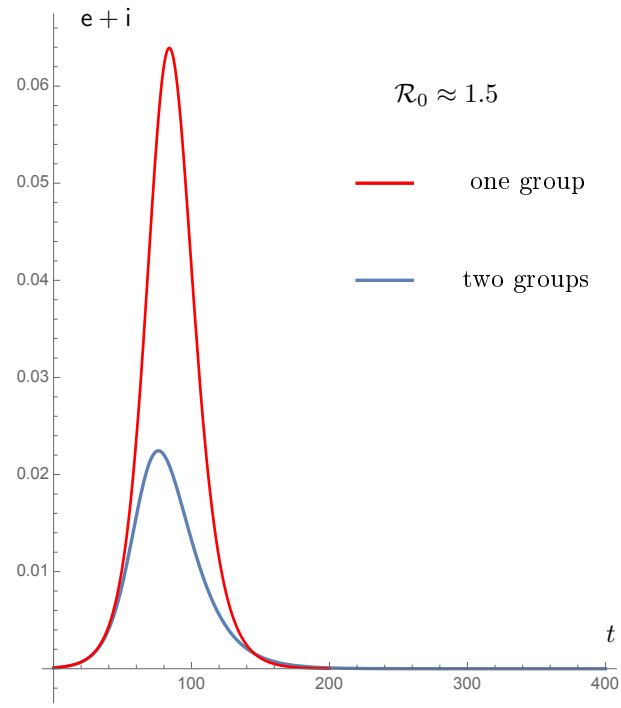

FIGURE 5. A striking effect of the heterogeneity is the flattening of the curve. For the same value of $\mathcal{R}_0 \approx 1.5$, the curve $t \mapsto \mathsf{e}(t) + \mathsf{i}(t)$ is represented in red for a single group (this corresponds here to a factor of reduction of social interactions $q \approx 1.55$) and in blue in the case of two groups (with parameters $q_1 = 2.4$, $q_2 = 0.2$ and $p = 0.05$). Note that the blue curve is the same as in Fig. 4 with $p = 5\%$.

Now, let us describe how the epidemic peak depends on the parameters in our two groups example. For a given value of β_2, larger values of p mean a larger $\mathcal{R}_0$, with a linear dependence given by (2), a larger epidemic size (see Fig. 2), but also a larger epidemic peak (see Fig. 6).

Let us denote by $t_\star$ the date of the peak and let

$$\mathsf{m}_\star := \mathsf{e}(t_\star) + \mathsf{i}(t_\star) = \max_{t \geq 0} \big(\mathsf{e}(t) + \mathsf{i}(t)\big) .$$

We observe that $p \mapsto \mathsf{m}_\star(p)$ is increasing with p and $p \mapsto t_\star(p)$ is non-increasing with p for sufficiently large values of p.

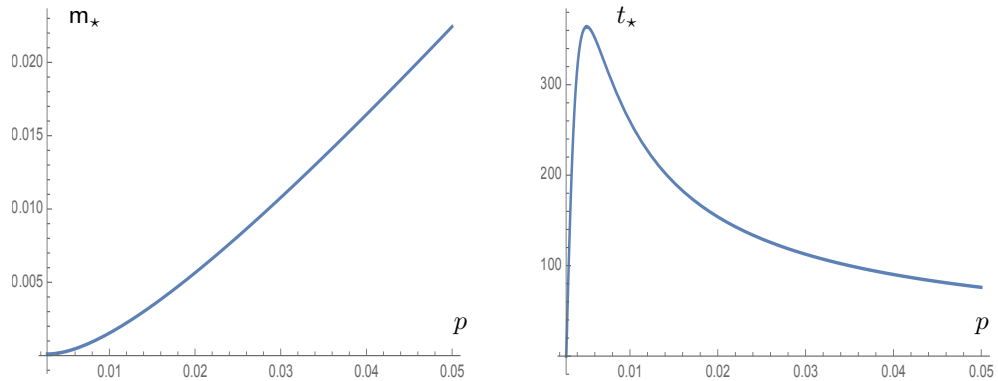

FIGURE 6. Size (left) and date (right) of the epidemic peak as function of p in the range 0.27% to 5%. With our choice of the parameters ($q_1 = 2.4$ and $q_2 = 0.2$), there is no outbreak if $p \leq 0.27\%$, because $\mathcal{R}_0 \leq 1$ in that case.

It is also interesting to make comparisons with the same value of $\mathcal{R}_0$. In our model case for $n = 2$, by inverting (2), we can for instance choose q_2 as a function of p to achieve a fixed value of $\mathcal{R}_0$ and get

$$q_2 = \frac{\beta\, p\, q_1}{\gamma\, q_1\, \mathcal{R}_0 - (1-\beta)\, p}\,. \tag{6}$$

See Fig. 7 for some qualitative results on the dependence of $\mathsf{m}_\star$ and $t_\star$ in $\mathcal{R}_0$, for various values of p.

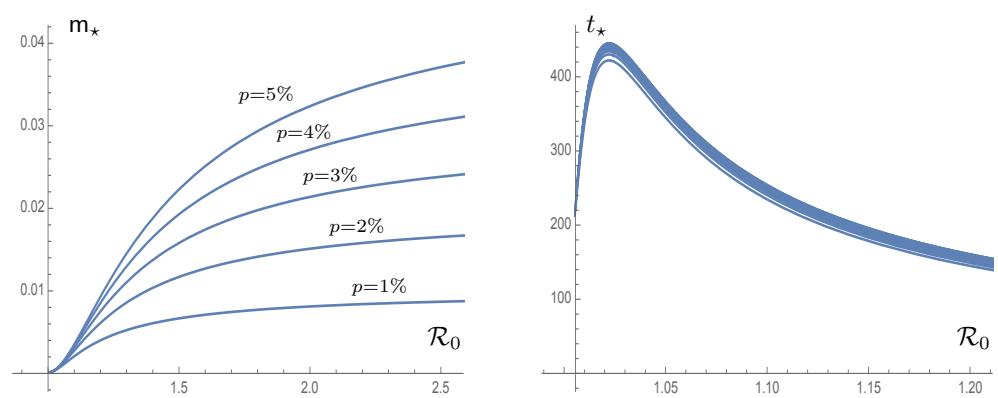

FIGURE 7. Left: size of the epidemic peak as a function of $\mathcal{R}_0$. Although not linear, it is a clearly increasing function of p in the outbreak regime. Right: date of the epidemic peak as a function of $\mathcal{R}_0$. It is an empirical but remarkable fact that, for our set of data, the date almost does not depend on p. Plots correspond to $p = 0.01,\, 0.02, \ldots 0.05$.

The $\mathcal{R}_0$ is not the message. As seen in Fig. 7, $\mathcal{R}_0$ does not retain all relevant information and for a given q_1 and $\mathcal{R}_0$, one may still vary either p or q_2 and eliminate the other one using (2). Here we shall rely on (6) for a few more plots which summarize our findings.

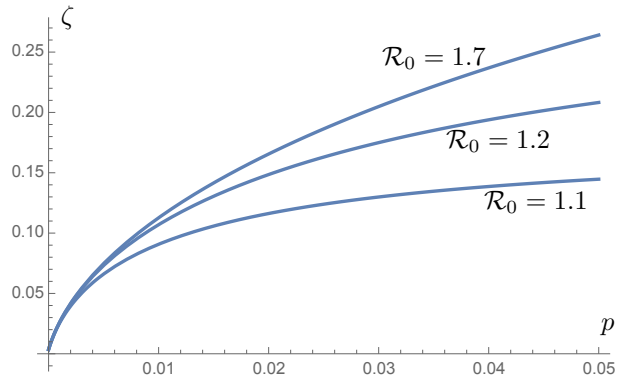

FIGURE 8. The epidemic size ζ as a function of p ranging from 0 to 0.05, where q_2 is adjusted according to (6) so that $\mathcal{R}_0$ is given and equal to 1.1, 1.2 and 1.7.

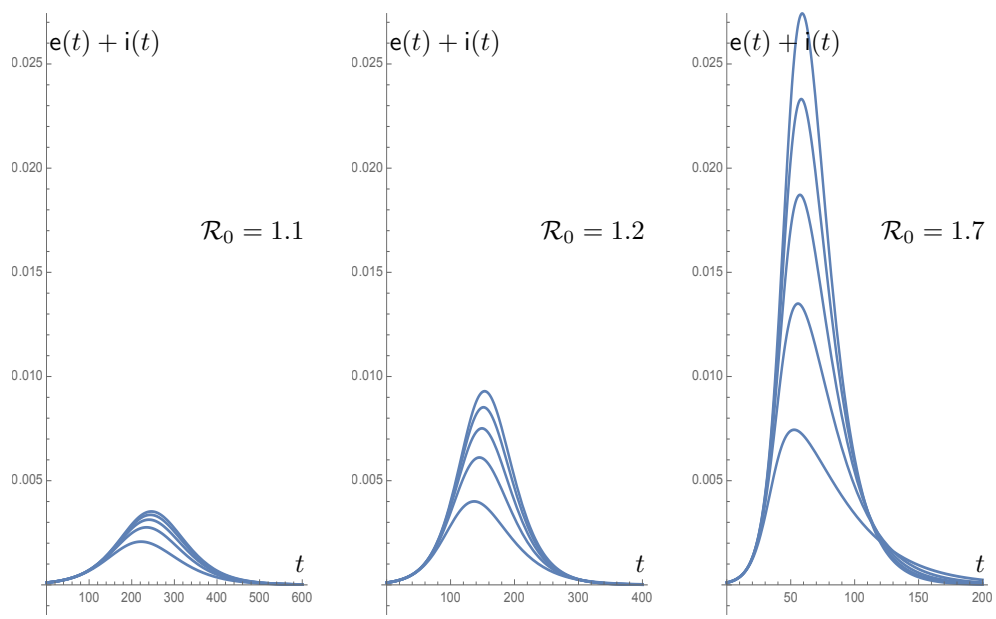

FIGURE 9. The epidemic curve and the peak strongly depends on $\mathcal{R}_0$. Notice that the scale on the vertical axis is the same in the three cases corresponding to $\mathcal{R}_0 = 1.1$, 1.2 and 1.7, while the time scale is respectively $(0, 200)$, $(0, 400)$, and $(0, 600)$. On each plot, the curves correspond to $p = 0.01$, $0.02, \ldots 0.05$.

Discussion. The goal of a lockdown is to reduce the basic reproduction ratio of the epidemic to a value less than 1 and drive the disease to extinction, or at least decrease it in such a way that "the curve is flattened". The efficiency of the lockdown is achieved by a reduction of social interactions, which is measured by a factor q. Here we question whether such an average factor makes sense or not. Beyond the difficult issue of giving realistic values to q, we study the impact of social heterogeneities when the population is divided into groups for which q takes different values.

We observe that the basic reproduction ratio $\mathcal{R}_0$ is computed as an average of the basic reproduction ratios for each group, or on the basis of an average transmission rate, and not as a global ratio based on an average q factor. A small group with a high transmission rate eventually triggers an outbreak even if the basic reproduction ratio $\mathcal{R}_0$ of the majority would be below 1.

A second message is the following: the equilibrium in presence of heterogenous groups is not the same as when considering a single group with **the same** (averaged) basic reproduction ratio. The dynamics of the outbreak, for instance the height of the epidemic peak, is also changed. A model with only one group and fitting the observed data in the initial phase of the outbreak will be more pessimistic concerning the epidemic outcomes than a heterogeneous model. This is even more true after lockdown when social distancing measures have been enforced, the lockdown being by its nature a creator of heterogeneity. In terms of public health, this also underlines the importance of targeting prevention measures on individuals with a high level of social interactions. Further details and mathematical proofs can be found in [4].

Acknowledgment: This work has been partially supported by the Project EFI (ANR-17-CE40-0030) of the French National Research Agency (ANR). The authors thank the MODCOV19 platform [1] for support.

© 2020 by the authors. This paper may be reproduced, in its entirety, for non-commercial purposes.

References

[1] *Modcov19*. https://modcov19.math.cnrs.fr.

[2] N. Bacaër, *Un modèle mathématique des débuts de l'épidémie de coronavirus en France*, Mathematical Modelling of Natural Phenomena, 15 (2020), p. 29.

[3] O. Diekmann, J. A. P. Heesterbeek, and M. G. Roberts, *The construction of next-generation matrices for compartmental epidemic models*, Journal of The Royal Society Interface, 7 (2009), pp. 873–885.

[4] J. Dolbeault and G. Turinici, *Heterogeneous social interactions and the COVID-19 lockdown outcome in a multi-group SEIR model*. Preprint hal-02559938, Apr. 2020.

[5] H. Salje, C. Tran Kiem, N. Lefrancq, N. Courtejoie, P. Bosetti, J. Paireau, A. Andronico, N. Hoze, J. Richet, C.-L. Dubost, Y. Le Strat, J. Lessler, D. L. Bruhl, A. Fontanet, L. Opatowski, P.-Y. Boëlle, and S. Cauchemez, *Estimating the burden of SARS-CoV-2 in France*. Preprint https://hal-pasteur.archives-ouvertes.fr/pasteur-02548181, Apr. 2020.

[6] Santé Publique France, *French national public health agency*. https://dashboard.covid19.data.gouv.fr, 2020.