MathematicS In Action
Version de travail

A simple method to integrate social factors into SIR contact matrix: Insights from COVID-19

Vincent Cardon
Gauthier Delvoye
Hervé Le Meur
Youcef Mammeri
Sylvain Thine

[*] Centre Universitaire de Recherches sur l'Action Publique et le Politique, CNRS UMR 7319, Université de Picardie Jules Verne, 80069 Amiens, France
Email address: vincent.cardon@u-picardie.fr
[†] Laboratoire Amiénois de Mathématique Fondamentale et Appliquée, CNRS UMR 7352, Université de Picardie Jules Verne, 80069 Amiens, France
Email address: gauthier.delvoye@u-picardie.fr
[‡] Laboratoire Amiénois de Mathématique Fondamentale et Appliquée, CNRS UMR 7352, Université de Picardie Jules Verne, 80069 Amiens, France
Email address: herve.le.meur@u-picardie.fr
[§] Université Jean Monnet, Institut Camille Jordan UMR5208, CNRS, Ecole Centrale de Lyon, INSA Lyon, Universite Claude Bernard Lyon 1, 42023 Saint-Etienne, France
Email address: youcef.mammeri@math.cnrs.fr
[¶] Centre Universitaire de Recherches sur l'Action Publique et le Politique, CNRS UMR 7319, Université de Picardie Jules Verne, 80069 Amiens, France
Email address: sylvain.thine@u-picardie.fr.

Abstract. The COVID-19 pandemic has underscored the importance of social structure in shaping epidemic dynamics, beyond age alone. While recent models increasingly incorporate social heterogeneity, many rely on complex frameworks that limit their practical use. In this paper, we propose a simple and tractable method to integrate alternative social representations into compartmental epidemic models using standard age-stratified contact matrices. Our approach relies on an algebraic transformation based on a transition matrix, allowing social dimensions such as occupation, education, or health status to be incorporated without requiring additional data on contact mixing patterns. This formulation preserves the classical contact matrix structure and is straightforward to implement. We illustrate the method through numerical simulations accounting for age distribution, telework feasibility, school closures, and health status. The results demonstrate how targeted interventions can be evaluated within a unified and accessible modeling framework, providing a practical tool for incorporating social heterogeneity into epidemic modeling.

1. Introduction

The COVID-19 pandemic has starkly revealed the profound influence of social class on health outcomes, bringing long-standing social inequalities into sharp focus [38, 29]. Beyond its unprecedented pressure on healthcare systems and economies, the pandemic exposed systematic disparities in exposure, vulnerability, and mortality across socio-economic groups. Individuals'

The authors were funded by the Agence National de la Recherche and Région Hauts-de-France, projet Space-covid ANR Résilience.
Keywords: Contact matrix, Social class, Age, COVID-19, SEI_sI_aR model .
Math. classification: 92D30, 37N25, 35K51, 35Q92.

capacity to comply with public health measures such as physical distancing, teleworking, or isolation varied strongly by social position, reflecting differences in occupation, housing conditions, income security, and access to healthcare. As a result, marginalized and working-class populations experienced disproportionately higher infection rates, worse health outcomes, and more severe socio-economic consequences. These observations underscore the necessity of explicitly accounting for social structure when analyzing epidemic dynamics and designing effective public health interventions.

A growing body of empirical and modeling research has documented the heterogeneous social impacts of the COVID-19 crisis. Extended school closures disproportionately affected children from economically disadvantaged backgrounds, highlighting the role of educational institutions as essential social infrastructure [13]. University students from working-class backgrounds faced additional barriers during the transition to remote learning, including digital exclusion and cultural mismatches [31]. Modeling studies have explored school reopening strategies, showing that while reopening may increase transmission, well-designed mitigation protocols can control epidemic spread [10, 15, 21, 12]. Social class has also been shown to shape compliance with public health measures, vaccine hesitancy, and risk perception [7, 5, 18], as well as gendered and class-specific behavioral responses during the pandemic [27]. More broadly, working-class populations faced heightened exposure and compounding inequalities, contributing to increased socio-economic gradients in COVID-19 mortality and healthcare access [33, 26, 2, 19, 16]. Such a discrepancy may stem from differences in working conditions and the need for in-person collaboration, whereas other roles may involve working independently in front of a screen with little to no interaction with colleagues [6].

From a modeling perspective, recent epidemic models increasingly incorporate heterogeneous mixing patterns stratified by age and socio-economic variables, acknowledging that contact structures are shaped by multiple, intersecting social dimensions [3, 39, 36]. In this context, Manna et al. [28] introduced the framework of generalized contact matrices, which represent contact patterns across an arbitrary number of social dimensions arranged in a tensor-like structure. This approach offers a powerful and flexible way to capture complex social heterogeneity but comes at the cost of increased mathematical and computational complexity, which may limit its accessibility and practical use in applied settings.

The present study proposes a simpler and more tractable approach that can be viewed as a reduced case of the generalized framework proposed by Manna et al. [28]. Specifically, we retain the familiar contact matrix structure commonly used in age-structured SIR models, while allowing social stratification to be incorporated through an algebraic transformation. Changes in the partition of social classes are achieved by defining an appropriate transition matrix, constructed from the proportion of individuals belonging to each social class within a given age group—formally, the conditional probability that an individual belongs to a given class given their age group. This procedure preserves the interpretability and numerical convenience of classical contact matrices, while enabling the integration of additional social dimensions in a way that is both easy to implement and computationally efficient.

The objectives of this paper are therefore twofold. First, we explore alternative matrices of social representation beyond purely age-based stratifications. Second, we formalize the process by which one social representation can be transformed into another within compartmental epidemic models. Once the suitable social representation is established (Section 2), numerical simulations of disease propagation can be carried out (Section 3). These simulations rely on contact matrices derived from the chosen representation to model transmission dynamics across heterogeneous

population groups. By explicitly incorporating dimensions such as telework feasibility, school closures, and health status, the proposed framework allows to assess the epidemiological impact of targeted public health interventions.

2. Materials and Methods

2.1. Description of the model

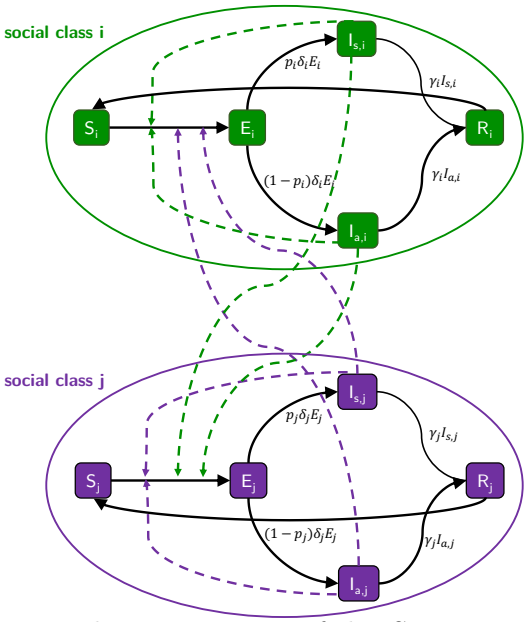

FIGURE 1. Compartmental representation of the SEI_sI_aR-model across social classes, highlighting how infection can spread between individuals of varying social classes.

We assume a partition of the society into two distinct compartmental representations: one based on non-overlapping immunological states, and another based on social classes. These representations result in separate partitions. Let n be the number of social classes. Each social class i is divided in five components of the epidemic flow (Figure 1), *i.e.* the densities of Susceptible individuals (S_i), Exposed individuals (E_i), symptomatic Infected individuals ($I_{s,i}$), asymptomatic Infected individuals ($I_{a,i}$), and Removed individuals(R_i). To build the mathematical model, we followed the standard strategy developed in the literature concerning SIR model [17, 1]. We assumed that susceptible individuals can be infected by infected individuals, while removed individuals can become susceptible again. The dynamics is governed by a system of $5n$ ordinary differential equations (ODE) as follows, $t > 0, 1 \leq i \leq n$,

$$\begin{cases} S_i'(t) = -\beta \left(\sum_{j=1}^{n} \kappa_{ij}(I_{s,j} + I_{a,j}) \right) \dfrac{S_i}{N} + \theta R_i \\ E_i'(t) = \beta \left(\sum_{j=1}^{n} \kappa_{ij}(I_{s,j} + I_{a,j}) \right) \dfrac{S_i}{N} - \delta E_i \\ I_{a,i}'(t) = (1-p)\delta E_i - (\gamma - \mu_i) I_{a,i} \\ I_{s,i}'(t) = p\delta E_i - (\gamma - \mu_i) I_{s,i} \\ R_i'(t) = \gamma(I_{a,i} + I_{s,i}) - \theta R_i. \end{cases} \tag{2.1}$$

The parameters are non-negative. The total living population is $N = \sum_i S_i + E_i + I_{s,i} + I_{a,i} + R_i$. The infection rate is denoted by β and the matrix of contact rate between the class i and the class j is $(\kappa_{ij})_{i,j\in[1\dots n]}$. This matrix is necessarily symmetric and nonnegative and we assume it is normalized in a way to ensure identifiability of β. The waning rate from recovered to susceptible is denoted by θ. The latency rate is denoted by δ and the recovery rate by γ. The ratio of symptomatic among all infected is p. The variation of number of deaths in the class i is assumed to be proportional to the number of infected individuals, *i.e.* $D_i'(t) = \mu_i(I_{s,i} + I_{a,i})$. No new recruit nor immigration is added.

2.2. **Building the contact matrices**

Matrices detailing contact between different age groups can indeed be found in the literature [4, 30]. Below, we first explain how to transform the sampled results into statistics suitable for a given population. Then we describe how to construct the contact matrix, *i.e.* how to derive the matrix of conditional probabilities that an individual belongs to a given class, given his age group, for a specified population partition, using the sampled contact matrix. Finally, we introduce a mathematical procedure to derive contact matrices between n social classes based on the contact matrices established by age. This method leverages the proportion of individuals within each age group belonging to each social class.

2.2.1. *The contact matrix*

Let us denote by N_{samp} the total number of participants in a given survey, by $N_{samp,i}$ the number of participants in the age class i, by N the total population, and N_i the number of individuals in age class i (see Appendix for an example). To build the contact matrix between age classes, we need for a given population:

- An initial sampled contact matrix from a measurement, $B = (b_{ij})_{1\leq i,j\leq n}$, whose coefficients b_{ij} represent the total number of contacts between the individuals of age i in the sample with the individuals of age j in the society, and per day.

- A vector $\pi = (\frac{N_i}{N})_{1\leq i\leq n}$ whose coefficients represent the proportion of the class i in the full population and N is the total number.

One may define the matrix $C = (c_{ij})_{1\leq i,j\leq n}$ such that

$$c_{ij} = \frac{b_{ij}/N_{samp,\,i}}{\sum_{m=1}^{n} b_{im}/N_{samp,\,i}} = \frac{b_{ij}}{\sum_{m=1}^{n} b_{im}}. \tag{2.2}$$

This matrix provides the probability per day of contacts made by an individual from class i with an individual of class j. Since we are using a frequency-dependent model, the contact matrix κ in (2.1) corresponds to the inter-class contacts relative to the class sizes, in other words, for $1 \leq i, j \leq n$, we have

$$\begin{aligned}\kappa_{ji} &= \frac{\text{(probability of contacts from class } j \text{ with class } i) \times \text{(normalized number of contacts with class } j)}{\text{proportion of class i}} \\ &= c_{ji} \times \left(\frac{\sum_{m=1}^{n} b_{jm}}{N_{samp,j}} \right) / \left(\frac{N_i}{N} \right). \end{aligned} \tag{2.3}$$

Remark 2.1. The number of contacts between the social classes i and j is supposed to be equal to the number of contacts between the social classes j and i. Given the relative proportions of social classes π, we can write for all $1 \leq i, j \leq n$,

$$\kappa_{ji} = \kappa_{ij} \Leftrightarrow c_{ji} \times \left(\frac{\sum_{m=1}^{n} b_{jm}}{N_{samp,j}} \right) / \left(\frac{N_i}{N} \right) = c_{ij} \times \left(\frac{\sum_{m=1}^{n} b_{im}}{N_{samp,i}} \right) / \left(\frac{N_j}{N} \right). \tag{2.4}$$

Using this construction, the matrix κ is symmetric. This models the symmetry of contacts. In practice, the relation (2.4) is not satisfied because the matrix B is not exactly symmetric. We impose symmetry by averaging of the extra-diagonal coefficients.

Remark 2.2. A similar construction of contact matrices was proposed in *socialmixr* [20] with the choice

$$\kappa_{ji} = c_{ji} \times \left(\frac{\sum_{m=1}^{n} b_{jm}}{N_{samp,j}} \right).$$

With this choice, it becomes necessary to modify the infection rate $\beta \left(\sum_{j=1}^{n} \kappa_{ij}(I_{s,j} + I_{a,j}) \right) \frac{S_i}{N}$ by $\beta \left(\sum_{j=1}^{n} \kappa_{ij}(I_{s,j} + I_{a,j}) \right) \frac{S_i}{N_i}$ [22, 23, 9].

2.2.2. *Change of classes*

Given a contact matrix B_σ between social classes $(\sigma_i)_{1\leq i\leq n}$, we present an algebraic procedure to build a contact matrix B_τ between other social classes $(\tau_i)_{1\leq i\leq m}$. Each social classes must be a partition of the population (see Appendix for an example).

Theorem 2.3. *Let $\sigma = (\sigma_i)_{1\leq i\leq n}$ and $\tau = (\tau_j)_{1\leq j\leq m}$ be two subdivisions in social classes of the population. We denote by B_σ, respectively B_τ, the contact matrix between individuals of classes σ, respectively τ. Then, there exists an $n \times m-$matrix $R_{\tau\to\sigma}$ such that*

$$B_\tau = R_{\tau\to\sigma}^T B_\sigma R_{\tau\to\sigma}, \tag{2.5}$$

where its coefficients $r_{ij} = P(\sigma = \sigma_i | \tau = \tau_j)$ are the probability that an individual belongs to the class σ_i knowing that he belongs to the class τ_j.
Moreover, we have:

(i) $\sum_{1 \leq k,l \leq p} (b_\tau)_{kl} = \sum_{1 \leq i,j \leq n} (b_\sigma)_{ij}$.

(ii) *If B_σ symmetric, then B_τ is symmetric.*

Proof. (i) We have

$$\begin{aligned}\sum_{u,v} (b_\tau)_{uv} &= \sum_{u,v} \sum_{k,l} r_{ku} (b_\sigma)_{kl}\, r_{l,v} = \sum_{k,l,v} (b_\sigma)_{kl}\, r_{lv} \sum_u r_{ku} \\ &= \sum_{k,l} (b_\sigma)_{kl} \sum_v r_{lv} = \sum_{k,l} (b_\sigma)_{kl}.\end{aligned}$$

(ii) Suppose that B_σ is symmetric. For all $u, v \in \{1, \cdots, m\}$, we have

$$(b_\tau)_{vu} = \sum_{k,l} r_{vu} (b_\sigma)_{kl}\, r_{lu} = \sum_{k,l} r_{lu} (b_\sigma)_{lk}\, r_{ku} = (b_\sigma)_{uv}.$$

□

Similarly, we define the *change of classes* $R_{\sigma \to \tau}$ as the $m \times n$ matrix whose coefficients are the probability that an individual belongs to the class τ_i knowing that he belongs to the class σ_j.

Proposition 2.4. *Let $B_{0,\sigma}$ a raw contact matrix, $R_{\tau \to \sigma}$, respectively $R_{\sigma \to \tau}$, matrix of change of classes from σ to τ, respectively from τ to σ. If the product $R_{\tau \to \sigma} R_{\sigma \to \tau}$ is irreducible, and aperiodic, then the sequence $((R_{\tau \to \sigma} R_{\sigma \to \tau})^T B_{0,\sigma} (R_{\tau \to \sigma} R_{\sigma \to \tau}))^n$ has a limit. This limit matrix encodes the idea that we perform infinitely many alternating transformations from σ to τ and back again.*

Proof. By induction, starting from $B_{0,\sigma}$ being the contact matrix between individuals of classes σ, we define new contact matrices by

$$\begin{array}{ccccccc} B_{0,\sigma} & \to & R_{\tau \to \sigma}^T B_{0,\sigma} R_{\tau \to \sigma} = B_{0,\tau} & \to & R_{\sigma \to \tau}^T (R_{\tau \to \sigma}^T B_{0,\sigma} R_{\tau \to \sigma}) R_{\sigma \to \tau} & & \\ & & & & = R_{\sigma \to \tau}^T B_{0,\tau} R_{\sigma \to \tau} = B_{1,\sigma} & \to & \cdots \\ \text{class } \sigma & & \text{class } \tau & & \text{class } \sigma & & \end{array}$$

The matrices $R_{\tau \to \sigma}$ and $R_{\sigma \to \tau}$ being stochastic by construction, $R_{\tau \to \sigma} R_{\sigma \to \tau}$ is stochastic. Since $R_{\tau \to \sigma} R_{\sigma \to \tau}$ is irreducible and aperiodic, there exists, thanks to the Perron-Frobenius theorem, the stochastic matrix $(R_{\tau \to \sigma} R_{\sigma \to \tau})^\infty$ such that $\lim_{n \to \infty} (R_{\tau \to \sigma} R_{\sigma \to \tau})^n = (R_{\tau \to \sigma} R_{\sigma \to \tau})^\infty$.

□

Remark 2.5. From Bayes' formula, we deduce the coefficients of $R_{\sigma \to \tau}$ from those of $R_{\tau \to \sigma}$ and the relative proportions π^σ of the classes $(\sigma_i)_{1 \leq i \leq n}$ and the relative proportions π^τ of the classes $(\tau_j)_{1 \leq j \leq m}$. Let π_i^σ denote the probability to be of the σclass σ_i and π_j^τ the probability to be of the τclass τ_j. Then

$$(R_{\sigma \to \tau})_{ji} = P(\tau = \tau_j | \sigma = \sigma_i) = P(\sigma = \sigma_i | \tau = \tau_j) \frac{\pi_j^\tau}{\pi_i^\sigma} = (R_{\tau \to \sigma})_{ji} \frac{\pi_j^\tau}{\pi_i^\sigma}.$$

It is worthnoting that the iterative change of representation may lead to a loss of information (see Appendix).

2.3. **Global well-posedness and basic reproduction numbers**

For ease of reading, we briefly outline the derivation of global well-posedness and the basic reproduction number, following Anderson and May [1]. It is standard to check that the domain

$$\Omega = \{(S_i, E_i, I_{s,i}, I_{a,i}, R_i)_{1\leq i\leq n} \in \mathbb{R}_+^{5n}; 0 \leq \sum_{i=1}^{n} S_i + E_i + I_{s,i} + I_{a,i} + R_i \leq N\}$$

is positively invariant. In particular, there exists a unique global in time solution in $\mathcal{C}(\mathbb{R}_+; \Omega)$ as soon as the initial condition lives in Ω.

The basic reproduction number $\mathcal{R}_0$ can be computed thanks to the next generation matrix of the model as in van den Driessche and Watmough [37]. Since the infected individuals are in $E_i, I_{a,i}$ and $I_{s,i}$, these components of infected individuals make the new infections ($\mathcal{F}$) and transitions between compartments ($\mathcal{V}$). They can be rewritten as

$$\mathcal{F} = \begin{pmatrix} \left(\beta \sum_{j=1}^{p} \kappa_{1j}(I_{s,j} + I_{a,j})\right) \frac{S_1}{N} & \cdots & \beta \left(\sum_{j=1}^{p} \kappa_{nj}(I_{s,j} + I_{a,j})\right) \frac{S_n}{N} \\ 0 & \cdots & 0 \\ 0 & \cdots & 0 \end{pmatrix},$$

$$\mathcal{V} = \begin{pmatrix} \delta E_1 & \cdots & \delta E_n \\ (\gamma + \mu_i) I_{a,1} - (1-p)\delta E_1 & \cdots & (\gamma + \mu_i) I_{a,n} - (1-p)\delta E_n \\ (\gamma + \mu_i) I_{s,1} - p\delta E_1 & \cdots & (\gamma + \mu_i) I_{s,n} - p\delta E_n \end{pmatrix}.$$

Thus, at Disease Free Equilibrium (DFE), we obtain

$$F = \begin{pmatrix} 0 & \frac{\beta}{N} F_n & \frac{\beta}{N} F_n \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \text{ and } V = \begin{pmatrix} \delta I_n & 0 & 0 \\ -(1-p)\delta I_n & J_n & 0 \\ -p\delta I_n & 0 & J_n \end{pmatrix},$$

where I_n is the identity matrix of size n, J_n is the diagonal matrix of size n whose coefficients $(J_n)_i = \gamma + \mu_i$, and

$$F_n = \begin{pmatrix} \kappa_{11} S_1 & \kappa_{12} S_1 & \cdots & \kappa_{1n} S_1 \\ \kappa_{21} S_2 & \kappa_{22} S_2 & \cdots & \kappa_{2n} S_2 \\ \vdots & \vdots & & \vdots \\ \kappa_{n1} S_n & \kappa_{n2} S_n & \cdots & \kappa_{nn} S_n \end{pmatrix}.$$

We deduce that the basic reproduction number is given by $\mathcal{R}_0 = \rho(FV^{-1})$, where FV^{-1} is the next generation matrix at the DFE. For uniform mortality rate $\mu_i = \mu$,

$$\rho(F\,V^{-1}) = \frac{\beta}{N(\gamma + \mu)} \rho(F_n).$$

Biologically speaking, the term $\frac{\beta}{\gamma+\mu}$ represents the transmission rate by infected individuals during the average infection period $1/(\gamma + \mu)$.
Note that the disease-free equilibrium is given by

$$\text{DFE} = (0, ..., 0, S_1^*, S_2^*, ..., S_n^*) \text{ with } \sum_{i=1}^{n} S_i^* = N,$$

while the endemic equilibrium can be written as $\text{EE} = (S_i^*, E_i^*, I_{s,i}^*, I_{a,i}^*, R_i^*)_{1\leq i\leq n}$ with

$$S_i^* = \frac{\delta N E_i^*}{\frac{\beta}{\gamma} \sum_{j=1}^{n} \kappa_{i,j} E_j^*}, I_{a,i}^* = \frac{(1-p)\delta}{\gamma} E_i^*, I_{s,i}^* = \frac{p\delta}{\gamma} E_i^*, R_i^* = \frac{\delta}{\theta} E_i^*.$$

Following Zhou et al. [40], we can prove the following result.

Theorem 2.6. *Assume that the matrix $K = (\kappa_{i,j})_{1\leq i,j\leq n}$ is irreducible.*

(1) *If $\mathcal{R}_0 \leq 1$, the disease-free equilibrium DFE is globally asymptomatically stable.*

(2) *If $\mathcal{R}_0 > 1$, the disease-free equilibrium DFE is unstable and the endemic equilibrium EE is globally asymptomatically stable.*

3. Results

3.1. Simulations involving age group contacts

The contact matrix is built for France from the French connexion matrix in Béraud et al. [4]. Recall that this matrix results from biased samples. We construct the contact matrix following the procedure explained in the section 2.2.1. The Infection Fatality Ratio (IFR) by age group [35] is used to compute the number of deaths. Parameters for the simulation are described in C. Simulations start with a initial proportion of infected equal to 0.001, evenly distributed among the classes. The code is available in our PLMLab repository. It includes code to simulate COVID-19 dynamics across social classes, along with a notebook specifically for modeling the spread among white- and blue-collar workers.

To simulate the reduction of contacts, a fraction ν of all contacts is removed. Specifically, if we intend to reduce the number of contacts of class i_0, and given the initial sampled contact matrix B, we define the new contact matrix B_ν as follows:

$$\begin{aligned}(b_\nu)_{i_0 i_0} &= (1-\nu) b_{i_0 i_0} \text{ (intra-class contacts)}\\ (b_\nu)_{ij} &= (1-\nu) b_{ij}, \text{ for } i = i_0 \text{ or } j = i_0 \text{ (inter-classes contacts)}\\ (b_\nu)_{ij} &= b_{ij}, \text{ for } i \neq i_0 \text{ and } j \neq i_0.\end{aligned}$$

We conduct simulations spanning $T = 600$ days to compute the number of infected individuals. The number of infected individuals is $\sum_{j=1}^{n} \int_0^T \frac{I_{s,j}(t) + I_{a,j}(t)}{N} dt$. This expression represents the cumulative number of individuals infected over time, where $I_{s,j}(t)$ and $I_{a,j}(t)$ denote the number of symptomatic and asymptomatic infected individuals in age group j at time t, respectively, and N is the total population. We also compute the maximum number of infected individuals as $\max_{0\leq t\leq T} (I_{s,j}(t) + I_{a,j}(t))$. Additionally, we determine the cumulative number of deaths over time due to the epidemic as $\int_0^T D(t)dt$. The population is subdivided into three age groups: 0 to 17 years old, 18 to 64 years old, and 65 years old or older ($n = 3$). These quantities are evaluated with respect to the fraction of limited contacts ν of individuals aged 65 years and older.

3.1.1. *Containment of the 65 years old or older*

Containment measures are targeted exclusively towards individuals aged 65 and above. They represent 20% of the French population.

In Figure 2, we can discern a notable trend: the proportion of infected individuals within the 65 and older age class demonstrates a nearly linear decline as ν, the proportion of contacts removed within this specific group, increases. As expected, when all contacts are restricted ($\nu = 1$), this

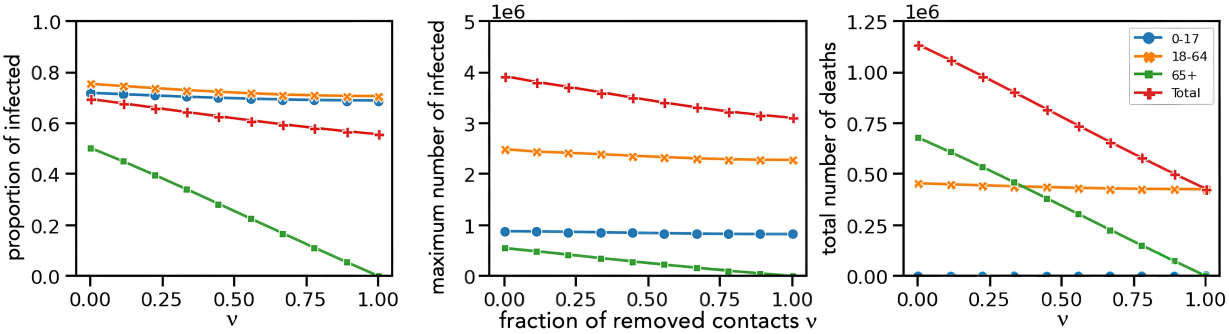

FIGURE 2. Effect of restricting contacts among individuals aged 65 and older on proportion of infected individuals (left), maximum over time number of infected (center), and total number of deaths (right) with respect to ν. The x-axis represents the fraction ν of contacts suppressed for the 65 and older age group. The outcomes are presented for each age group (0 to 17 years old in blue dot, 18 to 64 years old in orange cross, and 65 years old or older in green box) as well as for the entire population in red plus.

age class becomes entirely isolated, halting the spread of the epidemic within it. Interestingly, the containment measures targeting individuals aged 65 and above show minimal impact on other age demographics. Upon analyzing the entire population, we observe a significant 20% reduction in the overall proportion of infected individuals. Moreover, the maximum number of infected of the epidemic experiences a decline of approximately 21% when $\nu = 1$. This reduction primarily stems from the over 65 age group, with a secondary effect noticeable in the 18-64 age group. Finally, the containment strategy targeting individuals over 65 years old exerts a substantial influence on the overall mortality rate by the end of the epidemic, resulting in a remarkable decrease of approximately 62%. This effect is particularly pronounced due to the fact that this age group exhibits the highest Infection Fatality Rate (IFR) and consequently contributes the most to the total number of deaths.

3.1.2. *Containment of 0 to 17 years old*

The containment measures applied to individuals aged 0 to 17 years old, who also represent 20% of the French population, play an important role in epidemic control. By restricting contacts within this age group, we observe a significant impact on the dynamics of the epidemic.

As expected, implementing containment measures for 0-17 group results in a significant decline across various quantities reflecting the spread of the epidemic within this age group. Examining the proportion of infected individuals and maximum number of infected, as depicted in Figure 3, we observe that containing 0-17 group also influences the 18-64 age group, albeit to a lesser extent, and even marginally affects the over-65 age group. Remarkably, in the overall population, the reduction in both the proportion of infected individuals and the maximum number of infected exceeds that observed with the containment of individuals over 65 years old. This is because unconfined age groups experience a more pronounced impact from the containment of young people. This underscores the interdependency among different age demographics in epidemic transmission dynamics and emphasizes the importance of comprehensive containment strategies to effectively mitigate the spread of the virus across all age groups. The total number

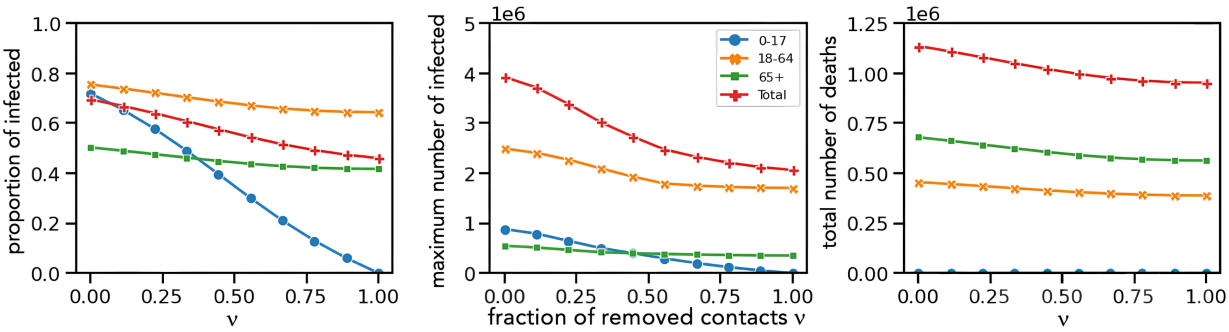

Figure 3. Effect of restricting contacts among individuals aged between 0 and 17 years old on proportion of infected individuals (left), maximum number of infected (center), and total number of deaths (right). The x-axis represents the fraction ν of contacts suppressed for the 65 and older age group. The outcomes are presented for each age group (0 to 17 years old in blue, 18 to 64 years old in orange, and 65 years old or older in green) as well as for the entire population (red).

of deaths at the conclusion of the epidemic tends to be higher when implementing containment measures for 0-17 rather than older people. It decreases of about 16% to be compared to the fall of 62% in the case of the containment of individuals over 65 years old. This outcome aligns with the understanding that young people typically exhibit the lowest IFR among age groups, consequently contributing less to overall mortality rates.

3.1.3. *Containment of 0-17 and 65+*

We implement containment measures for both 0 to 17 years old individuals and 65 years old and older individuals. A fixed proportion of suppressed contacts, denoted as $\nu = 0.25$, is implemented across the entire population. From this 25%, a specific fraction ν_{old} represents the suppressed contacts of individuals aged 65 years and older, while the remaining proportion $(1-\nu_{old})$ pertains to the suppressed contacts of individuals aged 0-17.

As shown in Figure 4, the proportion of infected individuals among those aged over 65 decreases as ν_{old} increases, whereas among 0-17 group, it increases. Interestingly, the proportion of infected individuals in the 18-64 age group initially decreases slightly with ν_{old}, reaches a minimum, and then starts to increase. A similar trend is observed for the maximum number of infected individuals. This phenomenon is further illustrated by the behavior of $\mathcal{R}_0$ as a function of ν_{old} in Figure 5. The value of the basic reproduction number $\mathcal{R}_0$ reaches its minimum at around 65% of suppressed contacts. However, the variation in $\mathcal{R}_0$ remains relatively modest across different values of ν_{old}.

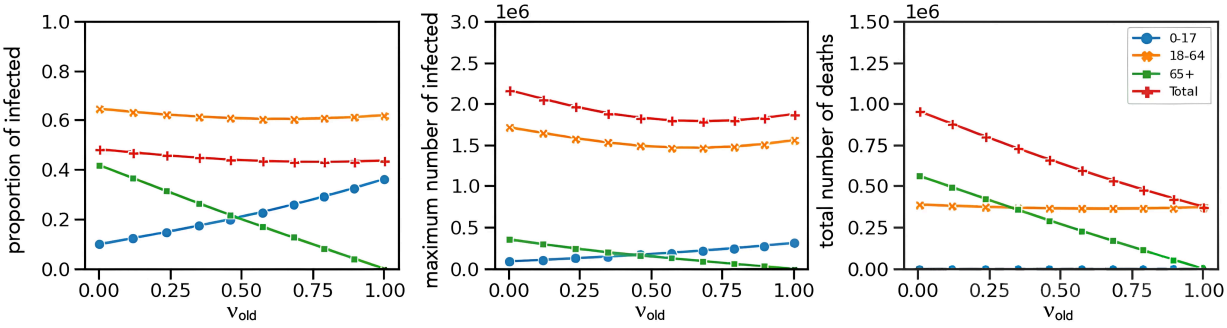

FIGURE 4. Effect of simultaneous restricting contacts among individuals aged between 0 and 17 years old and 65 and older on proportion of infected individuals (left), maximum number of infected (center), and total number of deaths (right). The x-axis represents the fraction ν_{old} of contacts suppressed for the 65 and older age group. The outcomes are presented for each age group (0 to 17 years old in blue dot, 18 to 64 years old in orange cross, and 65 years old or older in green box) as well as for the entire population (red plus).

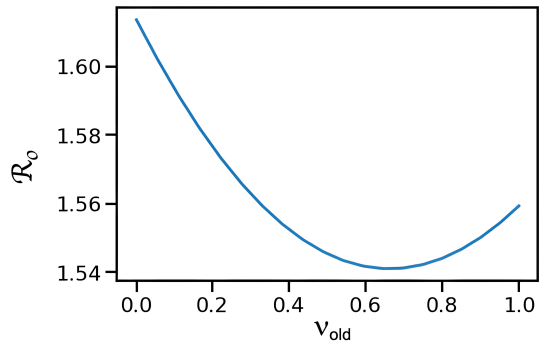

FIGURE 5. Impact of simultaneous restricting contacts among individuals aged between 0 and 17 years old and 65 and older on the basic reproduction number $\mathcal{R}_0$. The x-axis represents the fraction of contacts suppressed for the 65+ age group.

The mortality, significantly influenced by the containment measures targeting individuals over 65, declines with respect to ν_{old}. If we focus solely on the proportion and maximum count of infected individuals, an optimal approach entails limiting contacts across multiple demographic groups while maintaining a constant number of contacts. However, if we aim at minimizing the number of deaths, concentrating on restricting contacts within the oldest age group emerges as the most effective strategy. This is primarily due to the higher mortality rate in the 65+ age group.

3.2. **Simulations involving social class contacts**

We concentrate on three categories. Initially, we segment the population based on socio-professional categories. Subsequently, we divide the population according to their school attendance status. Finally, we categorize the population based on their self-perceived health status.

3.2.1. *White collars telework*

We simulate the telecommuting of white-collar workers. To construct the contact matrix, we utilize the French connection contact matrix of the French population ([4]), which is categorized into four age groups: 0 to 17 years old, 18 to 29 years old, 30 to 59 years old, and 60 years old and above. The relative proportions of white- and blue-collar workers are obtained from publicly available data (INSEE Professions et secteurs d'activité). Following the procedure outlined in Section 2.2.2, we partition the population into five classes: 0 to 17 years old, 30 to 60 years old white-collar workers, 30 to 60 years old blue-collar workers, and individuals aged 60 years and above. The impact of COVID19 on blue-collar workers is partially adressed in [33].
The matrix of class changes, denoted as R, is defined as follows:

	0-17	18-29	White-collar	Blue-collar	60+
0-17	1	0	0	0	0
18-29	0	1	0	0	0
30-60	0	0	0.45	0.55	0
60+	0	0	0	0	1

and the contact matrix κ is

	0-17	18-29	White-collar	Blue-collar	60+
0-17	24	6.2	7.4	7.4	2.6
18-29	6.2	34	11	11	3.6
White-collar	7.4	11	16	16	6.9
Blue-collar	7.4	11	16	16	6.9
60+	2.6	3.6	6.9	6.9	10

With this partition, white-collar workers represent 17% of the population, while white-collar workers represent 21% of the population. It is important to note that, under normal circumstances, white-collar and blue-collar workers have identical average number of contacts. The key difference lies in the white-collar workers' ability to adapt their work conditions, *e.g.* by using remote work reducing the need for commuting and physical presence in an office. Telework scenarios are modeled by adjusting the parameter ν, representing the fraction of restricted interpersonal interactions among white-collar workers.

White-collar telecommuting emerges as an influential intervention for this demographic segment, as depicted in Figure 6. Implementing telecommuting measures for white-collar workers yields significant declines in terms of reducing both the proportion of infected individuals and the maximum number of cases. The overall mortality rate in the population has decreased by 38%. This decline can be attributed to the survival of white-collar workers and a reduction in the number of individuals over the age of 60 affected by the epidemic.

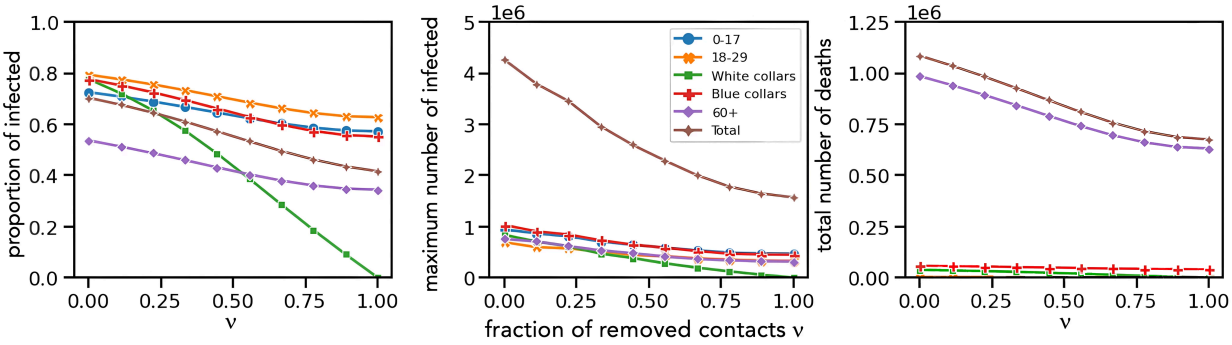

FIGURE 6. Effect of white-collar telework on proportion of infected individuals (left), maximum number of infected (center), and total number of deaths (right). The x-axis represents the fraction ν of contacts suppressed for the white-collar group. The outcomes are presented for each social class (0 to 17 years old in blue, 18 to 29 years old in orange, 30 to 59 years old white-collars in green, 30 to 59 years old blue-collars in red and 60 years old or older in violet) as well as for the entire population (purple).

3.2.2. *School closures*

We are now incorporating the total closure of schools into our simulation. To establish the contact matrix, we use the French connection contact matrix for the population of France [4], segmented into seven age groups: 0-2 years old, 3-6 years old, 7-10 years old, 11-14 years old, 15-17 years old, 18-24 years old, and 25 years old and above. The relative proportions of pupils and students are obtained from publicly available data (Ministère de l'éducation nationale - Repères et références statistiques). Employing the methodology introduced in Section 2.2.2 once again, we categorize the population into six groups: kindergarten, elementary school, middle school, high school, university, and individuals not affiliated with any of these educational institutions. This final class is necessary to maintain a complete partition of society. The matrix of change of classes R is

	Other	Kinder	Elementary	Middle	High	University
0-2	1	0	0	0	0	0
3-6	0	1	0	0	0	0
7-10	0	0	1	0	0	0
11-14	0	0	0	1	0	0
15-18	0.1	0	0	0	0.9	0
19-24	0.4	0	0	0	0	0.6
25+	0.95	0	0	0	0	0.05

and the contact matrix κ is

	Other	Kinder	Elementary	Middle	High	University
Other	9.5	7	5.7	5.1	7.1	9.3
Kinder	7	71	16	4.6	3.9	4.2
Elementary	5.7	16	62	13	4.5	4.1
Middle	5.1	4.6	13	99	22	5
High	7.1	3.9	4.5	22	150	17
University	9.3	4.2	4.1	5	17	23

School closures are modeled by setting the fraction of restricted interpersonal interactions to $\nu = 1$. The results are outlined in Table 1.

Confined class	% of the total population	Number of contacts	Reduction in the proportion of infected individuals	Reduction in the maximum number of infected	Reduction in the number of deaths
Kindergarten	3%	10.1	7%	13%	4%
Elementary	6%	10.8	13%	23%	7%
Middle school	6%	12.8	15%	3%	8%
High school	3%	14.3	10%	22%	6%
University	8%	10	19%	3%	14%

Table 1. For each scenario, we indicate the relative size of the class confined, the average number of contacts in this class and the percentage of decrease of the indicators between a simulation without contact limitation and a total limitation according to the chosen scenario.

Notably, closing kindergartens exhibits the least impact on mitigating the indicators of epidemic spread. Conversely, shuttering universities appears to be the most effective intervention. The effectiveness of closing a specific type of school is primarily elucidated by the relative size of the affected class. Additionally, the average number of contacts within the affected class also plays a role, albeit to a lesser degree. Lastly, the variation pertains to relatively small numbers of deaths, as the death rate among individuals in these classes is very low.

3.2.3. *Containment of the self-perceived most vulnerable population*

Here, our aim is to categorize individuals based on their self-perceived health status. The French connection contact matrix [4] is divided into fixed age groups: 0 to 15 years old, 16 to 29 years old, 30 to 49 years old, 50 to 74 years old, and 75 years old and above. Health status is classified into four categories: Very Good, Good, Medium, and Bad. The relative proportions of each health status and students are obtained from publicly available data (INSEE - Santé). The matrix of change of classes R is

	Very good	Good	Medium	Bad
0-15	0.504	0.395	0.083	0.018
16-29	0.504	0.395	0.083	0.018
30-49	0.307	0.467	0.182	0.046
50-74	0.118	0.467	0.309	0.079
75+	0.035	0.287	0.466	0.2165

and the contact matrix κ is

	Very good	Good	Medium	Bad
Very good	13	11	8.2	7.4
Good	11	9.7	8.6	8
Medium	8.2	8.6	8.5	8.2
Bad	7.4	8	8.2	8.1

We gradually, by a factor ν, isolate individuals from the "feeling bad" group. The results presented in Figure 7 reveal modest percentage reductions in the proportion of total infected individuals (11%), the maximum number of infected (18%), and the number of deaths (9%). The limited reduction in the number of deaths across the entire population can be attributed to the IFR within the isolated group. In this simulation, IFR is solely determined by an individual's age, with their perceived health status only influencing their classification into a specific group.

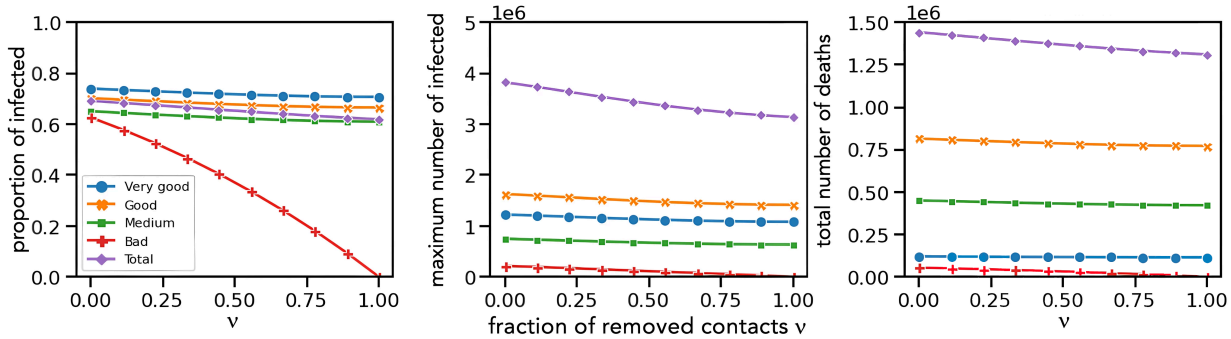

FIGURE 7. Effect of isolation of "feeling bad" group on proportion of infected individuals (left), maximum number of infected (center), and total number of deaths (right). The x-axis represents the fraction ν of contacts suppressed for the "feeling bad" group. The outcomes are presented for each social class (Very good in blue, Good in orange, Medium in green, Bad in red) as well as for the entire population (violet).

Scenario: Containment of	% of the total population	Nb of contacts	Reduction in the proportion of infected individuals	Reduction in the maximum number of infected	Reduction in the number of deaths
65+	20%	6.9	20%	21%	62%
0-17	20%	10.3	34%	47%	16%
white-collars	17%	10.1	41%	63%	38%
university	8%	10	19%	3%	14%
feeling bad	6%	2.7	11%	18%	9%

TABLE 2. For each scenario, we indicate the percentage of decrease of the indicators between a simulation without contact limitation and a total limitation according to the chosen scenario.

4. Discussion

Our results are summarized in Table 2. Exploring strategies that selectively confine specific population groups is particularly appealing, as it seeks to mitigate the economic and social costs associated with blanket lockdowns [14]. While differentiated strategies based on age or location have already been extensively studied [11, 8, 9], one of the main contributions of this work is to provide a simple and general methodology to extend such analyzes to alternative social partitioning, going beyond the mere splitting of rows in a contact matrix. We illustrate this approach by examining telecommuting among white-collar workers, the closure of different educational institutions, and isolation strategies based on self-perceived health status. Beyond age-based strategies, our framework enables the analysis of containment measures targeting other social dimensions in a unified and tractable manner. From a methodological standpoint, a key strength of our approach lies in the mathematical formulation of changes in social classification through the transition matrix R. This algebraic representation is particularly useful from a practical perspective: it allows to modify social stratifications without redefining the entire contact structure, while remaining within the familiar contact matrix framework. As a result, the procedure is straightforward to implement, computationally efficient, and easily adaptable to a wide range of policy scenarios.

Subdividing the population into age groups yields findings that are fully consistent with the existing literature. In particular, restricting contacts among the elderly leads to a more substantial reduction in the total number of deaths than limiting contacts among younger individuals [9]. By contrast, restricting interactions among young people primarily reduces the overall proportion of infected individuals and the peak number of infections. These indicators are driven jointly by the relative size of the restricted group and its average number of contacts. Mortality outcomes, on the other hand, are predominantly sensitive to the Infection Fatality Rate, which explains why confining younger individuals has a comparatively limited effect on the number of deaths. When restrictions target two age groups, strategies that minimize infections and epidemic peaks differ from those that minimize mortality, revealing trade-offs between public health goals.

When compared with age-targeted containment strategies, telecommuting among white-collar workers proves especially efficient in reducing both the proportion of infected individuals and the maximum number of cases (Table 2). Notably, the introduction of telework in this group generates substantial indirect effects, leading to declines in epidemic indicators across all other population classes. These spillover effects underscore the relevance of occupational structure in shaping epidemic dynamics. Our simulations indicate that telecommuting among white-collar workers produces one of the largest reductions in both the total proportion of infected individuals and the epidemic peak, ranking second only to the most stringent age-based strategies. Telework also exerts a meaningful influence on mortality outcomes, as it indirectly reduces exposure among more vulnerable groups, including individuals aged 65 and above..

When the population is stratified by educational attendance, university closures emerge as the most impactful intervention. This result is largely explained by the relatively large size of the university student population compared to other educational groups. By contrast, containment measures based on self-perceived health status do not yield statistically significant effects at the population level, primarily due to the limited size of the group identifying as unhealthy.

A central methodological assumption is that contact matrices for alternative social representations are derived from the original age-stratified matrix under a conditional proportional mixing assumption. This implies that individuals within the same age group have, on average, identical

contact patterns across social classes.Although this approach enables a simple and data-efficient implementation, it may not fully capture real-world heterogeneities in contact behavior. This assumption therefore limits the applicability of our results and motivates extensions that incorporate explicit class-based assortativity. An additional and conceptually important observation emerging from our analysis is that iterative changes of social representation may lead to a loss of information. While the transformation of contact matrices through successive applications of the transition matrix is algebraically well defined and computationally convenient, it is not necessarily information preserving.

From a modeling perspective, these observations clarify the strengths and limitations of the proposed framework. The approach is well suited to contexts in which age-based mixing dominates and social-class assortativity is weak, but it may underrepresent structured interactions when strong class-based preferences exist. Moreover, the choice of social representation is consequential, as information lost through aggregation cannot be recovered, motivating future extensions that incorporate richer assortative mixing data. Finally, the findings of this study should be interpreted qualitatively, in light of the simplifying assumptions inherent in the underlying SEIR framework. The model relies on average infection rates and partially observed contact patterns across social classes, potentially omitting other relevant drivers of epidemic dynamics. Moreover, the notion of contact is limited by survey data based on self-reported interactions within a 1.5-meter radius, without objective measurement, potentially leading to an underestimation of longer-range or indirect contacts. Despite these limitations, the proposed framework provides a practical and transparent tool for incorporating social heterogeneity into epidemic models, offering valuable insights into the differential impacts of targeted containment strategies. It also provides a systematic method for transitioning from one societal partition to another, regardless of how the population is grouped.

Bibliography

[1] Roy M Anderson and Robert M May. *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, 1991.

[2] Maria Antònia Barceló and Marc Saez. Impact of the covid-19 pandemic on the socioeconomic inequalities in mortality in spanish provinces. *Journal of epidemiology and global health*, pages 1–23, 2023.

[3] Jamie Bedson, Laura A. Skrip, Danielle Pedi, Sharon Abramowitz, Simone Carter, Mohamed F. Jalloh, Sebastian Funk, Nina Gobat, Tamara Giles-Vernick, Gerardo Chowell, João Rangel de Almeida, Rania Elessawi, Samuel V. Scarpino, Ross A. Hammond, Sylvie Briand, Joshua M. Epstein, Laurent Hébert-Dufresne, and Benjamin M. Althouse. A review and agenda for integrated disease models including social and behavioural factors. *Nature Human Behaviour*, 5(7):834–846, 2021.

[4] Guillaume Béraud, Sabine Kazmercziak, Philippe Beutels, Daniel Levy-Bruhl, Xavier Lenne, Nathalie Mielcarek, Yazdan Yazdanpanah, Pierre-Yves Boëlle, Niel Hens, and Benoit Dervaux. The french connection: the first large population-based contact survey in france relevant for the spread of infectious diseases. *PloS one*, 10(7):e0133203, 2015.

[5] M. Bisiada. Discourse and social cohesion in and after the covid-19 pandemic. *Media and Communication*, 10(2):204–213, 2022.

[6] Caroline Buckee, Abdisalan Noor, and Lisa Sattenspiel. Thinking clearly about social aspects of infectious disease transmission. *Nature*, 595(7866):205–213, July 2021.

[7] Bruno Buonomo, Rossella Della Marca, Alberto d'Onofrio, and Maria Groppi. A behavioural modelling approach to assess the impact of covid-19 vaccine hesitancy. *Journal of Theoretical Biology*, 534:110973, 2022.

[8] Daniela Calvetti, Alexander P Hoover, Johnie Rose, and Erkki Somersalo. Metapopulation network models for understanding, predicting, and managing the coronavirus disease covid-19. *Frontiers in Physics*, page 261, 2020.

[9] Maria Chikina and Wesley Pegden. Modeling strict age-targeted mitigation strategies for covid-19. *PloS one*, 15(7):e0236237, 2020.

[10] Robert M. Cohen, Christophe Delacourt, C. Gras-Le Guen, and Elise Launay. Covid-19 and schools. guidelines of the french pediatric society. *Archives De Pediatrie*, 2020.

[11] Pietro Coletti, Pieter Libin, Oana Petrof, Lander Willem, Steven Abrams, Sereina A Herzog, Christel Faes, Elise Kuylen, James Wambua, Philippe Beutels, et al. A data-driven metapopulation model for the belgian covid-19 epidemic: assessing the impact of lockdown and exit strategies. *BMC infectious diseases*, 21(1):1–12, 2021.

[12] Sara Comai, David Simeone, Silvia Mastrolembo Ventura, and Angelo Luigi Camillo Ciribini. Digital simulation modelling for a school re-opening during the covid-19 pandemic. *Proceedings of the Institution of Civil Engineers*, 2022.

[13] M. Crean, D. Devine, B.G. Moore, G. Martinez-Sainz, J. Symonds, S. Sloan, and E.E. Farrell. Social class, covid-19 and care: Schools on the front line in ireland during the covid-19 pandemic. *British Journal of Sociology of Education*, 44:452–466, 2023.

[14] Pragyan Deb, Davide Furceri, Jonathan D Ostry, and Nour Tawk. The economic effects of covid-19 containment measures. *Open Econ Rev*, 3:1–32, 2022.

[15] Laura Di Domenico, Giulia Pullano, Chiara E. Sabbatini, Pierre-Yves Boëlle, and Vittoria Colizza. Modelling safe protocols for reopening schools during the covid-19 pandemic in France. *Nature Communications*, 2021.

[16] Laura Di Domenico, Martina L. Reichmuth, and Christian L. Althaus. Individual and neighborhood based socioeconomic factors relevant for contact behaviour and epidemic control. *Communications Medicine*, 6(1):26, 2025.

[17] O. Diekmann and J.A.P. Heesterbeek. *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*. Chichester: John Wiley, 2000.

[18] Rey Audie S. Escosio, Olive R. Cawiding, Bryan S. Hernandez, Renier G. Mendoza, Victoria May P. Mendoza, Rhudaina Z. Mohammad, Carlene P.C. Pilar-Arceo, Pamela Kim N. Salonga, Fatima Lois E. Suarez, Polly W. Sy, Thomas Herald M. Vergara, and Aurelio A. de los Reyes. A model-based strategy for the covid-19 vaccine roll-out in the philippines. *Journal of Theoretical Biology*, 573:111596, 2023.

[19] Andreas Frey, Andrea M. Tilstra, and Mark D. Verhagen. Inequalities in healthcare use during the covid-19 pandemic. *Nature Communications*, 15(1894):1–30, 2024.

[20] Sebastian Funk, Lander Willem, and Hugo Gruson. *socialmixr: Social Mixing Matrices for Infectious Disease Modelling*, 2024. R package version 0.3.2.9000, https://epiforecasts.io/socialmixr/.

[21] Jennifer R. Head, Kristin L. Andrejko, Qu Cheng, Philip A. Collender, Sophie Phillips, Anna Boser, Alexandra K. Heaney, Christopher M. Hoover, Sean L. Wu, Graham R. Northrup, Karen Click, Naomi S. Bardach, Joseph A. Lewnard, and Justin V. Remais. School closures reduced social mixing of children during COVID-19 with implications for transmission risk and school reopening policies. *Journal of the Royal Society Interface*, 2021.

[22] Fatima-Zahra Jaouimaa, Daniel Dempsey, Suzanne Van Osch, Stephen Kinsella, Kevin Burke, Jason Wyse, and James Sweeney. An age-structured seir model for covid-19 incidence in Dublin, Ireland with framework for evaluating health intervention cost. *Plos one*, 16(12):e0260632, 2021.

[23] Mark Kimathi, Samuel Mwalili, Viona Ojiambo, and Duncan Kioi Gathungu. Age-structured model for covid-19: Effectiveness of social distancing and contact reduction in kenya. *Infectious Disease Modelling*, 6:15–23, 2021.

[24] Stephen M. Kissler, Christine Tedijanto, Edward Goldstein, Yonatan H. Grad, and Marc Lipsitch. Projecting the transmission dynamics of sars-cov-2 through the postpandemic period. *Science*, 368(6493):860–868, 2020.

[25] S. A. Lauer, K.H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, and J. Lessler. The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: Estimation and application. *Annals of Internal Medicine*, 03 2020.

[26] Pengyu Liu, Lisa McQuarrie, Yexuan Song, and Caroline Colijn. Modelling the impact of household size distribution on the transmission dynamics of COVID-19. *Journal of the Royal Society Interface*, 2021.

[27] Belinda Lunnay, Barbara Toson, Carlene Wilson, Carlene Wilson, Emma R. Miller, Samantha B Meyer, Ian N. Olver, Kristen Foley, Jessica Thomas, and Paul Ward. Social class and changes in australian women's affect and alcohol consumption during covid-19. *Frontiers in Public Health*, 9:645376–645376, 2021.

[28] Adriana Manna, Lorenzo Dall'Amico, Michele Tizzoni, Márton Karsai, and Nicola Perra. Generalized contact matrices allow integrating socioeconomic variables into epidemic models. *Science Advances*, 10(41):eadk4606, 2024.

[29] Michael F. McGovern and Keith A. Wailoo. Epidemic inequities: Social and racial inequality in the history of pandemics. *Isis*, 114(S1):S206–S246, 2023.

[30] Joël Mossong, Niel Hens, Mark Jit, Philippe Beutels, Kari Auranen, Rafael Mikolajczyk, Marco Massari, Stefania Salmaso, Gianpaolo Scalia Tomba, Jacco Wallinga, et al. Social

contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS medicine*, 5(3):e74, 2008.

[31] F. Müller, S. Goudeau, N. M. Stephens, C. Aelenei, and R. Sanitioso. Social-class inequalities in distance learning during the covid-19 pandemic: Digital divide, cultural mismatch, and psychological barriers. *Revue internationale de psychologie sociale*, 36(1), 2023.

[32] Daniel P. Oran and Eric J. Topol. Prevalence of Asymptomatic SARS-CoV-2 Infection. *Annals of Internal Medicine*, 173(5):362–367, 2020.

[33] Carl Edward Destin Pierre. Citizens facing covid-19: Inequalities and vulnerabilities of the working class the world before covid-19. 10(3):270–276, 2020.

[34] Ron Sender, Yinon Bar-On, Sang Woo Park, Elad Noor, Jonathan Dushoff, and Ron Milo. The unmitigated profile of COVID-19 infectiousness. *eLife*, 11:e79134, 2022.

[35] Reed J D Sorensen, Ryan M Barber, David M Pigott, Austin Carter, Cory N Spencer, Samuel M Ostroff, Reiner Robert C, Cristiana Abbafati, Christopher Adolph, and Adrien Allorant. Variation in the covid-19 infection–fatality ratio by age, time, and geography during the pre-vaccine era: a systematic analysis. *The Lancet*, 399:1469–88, 2022.

[36] Michele Tizzoni, Elaine O. Nsoesie, Laetitia Gauvin, Márton Karsai, Nicola Perra, and Shweta Bansal. Addressing the socioeconomic divide in computational modeling for infectious diseases. *Nature Communications*, 13(1):2897, May 2022.

[37] P van den Driessche and J. Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences*, 180(1-2):29–48, 2000.

[38] Elise Whitley, Gerard McCartney, Mel Bartley, and Michaela Benzeval. Examining the impact of different social class mechanisms on health inequalities: A cross-sectional analysis of an all-age UK household panel study. *Social Science & Medicine*, 312:115383, 2022.

[39] Jon Zelner, Nina B. Masters, Ramya Naraharisetti, Sanyu A. Mojola, Merlin Chowkwanyun, and Ryan Malosh. There are no equal opportunity infectors: Epidemiological modelers must rethink our approach to inequality in infection risk. *PLoS computational biology*, 18(2):e1009795, February 2022.

[40] Linhua Zhou, Yan Wang, Yanyu Xiao, and Michael Y. Li. Global dynamics of a discrete age-structured sir epidemic model with applications to measles vaccination strategies. *Mathematical Biosciences*, 308:27–37, 2019.

Appendix A. Example of construction of contact matrices

We explain the construction of a contact matrix, denoted as κ, utilizing synthetic data. The population is subdivided into two age groups: young (Y) and old (O). Suppose that the relative proportion of these two classes in the total population is $\pi = (\pi_Y, \pi_O) = (0.2, 0.8)$. Taking a

sample with $N_{samp} = 1200$ participants, whose $N_{samp,Y} = 1000$ and $N_{samp,O} = 200$, we measure the sampled contact matrix

$$B = \begin{pmatrix} 6000 & 4000 \\ 100 & 900 \end{pmatrix}.$$

From B, $N_{samp,Y}$ and $N_{samp,O}$ we derive the matrix

$$C = \begin{pmatrix} 0.6 & 0.4 \\ 0.1 & 0.9 \end{pmatrix}.$$

Finally, we build κ using C and π as

$$\kappa = \begin{pmatrix} 0.6 \times 10/0.2 & 0.4 \times 10/0.8 \\ 0.1 \times 5/0.2 & 0.9 \times 5/0.8 \end{pmatrix} = \begin{pmatrix} 30 & 5 \\ 2.5 & 5.625 \end{pmatrix},$$

which is symmetrized as follows

$$\kappa := \begin{pmatrix} 30 & \frac{2.5+5}{2} \\ \frac{5+2.5}{2} & 5.625 \end{pmatrix} = \begin{pmatrix} 30 & 3.75 \\ 3.75 & 5.625 \end{pmatrix}.$$

Appendix B. Change of classes

Let us assume the population is divided into two classes, denoted as $\sigma = (\sigma_1, \sigma_2)$, or into three classes, denoted as $\tau = (\tau_1, \tau_2, \tau_3)$, with known relative proportions π for each social class. It is important to note that iterative changes in representation may not be reversible and can result in information loss. This loss can be quantified by measuring the distance between matrix $R_{\tau\to\sigma}R_{\sigma\to\tau}$ and the identity matrix I_n.

First, consider another social class, denoted by τ, wherein an individual belonging to a category σ can be allocated to only one specific social class among the various classes of τ, *e.g.*

$$R_{\tau\to\sigma} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0.5 \end{pmatrix} = \begin{pmatrix} P(x \in \sigma_1 | x \in \tau_1) & P(x \in \sigma_1 | x \in \tau_2) & P(x \in \sigma_1 | x \in \tau_3) \\ P(x \in \sigma_2 | x \in \tau_1) & P(x \in \sigma_2 | x \in \tau_2) & P(x \in \sigma_2 | x \in \tau_3) \end{pmatrix}. \quad \text{(B.1)}$$

The matrix $R_{\sigma\to\tau}$ that changes σ to τ is

$$R_{\sigma\to\tau} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}. \quad \text{(B.2)}$$

Since $R_{\tau\to\sigma}R_{\sigma\to\tau} = I_2$, it follows that transitioning from the social classifications τ to σ can be achieved without any loss of information. Interverting the two changes of classes results in an error because transitioning from σ back to τ may not be lossless. Indeed, we have

$$R_{\sigma\to\tau}R_{\tau\to\sigma} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 \end{pmatrix} \neq I_3.$$

Now, suppose that an individual of each specific class σ_i belongs to more than one specific social class of the class τ with a non-zero probability. The matrix $R_{\tau\to\sigma}R_{\sigma\to\tau}$ can be written in the form

$$R_{\tau\to\sigma}R_{\sigma\to\tau} = \begin{pmatrix} p_{1,1} & p_{1,2} \\ p_{2,1} & p_{2,2} \end{pmatrix}, \quad \text{(B.3)}$$

and $\|I_2 - R_{\tau\to\sigma}R_{\sigma\to\tau}\| = \left((1-p_{1,1})^2 + (1-p_{2,2})^2 + p_{1,2}^2 + p_{2,1}^2\right)^{\frac{1}{2}} \neq 0$ since $p_{i,j} \neq 0$ for $1 \leq i, j \leq 2$. Starting from the contact matrix B_σ of the class σ, we observe that a first change followed

by its reversal produces in the following new contact matrix $(R_{\tau\to\sigma}R_{\sigma\to\tau})^T B_\sigma (R_{\tau\to\sigma}R_{\sigma\to\tau})$ of the class σ. Note that the information loss in this change of classes for our simulation may be significant, as highlighted in Table 3.

Scenario	$\|I_n - R_{\tau\to\sigma}R_{\sigma\to\tau}\|_2$	$\|B_\sigma - (R_{\tau\to\sigma}R_{\sigma\to\tau})^T B_\sigma (R_{\tau\to\sigma}R_{\sigma\to\tau})\|_2$
White-collars	0	0
Schools	1.63	0.05
Feeling bad	1.93	0.19

Table 3. Quantification of the information loss for the different proposed representation changes.

Appendix C. List of parameters

This appendix details the parameters used for the simulations.

Parameters	Description	Value	Reference
β	infection rate	0.0275	[15]
κ	matrix of contact	derived from age contact matrix	Section 2.2.1
θ	waning rate	1/360	[24]
δ	latency rate	1/5	[25]
p	ratio of symptomatic	0.52	[32]
γ	recovery rate	1/7	[34]
μ_i	death rate in the class i	computed from IFR by age	[35]