medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Optimal vaccination strategies on networks and in metropolitan areas

Lucas Machado Moschen[1] and María Soledad Aronna[2]

[1]Sorbonne University
[2]School of Applied Mathematics, Fundação Getulio Vargas

April 26, 2024

Abstract

This study presents a mathematical model for optimal vaccination strategies in interconnected metropolitan areas, considering commuting patterns. It is a compartmental model with a vaccination rate for each city, acting as a control function. The commuting patterns are incorporated through a weighted adjacency matrix and a parameter that selects day and night periods. The optimal control problem is formulated to minimize a functional cost that balances the number of hospitalizations and vaccines, including restrictions of a weekly availability cap and an application capacity of vaccines per unit of time. The key findings of this work are bounds for the basic reproduction number, particularly in the case of a metropolitan area, and the study of the optimal control problem. Theoretical analysis and numerical simulations provide insights into disease dynamics and the effectiveness of control measures. The research highlights the importance of prioritizing vaccination in the capital to better control the disease spread, as we depicted in our numerical simulations. This model serves as a tool to improve resource allocation in epidemic control across metropolitan regions.

Keywords: Optimal Control; Epidemiology; Vaccination Protocols; Commuting Patterns; Metropolitan Areas.

1 Introduction

Metropolitan areas consist of a densely populated urban core, such as a capital city, along with its surrounding territories that share social, economic, and infrastructural ties. One defining characteristic of these regions is commuting: the routine movement of individuals between their city of residence and the city of the workplace. For instance, in the Rio de Janeiro Metropolitan Area, over two million people commute daily, in a 13-million inhabitants region [34]. Such mobility creates a complex network of interconnected cities, where the links are weighted by the flux of people. The recent global health outbreaks, such as the Zika virus epidemic and the COVID-19 pandemic, have shown how these dense and interconnected urban areas can amplify disease transmissibility. Therefore, devising strategies that efficiently reduce the impact of outbreaks is essential, and mathematical modeling serves as an indispensable tool to perform this rigorously.

Mathematical Epidemiology has been instrumental in understanding how social interactions and human mobility influence outbreaks. The main tool in this field is compartmental models, which divide the population into groups based on their disease status (*e.g.,* susceptible, infectious, recovered) with differential equations representing the dynamics. They can integrate various real-life factors, including age, spatial effects, etc. The inclusion of spatial effects is particularly relevant as the transmission of infectious diseases often depends on proximity and movement patterns of individuals. The spatial aspect can be approached in either continuous or discrete ways. The latter, known as the *patch* or *metapopulation model*, first divides the population into distinct subpopulations, then further into compartments, resulting in a large set of equations, one for each compartment. Introduced in ecology to study competitive species [19], this approach gained prominence in epidemiology after the work of Rvachev and Longini Jr. (1985), who used the airplane network to predict the spread of influenza, focusing on the 1968 Hong Kong epidemic [31].

Building on this foundational work, Sattenspiel and Dietz (1995) and Arino and Van den Driessche (2003) discussed the integration of mobility between regions into epidemic models, with the latter arguing in favor of space-discrete models and obtaining inequalities for the basic reproduction number $\mathcal{R}_0$ [1, 32]. Extending the approaches into epidemic propagation, Takeuchi et al. (2006) analyzed the impact of transport-related infections on the disease transmission, being followed by Liu et al. [21, 22]. Considering the structure of the graph associated with the city's connectivity, Colizza and Vespignani [9, 10] calculated a global threshold for disease invasion and provided extensive Monte Carlo simulations to verify their findings. They also analyzed the effect of the topology of the graph on the phase diagram in metapopulation models. The works of Pastor-Satorras et al. [27, 28, 29] were also prominent in understanding infectious diseases in networks of individuals.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

We refer the reader to [41] to see the combination of metapopulation and agent-based models for two cities. Stolerman et al. (2015) introduced an SIR-network model to understand the impact of infection rates on $\mathcal{R}_0$, inferring that nodes with higher infection rates are the most important drivers of outbreaks [36].

To effectively tame the epidemic's outbreak, optimal control theory has shown to be useful for theoretical basis and practical solutions [7, 18, 35]. In optimal control formulations, one usually aims at minimizing the number of infections, deaths, or other epidemic-related quantity, at the expense of some control measure, such as quarantine, testing, treatment, and/or vaccination. Several works came in this sense, such as Behncke (2000) [5] who proved that the optimal strategy is to vaccinate as many people as possible as quickly as possible, subject to the resources. See also [13, 24] for other related references. The interplay of optimal control theory with metapopulation models in epidemiology has contributions from Ögren and Martin (2002), Asano et al. (2008), Rowthorn et al. (2009), with increasing traction after the COVID-19 pandemic. The focus of these studies was to improve vaccination strategies, especially in urbanized, mobile populations, and provide solutions tailored to economic constraints and quarantined settings [2, 26, 30].

However, most optimal control applications in epidemiology involved only simple constraints, while the interest in including real-world restrictions has grown. Hansen and Day (2011) notably included resource constraints in their SIR models. These advancements in constrained optimal control problems were studied by Biswas et al. (2014) and De Pinho et al. (2015). The COVID-19 pandemic accelerated these developments, with research like that of Avram et al. (2022) integrating constraints to simulate the capacity of intensive care units. Lemaitre et al. (2022) [17] developed a method that combines distributed direct multiple shooting, automatic differentiation, and large-scale nonlinear programming to optimally allocate COVID-19 vaccines, considering both supply and logistic constraints. When applying their approach to cities in Italy, they used a SEPIAHQRD-V model. They specified that individuals in certain compartments do not commute and accounted for mobility fluxes and infection forces that depend on each region. Another work that emerged to deal with coronavirus was the Robot Dance platform, developed by Nonato et al. (2022). This platform is a computational tool aiding policymakers in curating response strategies tailored to regional nuances, intercity commuting mobility, and hospital capacities, applied to the state of São Paulo, Brazil [25].

In this work, we delve into the intricate dynamics of epidemic spread in metropolitan areas, emphasizing the role of commuting patterns and vaccination strategies. We derive tight bounds for the basic reproduction number, for both a general network of cities and one associated with a metropolitan area. Utilizing optimal control theory, we devise efficient strategies for disease control in metropolitan areas, which we illustrate through a series of numerical experiments that corroborate our findings.

1.1 Contributions

In this work, we study a mathematical model that combines commuting and vaccination in a constrained optimal control problem to manage the propagation of infectious diseases within metropolitan areas, providing a more realistic representation of disease transmission and an efficient strategy to control it. Our main contributions are three-fold:

- Analysis of a mathematical model for epidemics that considers commuting dynamics in metropolitan areas and derivation of upper and lower bounds for the basic reproduction number.
- Formulation of a constrained control-affine optimization problem for finding optimal vaccination strategies in a network of cities.
- Extensive numerical simulations, illustrating various situations and validating the theoretical results.

2 Methods

The formulation of the vaccination model and the associated optimal control problem is composed of three elements: (I) the epidemiological model that combines a compartmental SIR model and commuting patterns to describe the transmission dynamics; (II) an objective function to be minimized, which contemplates the number of infections and applied vaccines; (III) a set of constraints on the state and control variables that ensure resource limitations for the production, distribution, and application of the vaccines.

2.1 Epidemiological modeling

Consider a directed graph $G = (V, E)$, where the set of nodes V corresponds to K interconnected cities, and the set of ordered pairs E indicates connectivity between cities. Each link (i, j) has a weight p_{ij} denoting the fraction of the population of city i that commutes daily to the city j. This forms a *weighted adjacency matrix* $[P]_{ij} = p_{ij}$, with P being a right stochastic matrix, *e.g.* the elements of each row sum up 1. Each city i has a

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

total population n_i divided into three compartments representing the proportion of susceptible (S_i), infectious (I_i), and recovered (R_i) individuals, satisfying:

$$S_i(t) + I_i(t) + R_i(t) = 1,$$

for all $t \geq 0$ and all $i = 1, \ldots, K$.

In this model, we incorporate a birth rate μ and a natural death rate μ for each compartment in each city. The rates are chosen such that the total population of each city remains constant over time. We also introduce α, a daily periodic function, such that $\alpha(t) = 0$ during daytime and $\alpha(t) = 1$ at night. For instance, a realistic setting for α could be $\alpha(t) = 0$ for $t \in (k + 1/4, k + 3/4)$, and $\alpha(t) = 1$ for $t \notin (k + 1/4, k + 3/4)$, for each day $k = 0, 1, \ldots, T$. During nighttime, the evolution of the system is not influenced by mobility, leading to the SIR model:

$$\begin{aligned}
\frac{dS_i}{dt}(t) &= \mu - \beta_i S_i I_i - \mu S_i, \\
\frac{dI_i}{dt}(t) &= \beta_i S_i I_i - \gamma I_i - \mu I_i, \\
\frac{dR_i}{dt}(t) &= \gamma I_i - \mu R_i,
\end{aligned}$$

where β_i represents city i's infection rate, which is directly affected by the population density and the probability of infection given contact, and γ^{-1} refers to the average infection period.

During the day, commuting changes the infection dynamics. We model this feature using the same approach as Nonato et al. (2022) [25]. The *effective* population in city i during working hours is given by

$$P_i^{\text{eff}} = \sum_{j=1}^{K} p_{ji} n_j,$$

which sums over the parcels of the populations that travel from any city j to i. The proportion of susceptible workers from city i that get exposed to the infectious individuals in city j is $p_{ij} S_i I_j^{\text{eff}}$, where I_j^{eff} is the effective proportion of infectious people in city j during the day and it is defined by

$$I_j^{\text{eff}} = \frac{1}{P_j^{\text{eff}}} \sum_{k=1}^{K} p_{kj} I_k n_k.$$

The complete model is

$$\begin{aligned}
\frac{dS_i}{dt}(t) &= \mu - \alpha(t)\beta_i S_i I_i - (1 - \alpha(t)) S_i \sum_{j=1}^{K} \beta_j p_{ij} I_j^{\text{eff}} - \mu S_i, \\
\frac{dI_i}{dt}(t) &= \alpha(t)\beta_i S_i I_i + (1 - \alpha(t)) S_i \sum_{j=1}^{K} \beta_j p_{ij} I_j^{\text{eff}} - \gamma I_i - \mu I_i, \\
\frac{dR_i}{dt}(t) &= \gamma I_i - \mu R_i,
\end{aligned} \tag{1}$$

for $i = 1, \ldots, K$, subject to the initial conditions $S_i(0) = S_{i0}, I_i(0) = I_{i0}$ and $R_i(0) = R_{i0}$, for $i = 1, \ldots, k$. We use the following vector notations $\boldsymbol{S}(t) = (S_1(t), \ldots, S_K(t)) \in \mathbb{R}^K$, $\boldsymbol{I}(t) = (I_1(t), \ldots, I_K(t)) \in \mathbb{R}^K$ and $\boldsymbol{R}(t) = (R_1(t), \ldots, R_K(t)) \in \mathbb{R}^K$. Figure 1 represents this model graphically.

Remark 2.1. In this study, we do not consider disease-specific mortality. This decision is based on two observations: firstly, including an additional mortality rate due to the disease does not change the disease dynamics. Secondly, our primary focus lies in analyzing the impact of vaccination strategies during the initial months of the epidemic, a period where the relative impact of disease-specific mortality on the outcomes is usually minimal. Our main objective is to control the rate of new infections, which is not influenced by the mortality rate.

It can be proved that

$$\mathcal{C} := \{X \in \mathbb{R}_{\geq 0}^{K \times 3} : X_{i1} + X_{i2} + X_{i3} = 1, \text{ for } i = 1, \ldots, K\}$$

is invariant under system (1). The complete statement and its proof are in Appendix B.

2.2 The basic reproduction number

Following the method developed by Van den Driessche and Watmough [39], we derive the basic reproduction number using the formula $\mathcal{R}_0 = \rho(FV^{-1})$. Here F and V are matrices associated with the Jacobian of the

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

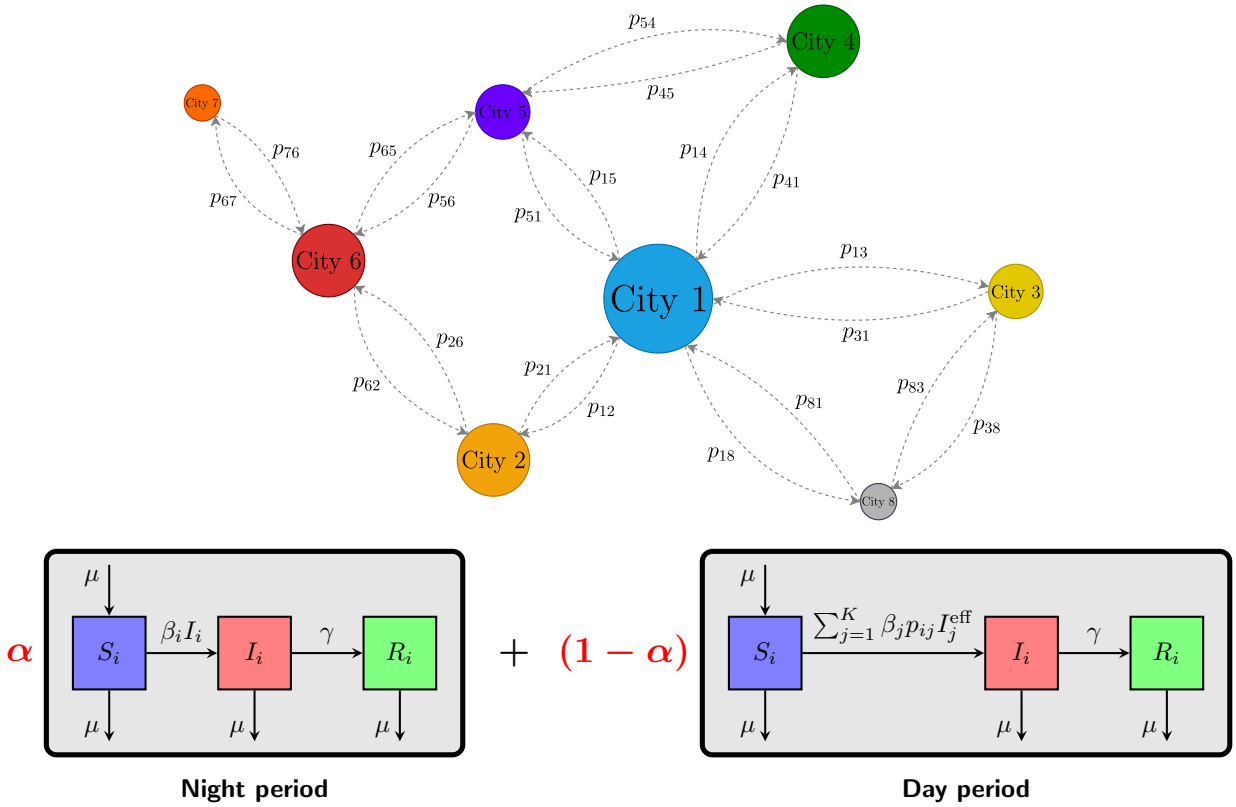


Figure 1: **Graphical representation of city interactions and the SIR model.**
The cities are connected composing a network, p_{ij} being the proportion of individuals from city i who work during the day at j. At night, cities do not interact, and the epidemic follows a standard SIR dynamic. During the day, some individuals commute, interact with infected individuals in the destination city, and may transport the infection.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

dynamics at the disease-free equilibrium (DFE). The detailed definitions are included in the Appendix A and the derivation is given in Subsection A.1. Given the problem's dimension, we could not obtain an explicit expression for $\mathcal{R}_0$. However, we obtained in Theorem 2.1, Corollary 2.1 and Theorem 2.2 below, sharp inequalities for $\mathcal{R}_0$ in terms of simple expressions involving the parameters.

Let $\mathcal{R}_0^i$ be the basic reproduction number of city i if the cities were isolated. This is given by the well-known formula

$$\mathcal{R}_0^i = \frac{\beta_i}{\gamma + \mu}. \tag{2}$$

We get the following result, the proof of which is in Appendix B .

Theorem 2.1. *The basic reproduction number $\mathcal{R}_0$ for system (1) satisfies the following inequalities*

$$\min_{1\le i\le K} \alpha\mathcal{R}_0^i + (1-\alpha)\sum_{k=1}^{K} p_{ik}\mathcal{R}_0^k \le \mathcal{R}_0 \le \max_{1\le i\le K} \alpha\mathcal{R}_0^i + (1-\alpha)\sum_{k=1}^{K} p_{ik}\mathcal{R}_0^k. \tag{3}$$

The inequality in (3) can be written as

$$\min_{1\le i\le K} w_i \le \mathcal{R}_0 \le \max_{1\le i\le K} w_i, \qquad \text{where } w := \frac{(\alpha I + (1-\alpha)P)\boldsymbol{\beta}}{\gamma+\mu} \in \mathbb{R}^K,$$

in which $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_K)$. The matrix $\alpha I + (1-\alpha)P$ is right stochastic and balances the static state, where people are in their residence cities, with probability α, and the transition matrix P, that represents commuting, with probability $1-\alpha$. Therefore, the entry w_i represents the expected basic reproduction number for city i, calculated as the average of $\mathcal{R}_0^j$ values for all cities j, with the weights determined by the probability of individuals from city i work in city j.

Corollary 2.1. *The following inequality holds:*

$$\min_{1\le i\le K} \mathcal{R}_0^i \le \mathcal{R}_0 \le \max_{1\le i\le K} \mathcal{R}_0^i.$$

In particular, if we consider, without loss of generality, that $\beta_1 = \max_{1\le i\le K} \beta_i$, we have that $\mathcal{R}_0 \le \mathcal{R}_0^1$.

The proof is in Appendix B.

From now on, unless indicated, we make the following assumption, which summarizes the idea that commuting majorly occurs between the capital and the other cities as illustrated by Figure 2.

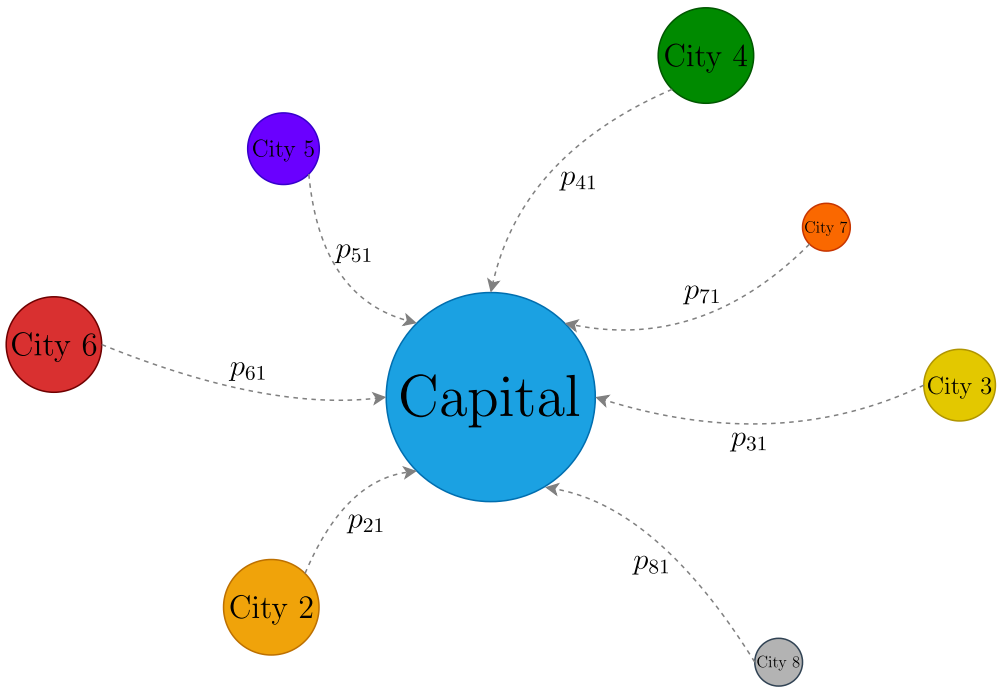


Figure 2: **Version of Figure 1 considering a metropolitan area and Assumption 2.1.**

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Assumption 2.1 (Metropolitan area structure). In a metropolitan area, there is a capital, which is a big city with a higher population density and a larger number of inhabitants, along with other connected cities. We assume that individuals residing in the capital city both stay and work there, while people in other cities either commute to the capital or work in their home city. We further enumerate the cities according to their populations in decreasing order of population size. Consequently, matrix P reduces to the form:

$$P = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ p_{21} & p_{22} & 0 & \cdots & 0 \\ p_{31} & 0 & p_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{K1} & 0 & 0 & \cdots & p_{KK} \end{bmatrix}.$$

The above assumption adjusts to the reality of several regions. For instance, Figure 3 shows the matrix P of the Rio de Janeiro metropolitan area through a heatmap. A similar mobility structure is present in the Buenos Aires metropolitan region [15]. Other patterns of mobility may occur in different regions. For instance, in Île-de-France, Paris metropolitan area, a portion of individuals from the capital work in the surroundings [11].

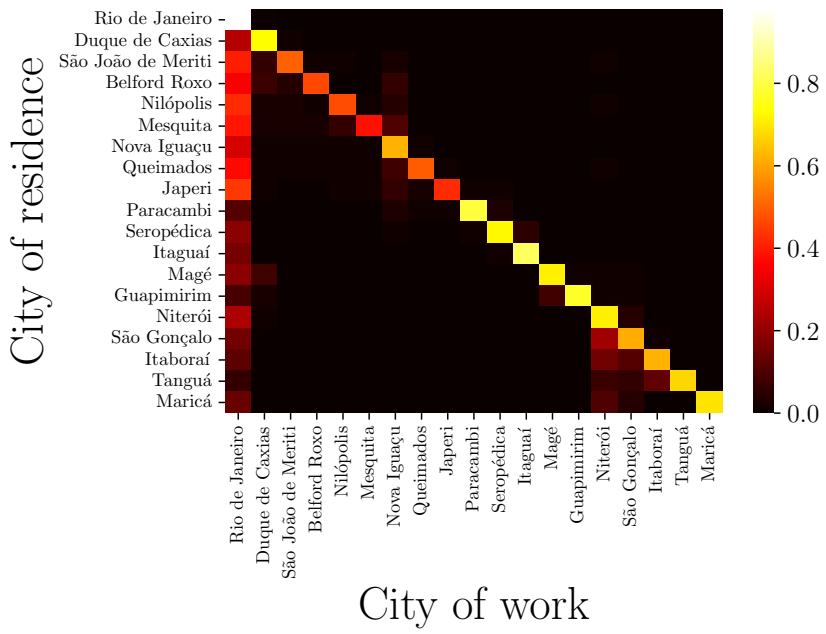


Figure 3: **Transition matrix of the Rio de Janeiro metropolitan area.**
A heatmap showing the matrix P in the case of the Rio de Janeiro metropolitan area (mobility data taken from [34]).

When applying the specific structure of Assumption 2.1 to the dynamics described in equation (1), we derive a different inequality for the reproductive number $\mathcal{R}_0$ (the proof is in Appendix B).

Theorem 2.2 (Inequalities for $\mathcal{R}_0$ under Assumption 2.1)**.** *The basic reproduction number $\mathcal{R}_0$ for system (1) under Assumption 2.1, satisfies the following inequalities*

$$\frac{\xi}{\gamma+\mu} \leq \mathcal{R}_0 \leq \frac{1}{\gamma+\mu}\left(\xi + (1-\alpha)\frac{\beta_1}{P_1^{\text{eff}}}\sum_{i=1}^{K} n_i p_{i1}^2\right), \tag{4}$$

where $\xi = \max_{1\leq i\leq K} \alpha\beta_i + (1-\alpha)p_{ii}\beta_i(1-\delta_{i1})$, where δ_{ij} is the Kronecker function.

Under Assumption 2.1, calculating $\mathcal{R}_0$ reduces to a problem of finding the spectral radius of a diagonal plus a rank-one matrix. As far as we know, there is no closed-form expression for it, and obtaining tighter bounds is an open question in the field of *Matrix Perturbation Theory* [20].

Remark 2.2. Higher values of γ correspond to faster recovery rates and lower values of $\mathcal{R}_0$. As α balances day and night periods, when it approaches 1, the night period becomes more relevant, and $\mathcal{R}_0$ tends to the maximum of $\mathcal{R}_0^i$, for $i = 1, \ldots, K$, meaning that infections in the residence city governed the dynamics. The parameters β_i directly amplify $\mathcal{R}_0$. Through numerical experiments, we observed that the highest value among the β_i parameters is the most significant driver (see *e.g.* Figure 7). Finally, although the individual population sizes n_i do not directly influence eigenvalue behaviors, the ratios n_i/n_1 offer a subdued effect, predominantly guided by $(1-\alpha)\max_{1\leq i\leq K}\beta_i$.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

2.3 Vaccination and optimal control problem

We implement vaccination as a control policy by introducing a time-dependent vaccination rate u_i in each city i. We consider that susceptible individuals can receive the vaccine and immediately move to compartment R_i. Once vaccinated, these individuals no longer contribute to the spread of the disease during the period under consideration. We could also suppose other sets of hypotheses or model formulations, as we did in [23]. The introduction of vaccination leads to a modified SIR epidemiological model, to which we add a vaccine counter V_i, resulting in the following set of equations:

$$\begin{aligned}
\frac{dS_i}{dt} &= \mu - \alpha\beta_i S_i I_i - (1-\alpha) S_i \sum_{j=1}^{K} \beta_j p_{ij} I_j^{\text{eff}} - u_i S_i - \mu S_i, \\
\frac{dI_i}{dt} &= \alpha\beta_i S_i I_i + (1-\alpha) S_i \sum_{j=1}^{K} \beta_j p_{ij} I_j^{\text{eff}} - \gamma I_i - \mu I_i, \\
\frac{dR_i}{dt} &= u_i S_i + \gamma I_i - \mu R_i, \\
\frac{dV_i}{dt} &= u_i S_i.
\end{aligned} \tag{5}$$

This framework assumes that infected individuals are recognizable and do not receive the vaccine. This assumption can be challenging in real-life scenarios, but we believe that vaccinating infectious and/or recovered individuals would not change the qualitative results of this work. In Appendix D we discuss a variant of this model, considering vaccination in the workplace.

For the control system (5), it can be shown that the set

$$\mathcal{C} = \{X \in \mathbb{R}_{\geq 0}^{K\times 4} : X_{i1} + X_{i2} + X_{i3} = 1, \text{ for } i = 1, \dots, K\}$$

is positively invariant under the flow of system (5), for each measurable function u. For details, see Appendix B.

Following the steps of Subsection A.1, we can calculate the basic reproduction number $\mathcal{R}_0^{\text{vac}}$ for the model (5) with constant vaccination. Although $\mathcal{R}_0^{\text{vac}}$ does not have a closed-form expression, the vaccination rate inversely impacts $\mathcal{R}_0^{\text{vac}}$. The inequality presented in equation (3) of Theorem 2.1 can be adapted for this context by replacing $\mathcal{R}_0^i$ with a modified version $\mathcal{R}_0^{\text{vac},i}$, where

$$\mathcal{R}_0^{\text{vac},i} = \frac{\mu}{\mu + u_i} \mathcal{R}_0^i = \frac{\mu}{\mu + u_i} \frac{\beta_i}{\gamma + \mu}.$$

Building upon the formulation of the dynamics, we develop an optimal control problem in which we vaccinate the population of the metropolitan area while considering the cost of vaccination and the cost of hospitalization of infected individuals. The number of applied doses in city i until time t is

$$n_i \int_0^t u_i(t) S_i(t)\, dt = n_i V_i(t).$$

We suppose that a rate r_h of infected individuals are hospitalized with a daily unity cost of c_h, and we assume that the unity cost of vaccination is c_v. Therefore, we define the cost functional

$$J[u_1, \dots, u_K] := \sum_{i=1}^{K} c_v n_i \int_0^T u_i(t) S_i(t)\, dt + c_h r_h n_i \int_0^T I_i(t)\, dt. \tag{6}$$

Due to limitations of resources of the health system, a limited number of people can be vaccinated at each time t, so we impose the *(mixed control-state)* constraint on the vaccination rate

$$u_i(t) S_i(t) \leq v_i^{\max}, \quad \text{for } t \in [0, T] \text{ and } i = 1, \dots, K, \tag{7}$$

which implies the following constraint on the daily cap

$$\int_d^{d+1} u_i(t) S_i(t)\, dt \leq v_i^{\max}, \qquad \text{for } d = 0, \dots, T-1.$$

Through the latter constraint, we indirectly ensure that a proportion of at most q individuals can be immunized by setting

$$\sum_{i=1}^{K} n_i v_i^{\max} \leq q \sum_{i=1}^{K} n_i,$$

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

for some $q \in (0,1)$.

We consider a scenario where cities receive weekly vaccine shipments from a centralized authority. We use $V_w^{\max}$ to denote the number of vaccines to be received at week $w = 0, \ldots, T/7 - 1$. From this, we define the function of cumulative vaccine shipments by

$$D(t) := \begin{cases} \sum_{n=0}^{w-1} V_n^{\max} + V_w^{\max}(t - 7w), & t \in [7w, 7w+1], \\ \sum_{n=0}^{w} V_n^{\max}, & t \in [7w+1, 7(w+1)], \end{cases}$$

for $w = 0, 1, \ldots, T/7 - 1$ and T a multiple of 7 for simplicity. The function D is piecewise linear non-decreasing. Note that, to account for the delay in delivering vaccines across cities and healthcare centers, D increases linearly on the first day of each week, after which it remains constant for the rest of the week. This induces the following constraint:

$$\sum_{i=1}^{K} n_i \int_0^t u_i(t) S_i(t)\, dt = \sum_{i=1}^{K} n_i V_i(t) \leq D(t). \tag{8}$$

Observe that any unused vaccines can be employed over the subsequent weeks. Figure 4 illustrates an example of this function.

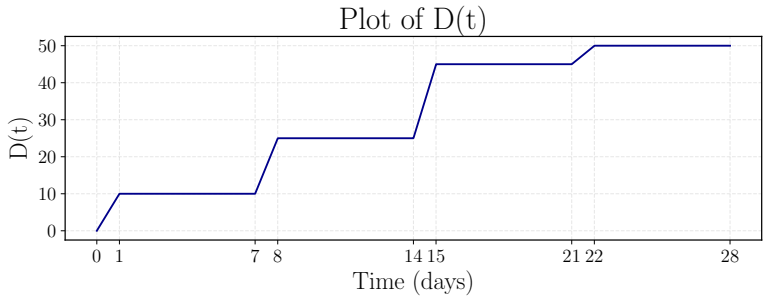


Figure 4: **Example of the function of delivered vaccines.**

Aggregating the dynamics in (5), the cost function in (6), the constraints in (7), (8) and $u_i \geq 0$, we get the following optimal control problem with u taken in the set of measurable functions from $[0,T]$ to $\mathbb{R}^K$,

$$\begin{aligned} \min \quad & \sum_{i=1}^{K} c_v n_i \int_0^T u_i(t) S_i(t)\, dt + c_h r_h n_i \int_0^T I_i(t)\, dt, \\ \text{s.t.} \quad & \dot{S}_i = \mu - \alpha\beta_i S_i I_i - (1-\alpha) S_i \sum_{j=1}^{K} \beta_j p_{ij} I_j^{\text{eff}} - u_i S_i - \mu S_i \\ & \dot{I}_i = \alpha\beta_i S_i I_i + (1-\alpha) S_i \sum_{j=1}^{K} \beta_j p_{ij} I_j^{\text{eff}} - \gamma I_i - \mu I_i \\ & \dot{R}_i = u_i S_i + \gamma I_i - \mu R_i \\ & \dot{V}_i = u_i S_i \\ & \sum_{i=1}^{K} n_i V_i(t) \leq D(t), \text{ a.e. } t \in [0,T] \\ & u_i(t) S_i(t) \leq v_i^{\max}, \text{ a.e. } t \in [0,T] \\ & u_i(t) \geq 0, \text{ a.e. } t \in [0,T] \\ & S_i(0) = s_{i0},\ I_i(0) = i_{i0},\ R_i(0) = r_{i0},\ V_i(0) = 0. \end{aligned} \tag{9}$$

Remark 2.3. In most of the studies involving vaccination and optimal control, a quadratic cost on the control is considered. That assumption guarantees the existence of a solution, stability of numerical methods, and direct application of Pontryagin's Maximum Principle to get a feedback expression for the optimal control. Nevertheless, we believe that a linear cost in this context is a more realistic representation of operational costs, since it computes the total amount of vaccines or hospitalizations, for instance. In particular, the cost of vaccination should be linear or concave with respect to the quantity of vaccines; a strongly convex assumption does not accurately mirror practical aspects.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

3 Results

The core objective of our study is to analyze the effects of different vaccination strategies on the spread of infectious diseases in a metropolitan region modeled by model (5). We have carried out a series of experiments that provide a practical platform to evaluate the efficacy of these strategies under controlled conditions. We start by presenting simulations that delineate the behavior of the basic reproduction number. This is followed by exploring the impact of constant vaccination rates on the disease dynamics. We then optimize vaccination strategies, by solving the optimal control problem (9). Lastly, we assess the performance of a practical feedback solution, which we develop by considering a modified problem.

For solving differential equations, we use the explicit Runge-Kutta method of order 5(4) ('rk45'), using the SciPy Python library [40]. For the optimal control solution, we adopt a *first discretize, then optimize* approach facilitated by the Gekko Python library [4], which uses IPOPT for non-linear optimization. All experiments were performed on a Linux PC equipped with an AMD Ryzen 9 5950X processor (16 cores) and 128 GB of memory. The computer code to reproduce the experiments and implement the proposed models is available under a license at `https://github.com/lucasmoschen/network-controllability`.

3.1 Simulations for the basic reproduction number

Before we introduce vaccination, we examine the impact of the parameters on the basic reproduction number $\mathcal{R}_0$, as described in Methods and the bounds for $\mathcal{R}_0$ provided by the aforementioned Theorems 2.1 and 2.2. For all the simulations throughout this article, we fix γ to $1/7$, which means that the recovery time from infection is on average 7 days, and $\mu = 3.6 \cdot 10^{-5}$ to represent approximately the average life expectancy of 75 years.

To open the discussion, we analyze the simpler situation of two cities, where the parameter p_{21}, the proportion of individuals that reside in city 2 and work in city 1, uniquely determines the transition matrix P. Let us set $n_2 = 1$ since only the ratio n_i/n_1 is relevant to $\mathcal{R}_0$, not the magnitude. Figure 5 presents the dynamics of the epidemic in two different scenarios for β, considering a single infected initial individual in the population of the first city. Notice that higher values of $\mathcal{R}_0$ give higher and faster epidemic peaks.

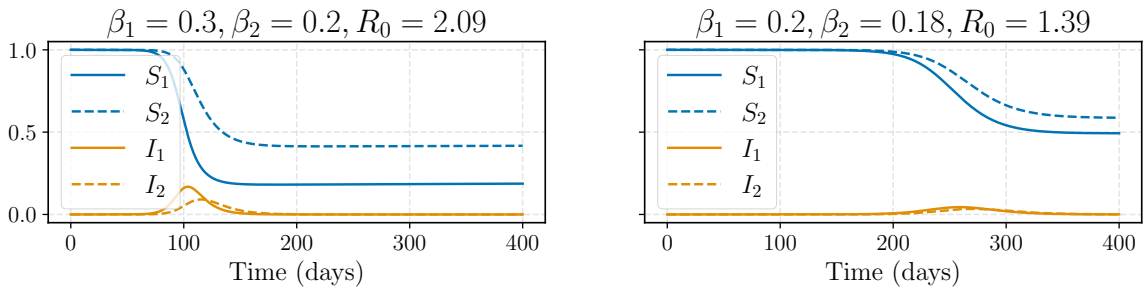


Figure 5: **Dynamics of the Epidemic in Two Cities.**
Epidemic dynamics in two cities under two different scenarios. The settings are $\alpha = 0.64, p_{21} = 0.2$ and $n_1 = 10$. The proportion of susceptible individuals is plotted in blue, and the proportion of infectious individuals in orange. The larger city is represented by a solid line, while the smaller city is represented by a dashed line.

In Figure 6 we show the level curves of $\mathcal{R}_0$ when p_{21}, n_1/n_2 and α vary. For the chosen parameters β, γ and μ, we have $\mathcal{R}_0^1 > 1$ and $\mathcal{R}_0^2 < 1$, which represents a situation for which Corollary 2.1 cannot determine whether $\mathcal{R}_0$ is greater or smaller than 1. Looking at Figure 6, we notice that higher values of α increase the value of $\mathcal{R}_0$. The impact of p_{21} is nonlinear and complex, but its behavior is similar to a quadratic function in $\mathcal{R}_0$. On the other hand, the impact of p_{21} and α is small when the population proportion n_1/n_2 increases. Additionally, see the histogram of Figure C.21, which illustrates the situation where $\mathcal{R}_0^1$ is close to 1.

In Figure 7, we analyze the bounds for $\mathcal{R}_0$ provided by Theorem 2.1 and Theorem 2.2. We mainly observe that $\mathcal{R}_0$ is primarily driven by β_1, the highest effective transmission rate, while α, p_{21}, and β_2 have little impact. Higher values of α improve the bounds of Theorem 2.2, depicted in orange, mainly due to the term factor of $(1-\alpha)$ appearing on the right-hand side of (4). However, the fourth graph (where β_2 varies) illustrates the fact that $\rho(\alpha B + (1-\alpha)\Lambda)$ remains constant until $\alpha\beta_2 + (1-\alpha)\beta_2 p_{22} < \alpha\beta_1$, beyond which the upper bound (in orange dashed line) gets worse.

In the context of multiple cities, we consider a scenario of $K = 5$ cities. We first examine five distinct mobility scenarios: (I) metropolitan area where 10% of the residents of each city work in the capital; (II) metropolitan area where 40% of each city works in the capital; (III) all cities are interconnected, where 60% of the residents work in their city, while the remaining are distributed evenly across the other four cities; (IV) a variant of the metropolitan area where 30% of each city's residents work in the capital and 10% of the capital's workforce is employed in the other cities; (V) no commuting between cities.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

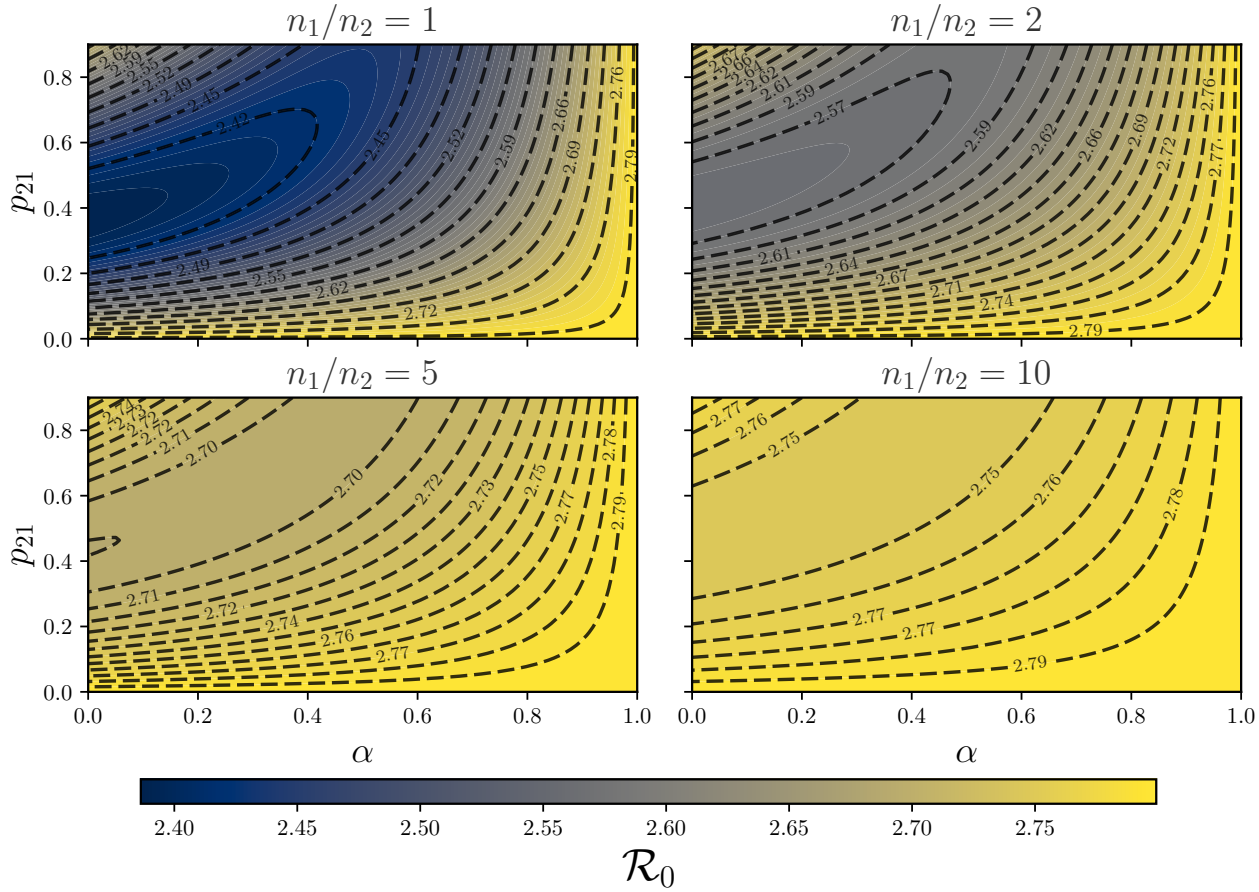


Figure 6: **Basic reproduction number as function of α and p_{21} for two cities.**
Contour plot of the values of $\mathcal{R}_0$ for different values of $\alpha \in (0,1)$, $p_{21} \in (0, 0.9)$ and $n_1/n_2 = (1, 2, 5, 10)$ considering $\beta = (0.4, 0.1)$.

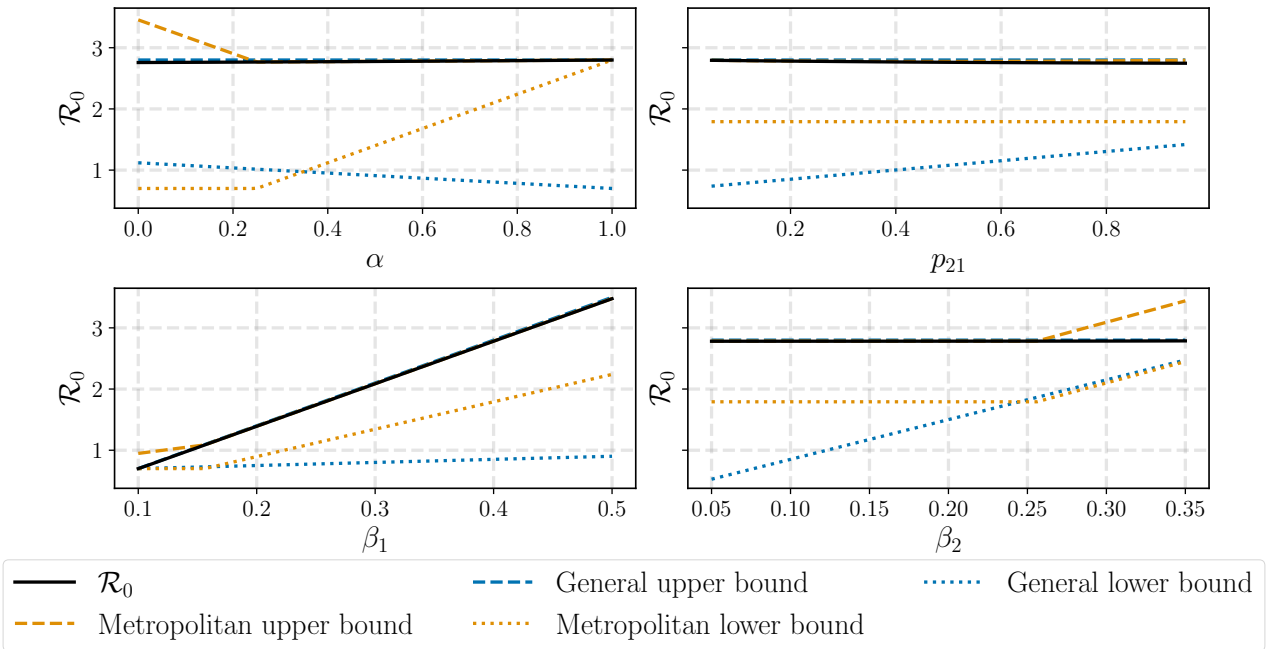


Figure 7: **Impact of the parameters on the bounds of $\mathcal{R}_0$.**
Using the same setting as in Figure 5, we visualize the impact of α, p_{21}, β_1 and β_2 on $\mathcal{R}_0$ (depicted in black). In orange, we see the bounds from Theorem 2.2, and in blue, the bounds from Theorem 2.1. The dotted lines are the lower bounds while the dashed lines are the upper bounds.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Our findings, summarized in Table C.2 and illustrated in Figure 8, reveal that in the capital city – compared to the other cities – the peak size is higher, the peak day occurs earlier, the outbreak (period between $t = 0$ and the time at which the proportion of infectious individuals falls below 10^{-5}) lasts less and the attack rate is larger. Additionally, higher commuting rates correlate with a shorter epidemic span, as scenarios (II) and (III) exhibit and the capital's peak day tends to dictate the overall trajectory of the disease spread across the network. Further, the outcomes of scenarios (II) and (III) are similar, which indicates that depending on the structure of the network, we may approximate it using the metropolitan hypothesis in Assumption 2.1.

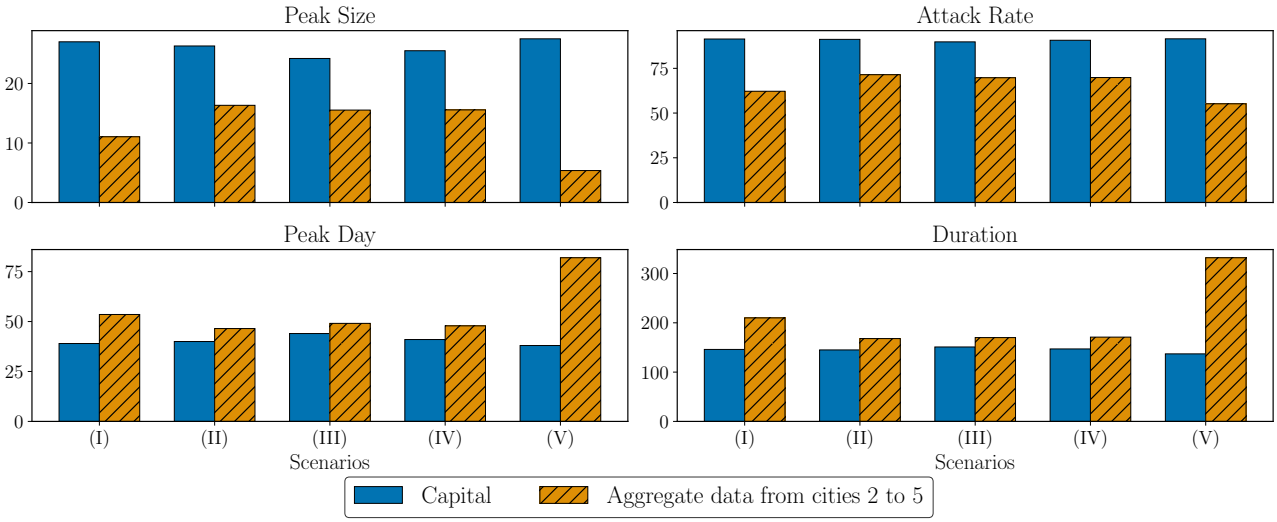


Figure 8: **Simulation results for different structures of the transition matrix.** Comparison of Peak Size (maximum proportion of infectious individuals), Peak Day (day of the peak size), Duration (time from epidemic onset to the day when the proportion of infectious individuals achieves 10^{-5}), and Attack Rate (proportion of individuals who contract the disease during 350 days of the epidemic) across different mobility scenarios. The aggregated data is the weighted average, considering their population size, from cities 2 to 5. For this experiment, $\beta = (0.4, 0.25, 0.2, 0.15, 0.1)$, $\alpha = 0.64$, and the population sizes are $10^5 \cdot (50, 10, 10, 1, 1)$. Lastly, the initial conditions are $I_1(0) = I_2(0) = I_3(0) = 10^{-4}$ and no recovered individuals.

In Figure 9, we vary the infection rate β. We consider an experiment of four different scenarios in five cities and depict only three cities to better exhibit the results. When comparing scenarios (I) and (II), both have $\mathcal{R}_0$ greater than 1. However, a minor adjustment in β_1 from 0.2 to 0.15 results in a nearly 100-fold decrease in peak size, changing the entire aspect of the epidemic. Interestingly, we observed that even if $\mathcal{R}_0^i < 1$, for $i = 2, \ldots, 5$, the epidemic still manages to spread in these cities. Turning our attention to scenario (III), this setting represents a situation where $\mathcal{R}_0 < 1$, achieved by modifying the city with the highest transmissibility rate: from the capital to the city 3, the second largest city. This implies that the capital, being the main hub of mobility, plays a significant role in accelerating the spread of the disease. Finally, scenario (IV) fixes the same infection rate for all cities. Notice in this case that the behavior of the epidemic is very similar for all cities and mobility does not play a big role in the capital.

3.2 Constant vaccination rates

We now turn our attention to the analysis of the vaccination policy considering a constant rate u_i for each city i, as outlined by the equations in (5).

In this scenario, we consider an ongoing epidemic that is subject to a vaccination strategy to control it, starting on a specified day s. In the two-city scenario, we set $\beta = (0.5, 0.3)$ to ensure that $\mathcal{R}_0^1 > 1$ and $\mathcal{R}_0^2 > 1$. The vaccination rates are varied within the range $u = (u_1, u_2) \in [10^{-6}, 10^{-3}]^2$ to generate Figure 10, which is presented on the log-log scale. For this setting, the basic reproduction number without vaccination is $\mathcal{R}_0 \approx 3.48$. Numerically, by a fitting method, we observe that

$$\mathcal{R}_0^{\text{vac}}(u_1, u_2) \approx \frac{\mathcal{R}_0}{1 + u_1 \cdot b(u_2)}, \tag{10}$$

where b is approximately a constant function of u_2, as we observe in Figure 10. By this figure and by equation (10), we infer that the vaccination rate of the capital city governs $\mathcal{R}_0^{\text{vac}}$.

To analyze the significance of vaccinating the capital, we perform a simplified version of the optimal control problem introduced in equation (9) only considering the constant vaccination strategies. We use a constant

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

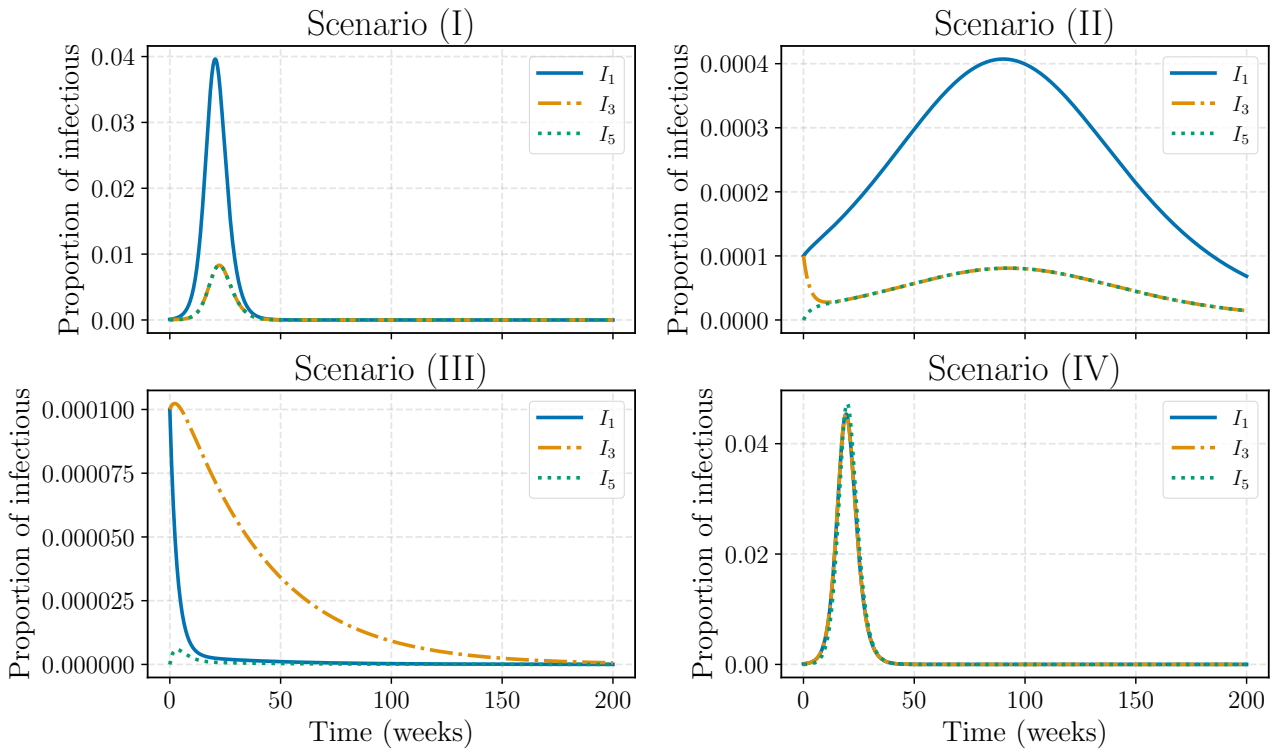


Figure 9: **Infectious curves for different scenarios of β.**
Curves of infectious individuals from cities 1, 3, and 5 under four scenarios of the β specification. Here $\alpha = 0.64, p_{i1} = 0.2$ for $i \geq 2$ and $n_1 = 5n_2 = 5n_3 = 50n_4 = 50n_5 = 5 \cdot 10^6$. The four scenarios considered are (I) $\beta = (0.2, 0.1, 0.1, 0.1, 0.1)$, indicating $\mathcal{R}_0^1 > 1$ and $\mathcal{R}_0^i < 1$ for $i \neq 1$; (II) $\beta = (0.15, 0.1, 0.1, 0.1, 0.1)$, a situation similar to (I), but with a slower transmissibility rate, even though $\mathcal{R}_0 > 1$; (III) $\beta = (0.1, 0.1, 0.15, 0.1, 0.1)$, a situation similar to (II), but with the highest transmissibility rate not in the capital. This scenario results in $\mathcal{R}_0 < 1$; (IV) $\beta = (0.2, 0.2, 0.2, 0.2, 0.2)$, where all cities have the same transmissibility rate.

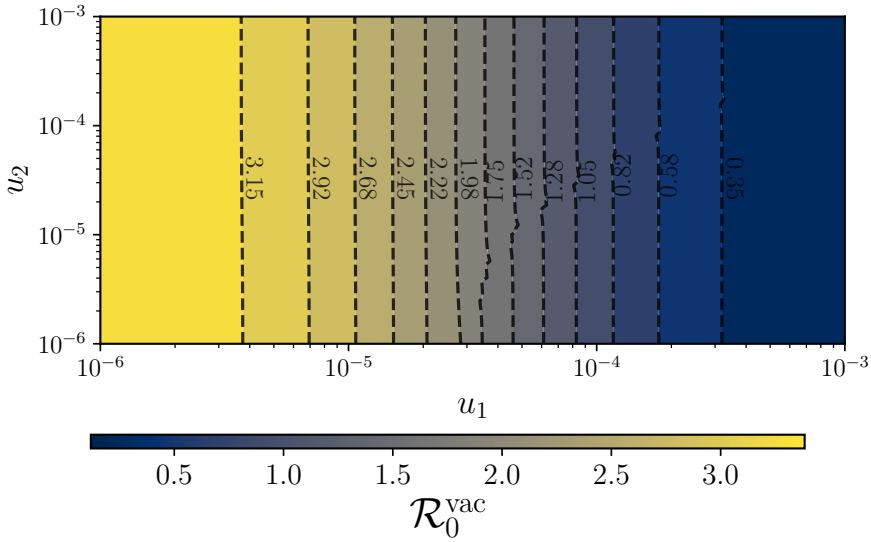


Figure 10: **Basic reproduction number $\mathcal{R}_0^{\text{vac}}$ in the vaccination model.**
The basic reproduction number $\mathcal{R}_0^{\text{vac}}$ for the model described in (5) as a function of u_1 and u_2 across $K = 2$ cities. The settings are: $\beta = (0.5, 0.3)$, $\alpha = 0.64$, $p_{21} = 0.2$ and the ratio $n_1/n_2 = 10$. Both axes u_1 and u_2 are in the logarithmic scale.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

vector $u \in [0, 0.05]^K$ as the decision variable. In this constant vaccination case, we can simplify the constraints and consider the following:

$$\sum_{k=1}^{K} V_k(T) n_k \leq 0.8 \sum_{k=1}^{K} n_k$$

to ensure that the total number of vaccinated individuals at the final time T is at most 80% of the population. We set $T = 8 \times 7 = 56$ days, the hospitalization rate $r_h = 0.1$ and the hospitalization cost $c_h = 1000$. Vaccination starts at the 70th day after the first infection in the capital. Finally, we consider three different scenarios for the parameter β and three specifications of the unity cost of a vaccine c_v. As we can see in Table 1, it is evident that in all scenarios, the capital city not only receives the majority of vaccines but also achieves the highest vaccination rate, particularly reaching the upper bound in $\beta^{(2)}$ and $\beta^{(3)}$ scenarios. In the first scenario, which is characterized by a uniform infection rate across the cities, vaccine distribution is more equitable, yet the capital still maintains an advantage. The behavior is similar when u_j is allowed to vary up to 0.1 for each city j. However, it is important to note that the relevance of vaccinating the capital can diminish in scenarios where the susceptible population decreases rapidly, as vaccine allocation depends significantly on the size of the available population.

β	c_v	Capital	City 2	City 3	City 4	City 5
$\beta^{(1)}$	c_v^1	$\mathbf{4.26 \cdot 10^{-2}}$	$3.41 \cdot 10^{-2}$	$3.41 \cdot 10^{-2}$	$3.41 \cdot 10^{-2}$	$3.41 \cdot 10^{-2}$
	c_v^2	$\mathbf{4.26 \cdot 10^{-2}}$	$3.41 \cdot 10^{-2}$	$3.41 \cdot 10^{-2}$	$3.41 \cdot 10^{-2}$	$3.41 \cdot 10^{-2}$
	c_v^3	$\mathbf{4.26 \cdot 10^{-2}}$	$3.41 \cdot 10^{-2}$	$3.41 \cdot 10^{-2}$	$3.41 \cdot 10^{-2}$	$3.41 \cdot 10^{-2}$
$\beta^{(2)}$	c_v^1	$\mathbf{5.00 \cdot 10^{-2}}$	$\mathbf{5.00 \cdot 10^{-2}}$	$9.63 \cdot 10^{-3}$	$9.63 \cdot 10^{-3}$	$3.42 \cdot 10^{-8}$
	c_v^2	$\mathbf{5.00 \cdot 10^{-2}}$	$\mathbf{5.00 \cdot 10^{-2}}$	$9.63 \cdot 10^{-3}$	$9.63 \cdot 10^{-3}$	$3.02 \cdot 10^{-7}$
	c_v^3	$\mathbf{5.00 \cdot 10^{-2}}$	$1.38 \cdot 10^{-2}$	$1.65 \cdot 10^{-8}$	$1.70 \cdot 10^{-7}$	$1.62 \cdot 10^{-7}$
$\beta^{(3)}$	c_v^1	$\mathbf{5.00 \cdot 10^{-2}}$	$\mathbf{5.00 \cdot 10^{-2}}$	$1.06 \cdot 10^{-2}$	$1.54 \cdot 10^{-7}$	$5.01 \cdot 10^{-8}$
	c_v^2	$\mathbf{5.00 \cdot 10^{-2}}$	$\mathbf{5.00 \cdot 10^{-2}}$	$1.06 \cdot 10^{-2}$	$1.76 \cdot 10^{-7}$	$3.36 \cdot 10^{-8}$
	c_v^3	$\mathbf{5.00 \cdot 10^{-2}}$	$7.65 \cdot 10^{-9}$	$7.65 \cdot 10^{-9}$	$7.91 \cdot 10^{-8}$	$6.20 \cdot 10^{-8}$

Table 1: **Optimal constant vaccination rates across cities for different parameter settings:** vaccination rates in five different cities under nine parameter settings combining three values for β and three for c_v. We set $\beta^{(1)} = (0.3, 0.3, 0.3, 0.3, 0.3)$; $\beta^{(2)} = (0.3, 0.2, 0.12, 0.12, 0.1)$; and $\beta^{(3)} = (0.3, 0.12, 0.12, 0.1, 0.05)$, and $c_v^1 = 0.001$; $c_v^2 = 0.1$; and $c_v^3 = 10$. Each row corresponds to a different combination of β and c_v values, representing different scenarios of disease transmission rate and vaccination cost. For each row, in bold, we highlight the highest rate.

As seen in Table 1, when the first day of vaccination is determined by the number of infectious individuals, for instance, the day on which 1% of the total population is infected, all scenarios indicate that the capital should receive vaccines at the highest rate, regardless of c_v. The same result is obtained when the initial susceptible population is uniform across the cities. This analysis showcases that it is preferable to vaccinate the capital at a higher rate.

3.3 Performance of time-variable vaccination strategies

In this section, we perform numerical simulations that incorporate vaccination as a time-variable control function. By integrating vaccination strategies into our metropolitan area model, the goal is to provide a comprehensive understanding of disease dynamics and control measures. We start by examining the interaction between two cities, as shown in Figure 11. This figure illustrates the optimal trajectory and control for a capital city with a population 10 times greater than the second city and a higher infection rate. Vaccination begins on the 100th day following the arrival of the first infected individual in the capital.

Notably, during the initial two weeks, the optimal solution allocates all vaccines to the capital, administering them at the maximum possible rate according to the weekly cap. Subsequently, as the susceptible population in the capital decreases to around 60%, it becomes advantageous to start vaccinating the second city, which has a lower daily cap. The fourth graph illustrates the *bang-bang* [33] behavior of the *effective vaccination rate* $u_i S_i$, this is, in which $u_i^*(t) S_i^*(t) \in \{0, v_i^{\max}\}$ holds for $i = 1, 2$, along the whole interval $[0, T]$. The gap of the weekly cap constraint is depicted in Figure 12. We assign a high unity cost to hospitalization $c_h = 1000$ in contrast to a relatively low unity cost for vaccination, with $c_v = 0.01$ to reflect the burden of hospitalization. Both healthcare resources and patient well-being are comparatively way more expensive than the minor inconvenience and price of vaccination. Furthermore, the benefits of vaccination are amplified when administered on a large scale. This simulation took around 33 seconds to run.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

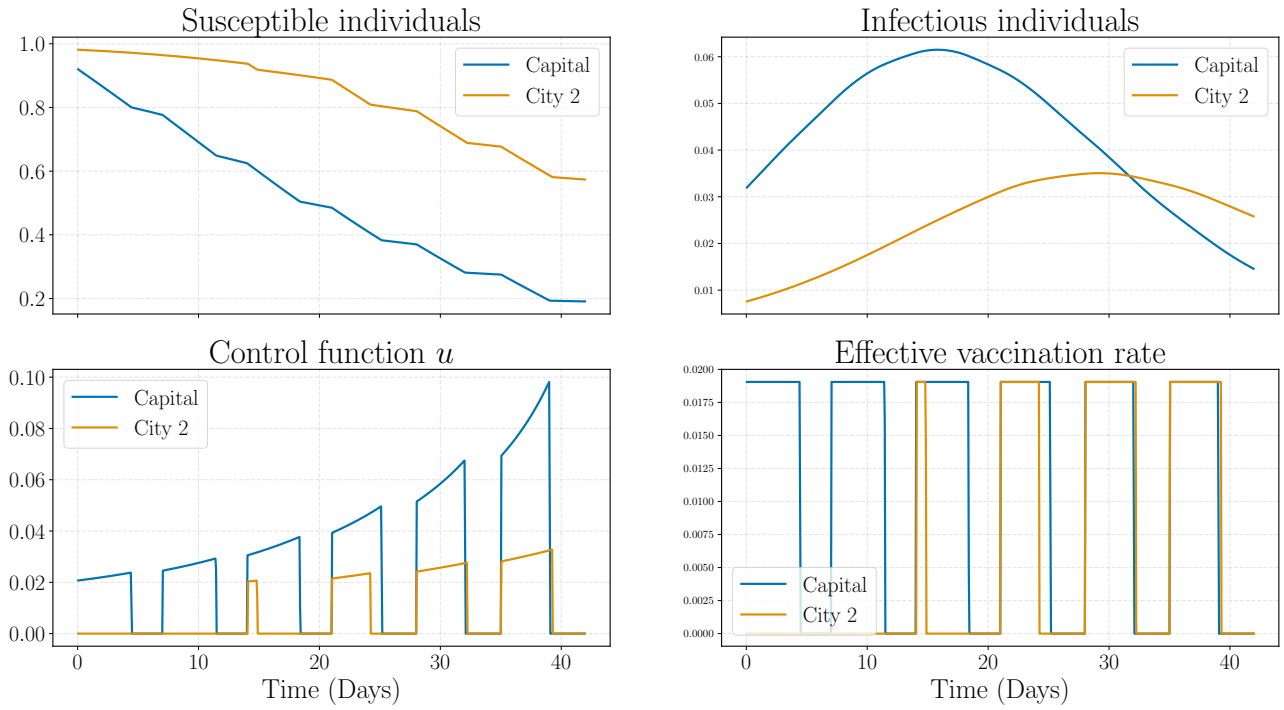


Figure 11: **Optimal trajectories and control for two-city interaction.**
Optimal trajectory of the proportion of susceptible and infectious individuals in two cities, the capital in orange and the second city in blue. The fourth subplot illustrates the variables $u_i(t)S_i(t) \in \{0, v_i^{\max}\}$, $i = 1, 2$. The settings for this experiment are: $\beta = (0.25, 0.18)$, $\alpha = 0.64$, $p_{21} = 0.2$, $n_1 = 10^6$ and $n_2 = 10^5$. For the optimal control problem, $c_v = 0.01$, $c_h = 1000$, $v_1^{\max} = v_2^{\max} = 0.8/42$ and a weekly cap allowing the vaccination of at most $1/13$ of the population per week.

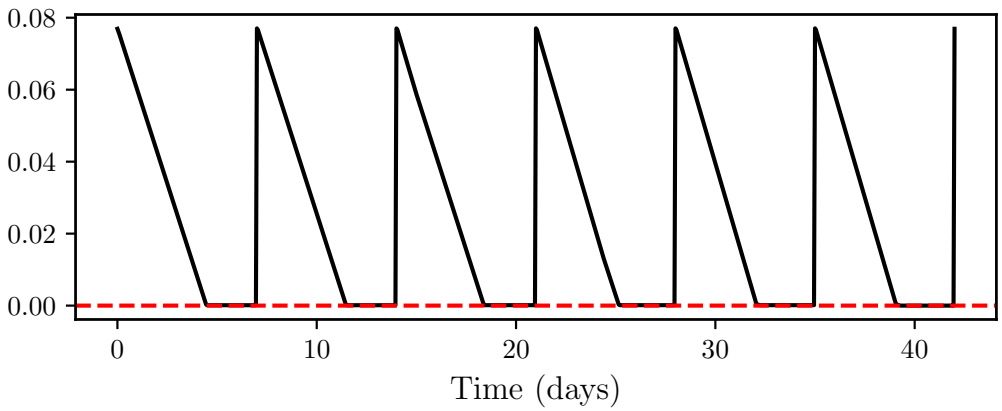


Figure 12: **Weekly constraint analysis.**
Function $D(t) - \sum_{k=1}^{K} n_k V_k(t)$ for the experiment presented in Figure 11.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

When we equalize the infection rates of the two cities, setting $\beta_1 = \beta_2 = 0.25$, while keeping the other parameters fixed, we obtain the solution in Figure 13. One can see that the behavior of the solution is similar to that of the previous experiment but with a uniform optimal solution across both cities. This suggests that the transmission rate β is the most important driver when deciding vaccination strategies. Our prior conclusion about the importance of the susceptible population in deciding which city retains the vaccination is also evident in the figure.

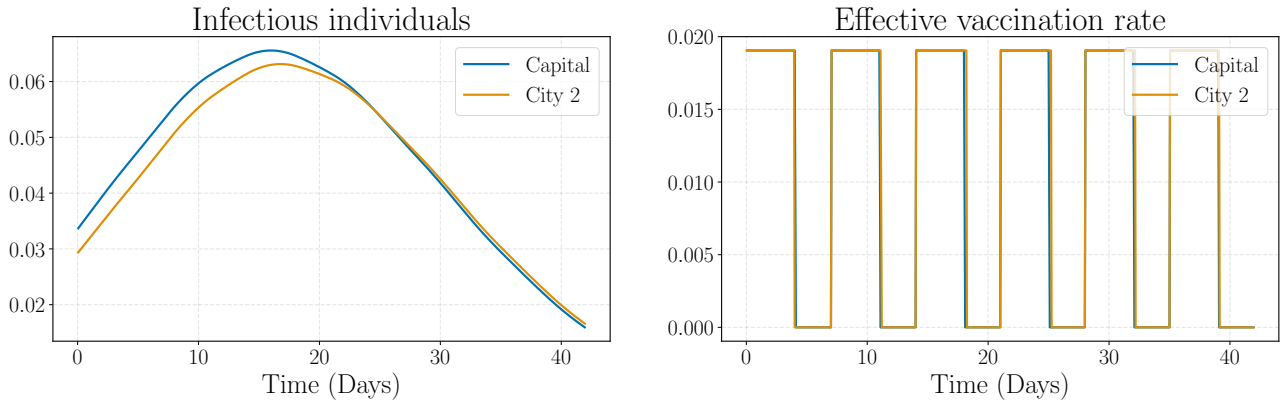


Figure 13: **Optimal trajectories and control for two-city interaction for uniform infection rates.** Optimal trajectory of the proportion of susceptible and infectious individuals in two cities and the optimal control using the same settings as Figure 11 except $\beta = (0.25, 0.25)$.

Finally, we consider four different scenarios: first infection in the second city (Scenario (I)), uniform initial conditions for the control problem (Scenario (II)), and two that express higher transmissibility in the capital, but a lower vaccination rate limitation (Scenarios (III) and (IV)). The result is shown in Figure 14. It establishes that the vaccination always starts in the capital in all scenarios, even with the infection starting in the second city. Moreover, Scenarios (III) and (IV) confirm that this behavior does not come from the maximum effective vaccination rate, highlighting that even though we can vaccinate more rapidly in the second city, the capital starts first.

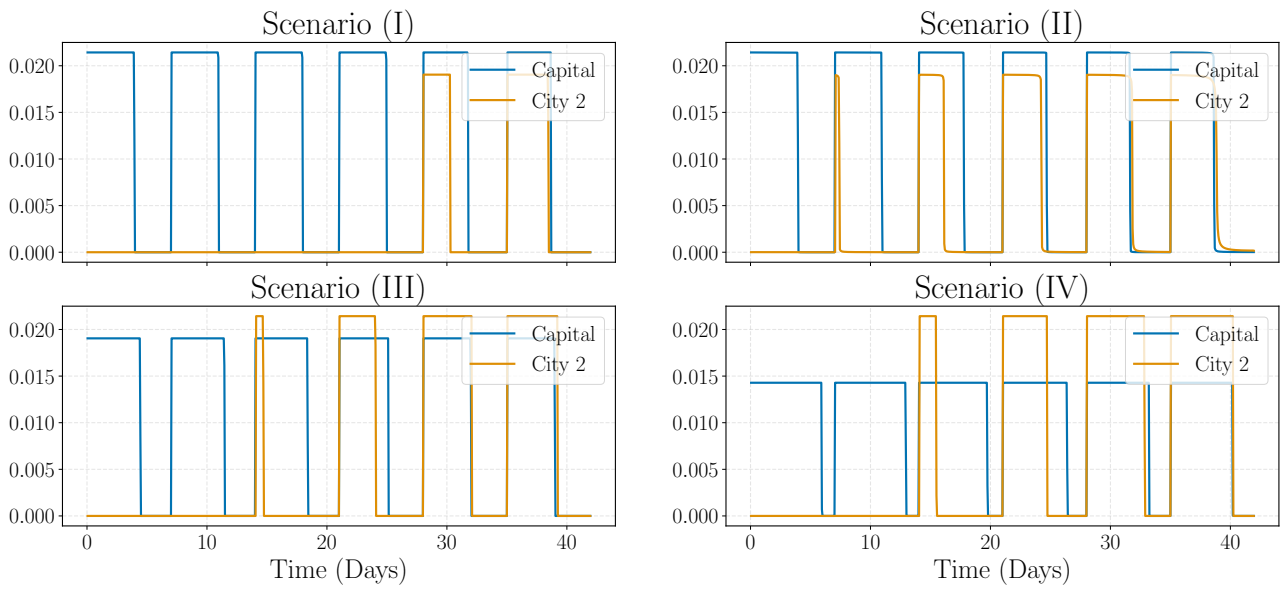


Figure 14: **Optimal effective vaccination rate for two-city interaction.**
Scenario (I) considers the same setting as Figure 11 except for the first infection happening in the second city and $v_1^{\max} = 0.9/42$ (while $v_2^{\max} = 0.8/42$). Scenario (II) considers $I_1(0) = I_2(0) = 0.05$ and $R_1(0) = R_2(0) = 0.02$. Scenarios (III) and (IV) fix the same parameters of Figure 11 but with different maximum rates $v_1^{\max}$ and $v_2^{\max}$. In (III), we set $v_1^{\max} = 0.8/42 < v_2^{\max} = 0.9/42$, while in (IV), $v_1^{\max} = 0.6/42 < v_2^{\max} = 0.9/42$.

Having considered the simplified two-city case, we extend our analysis to more cities in the metropolitan area. We first observe that the dimensionality of the problem escalates quickly, which makes efficient implementations a challenging aspect yet to be addressed in the literature. Lemaitre et al. (2022) [17] consider a similar optimal

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

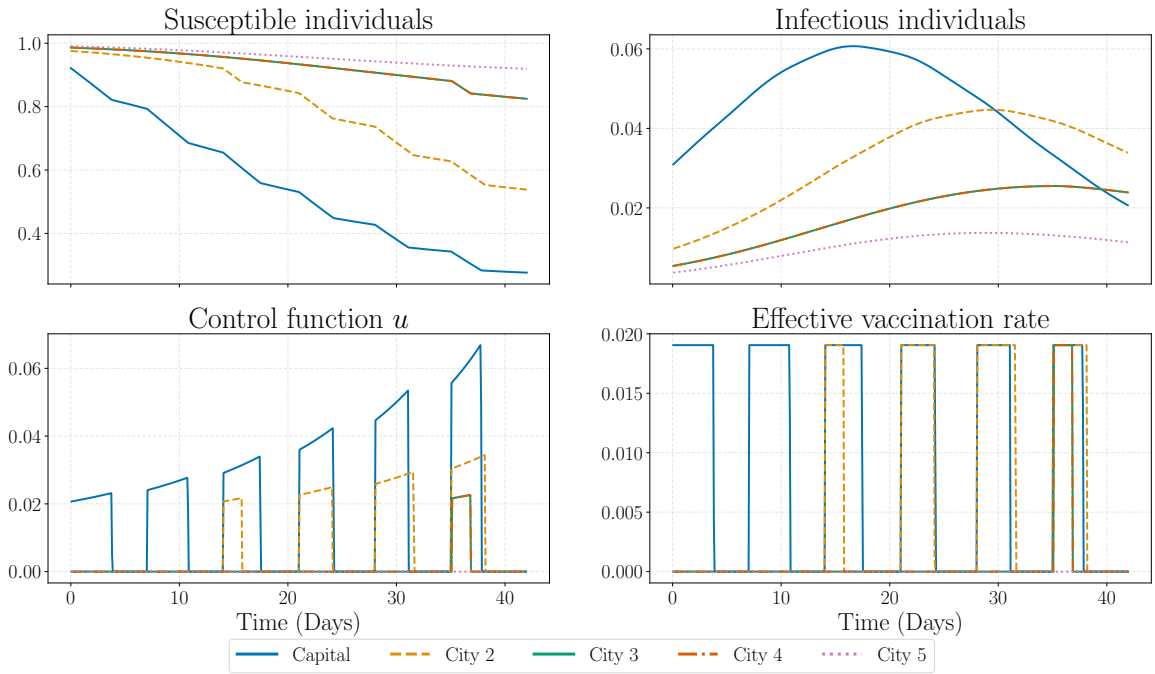


Figure 15: **Optimal trajectories and control functions for 5 cities.**
Optimal trajectories and controls for a problem involving 5 cities. The settings mirror those in Figure C.22, with the following specifications: $c_v = 0.01$, $c_h = 1000$, and $r_h = 0.1$. The model allows for a weekly vaccination of up to 1/20 of the susceptible population. The maximum rate of vaccinated individuals is set at $v_1^{\max} = \cdots = v_5^{\max} = 0.8/42$.

control problem and offer a viable solution, which we could further explore in future research.

For a simulation with $K = 5$ cities, refer to Figure 15. The results contained in the image are parallel to those in the two-city scenario in which the vaccines are preferentially allocated to the capital in contrast to other cities, vaccinating as fast as possible saturating the constraints. During the fourth week, when the susceptible population is reduced to less than 60% in the capital, the vaccination expands to the second city, which has the second highest transmission rate. The behavior of the control function aligns again with the *bang-bang* type of solution. In total, the experiment required approximately 3 minutes and 31 iterations.

In addition to the analysis of the optimal control problem in multi-city interaction, we introduce a comparative visualization depicted in Figure 16. This figure presents a dual-bar plot comparison between the impact of two distinct vaccination strategies on the total number of infections across varying transmission rates of the capital city. The transmission rate of the capital was chosen as a comparison factor since it drives most of the epidemic. The bars in blue correspond to the application of the optimal control strategy as the solution of problem (9). In orange, we consider a constant vaccination approach, where the vaccination rate is uniform in different cities and over time. This comparison highlights that following the optimal strategy leads to fewer people having the disease during the epidemic. Additionally, the higher the transmission rate in the capital, the greater the impact of the optimal solution.

3.4 Performance of feedback practical solution

We bring to attention the fact that the optimal solution is hard to implement in real-world applications since the exact number of susceptible, infected, and recovered individuals is difficult to estimate. Because of this, we propose a numerical simulation that substitutes the constraint $u_i(t)S_i(t) \leq v_i^{\max}$ by

$$u_i(t)(1 - V_i(t)) \leq v_i^{\max}$$

which also leads to a bang-bang type of solution where we get $u_i^*(t)(1 - V_i^*(t)) \in \{0, v_i^{\max}\}$ for $t \in [0, T]$. If $\mu = 0$ (which is a reasonable approximation for short horizons), this solution is an admissible control to the original problem since

$$\frac{d}{dt}(V_i + S_i + I_i) = \mu R_i - \gamma I_i \leq 0,$$

which implies

$$V_i(t) + S_i(t) \leq V_i(t) + S_i(t) + I_i(t) \leq V_i(0) + S_i(0) + I_i(0) \leq 1$$

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

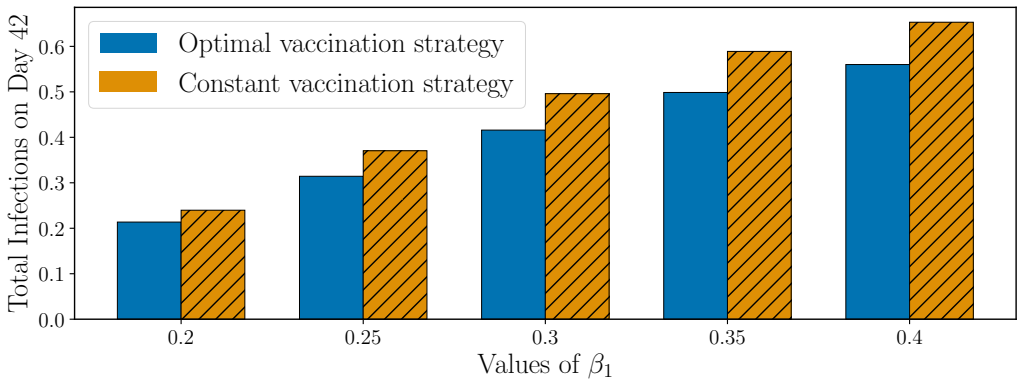


Figure 16: **Comparison between optimal and constant vaccination strategies for different values of β_1.**
Comparison among the total number of infections in the metropolitan area, which is calculated by integrating the new infections and taking the mean over the cities, for two scenarios. The first, in blue, solves the optimal control problem (9), and the second, in orange, chooses the best constant vaccination rate across all cities.

and, therefore,

$$u_i(t)(1 - V_i(t)) \geq u_i(t)S_i(t).$$

Moreover, it is possible to apply in a real-world scenario, because $V_i(t)$ is naturally known by the Health System. Figure 17 shows that this solution yields very similar results regarding the optimal trajectories since the curves of susceptible and infectious individuals end very closely. The difference appears in the optimal vaccination policy since fewer vaccines can be administrated daily, but these are compensated by vaccinating during more days in the week.

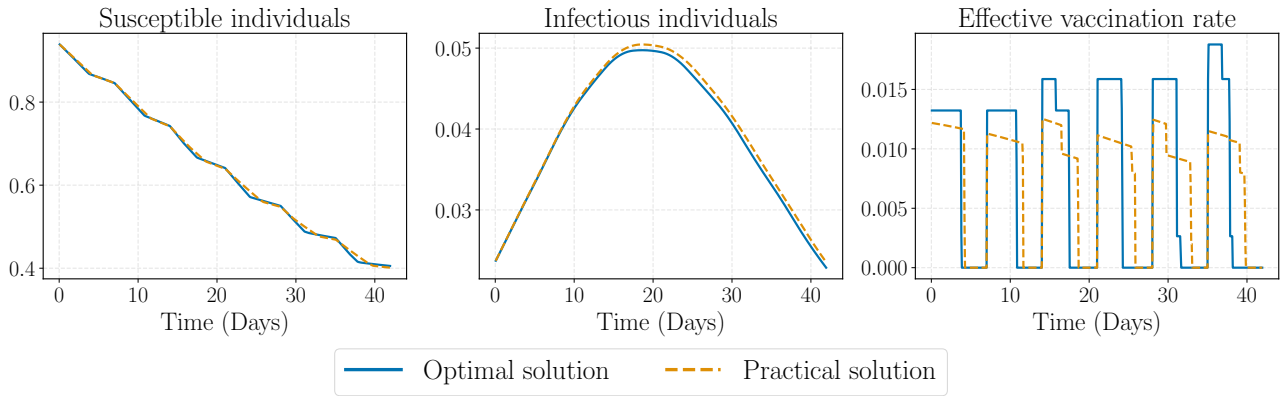


Figure 17: **Comparison between optimal and practical solution in the metropolitan area.**
Optimal trajectories and effective vaccination rate for the metropolitan area considering the settings of Figure 15. It compares the optimal solution with a more practicable solution, which is only based on the knowledge of vaccinated individuals. We observe that the effective vaccination rate here is the daily proportion of vaccinated individuals in the metropolitan area as a whole.

3.5 Evaluating Vaccination Strategies in the Rio de Janeiro Metropolitan Area

The last section of the numerical experiments considers real data on commuting patterns and population dynamics within the Rio de Janeiro metropolitan area [34]. We examine the mobility matrix P, as depicted in Figure 3. Despite the similarity of this matrix and the metropolitan structure outlined in Assumption 2.1, the differences warrant further analysis. Figure 18 shows a comparative study of the optimal vaccination strategies derived from both the actual matrix P and its approximation considering the metropolitan hypothesis of Assumption 2.1. This comparison reveals that the metropolitan structure approximation of the matrix P yields an optimal solution that is closely similar to the one obtained using the original matrix.

Remark 3.1. The parameter β is proportional to the average contact rate of a random person. It is generally an increasing function of the population density. Given the heterogeneous density in Rio de Janeiro's metropolitan area, which includes large green spaces like the Tijuca Forest in the capital city, it would not be enough only to

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

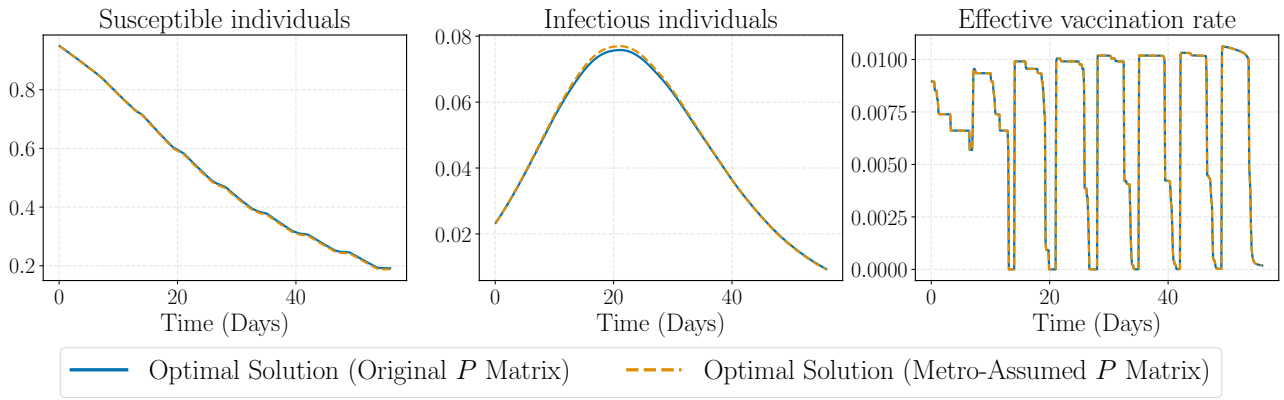


Figure 18: **Comparative Analysis of Optimal Vaccination Strategies in the Rio de Janeiro Metropolitan Area.**
Optimal vaccination strategies for Rio de Janeiro, through both the original and approximated P matrices. The settings are $\alpha = 0.64$, a maximum vaccination rate of $0.6/56$ per city, and weekly vaccine shipments for vaccinating up to $1/20$ of the total susceptible population. The transmission rate β is randomly chosen between 0 and 0.3, and sorted by city population size. Vaccination starts when the susceptible population falls below 95% following the initial infection in the capital. The approximation of P in a metropolitan context is calculated by normalizing p_{i1} and p_{ii} by its sum.

consider β as proportional to the average density of each city. For this reason, we consider random values for β and sort them by population size.

Vaccination in the capital and the other cities is compared in Figure 19. This analysis tracks on a daily basis the number of cities that vaccinate while the capital is not, i.e., $u_1(t) = 0$ and some $u_i(t) \neq 0$ for $i \geq 2$. We notice that most of the time, either the capital is vaccinating or none of the cities are. The first deviation from this fact happens when the susceptible population is less than 60%, which corroborates our findings on the behavior of the optimal control function.

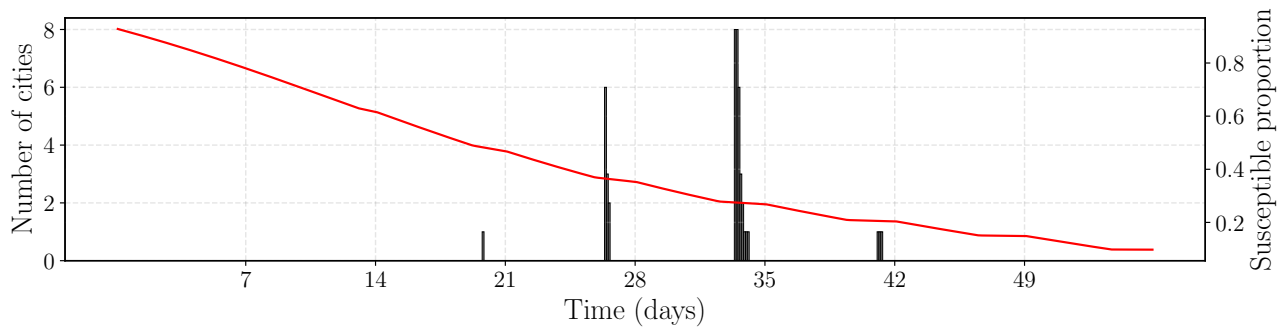


Figure 19: **Analysis of Effective Vaccination Rates in the Rio de Janeiro Metropolitan Area.**
The red graph represents the proportion of susceptible individuals in Rio de Janeiro city, while the gray bars show the number of cities vaccinating when the capital is not. The settings of this experiment are identical to those in Figure 18.

We conclude this section by presenting the evolution of the disease across the Rio de Janeiro metropolitan area in Figure 20. This figure compares three scenarios: the optimal solution of problem (9), a uniform vaccination strategy across all cities and over time, and a scenario with no vaccination. We verify that the optimal solution induces a lower incidence of infections over the period. Moreover, it brings attention to the fact that smaller cities, such as Maricá, exhibit a higher proportion of infections under the optimal solution (23% over 56 days) compared to the uniform vaccination strategy (18%), due to the weighting of population sizes in the cost function (6).

4 Discussion

Our research has led to significant advances in comprehending how epidemics behave within a metropolitan area, especially in relation to the most effective vaccination strategies. The key achievement of our study is the theoretical analysis of a mathematical model that combines commuting patterns with constrained optimal

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

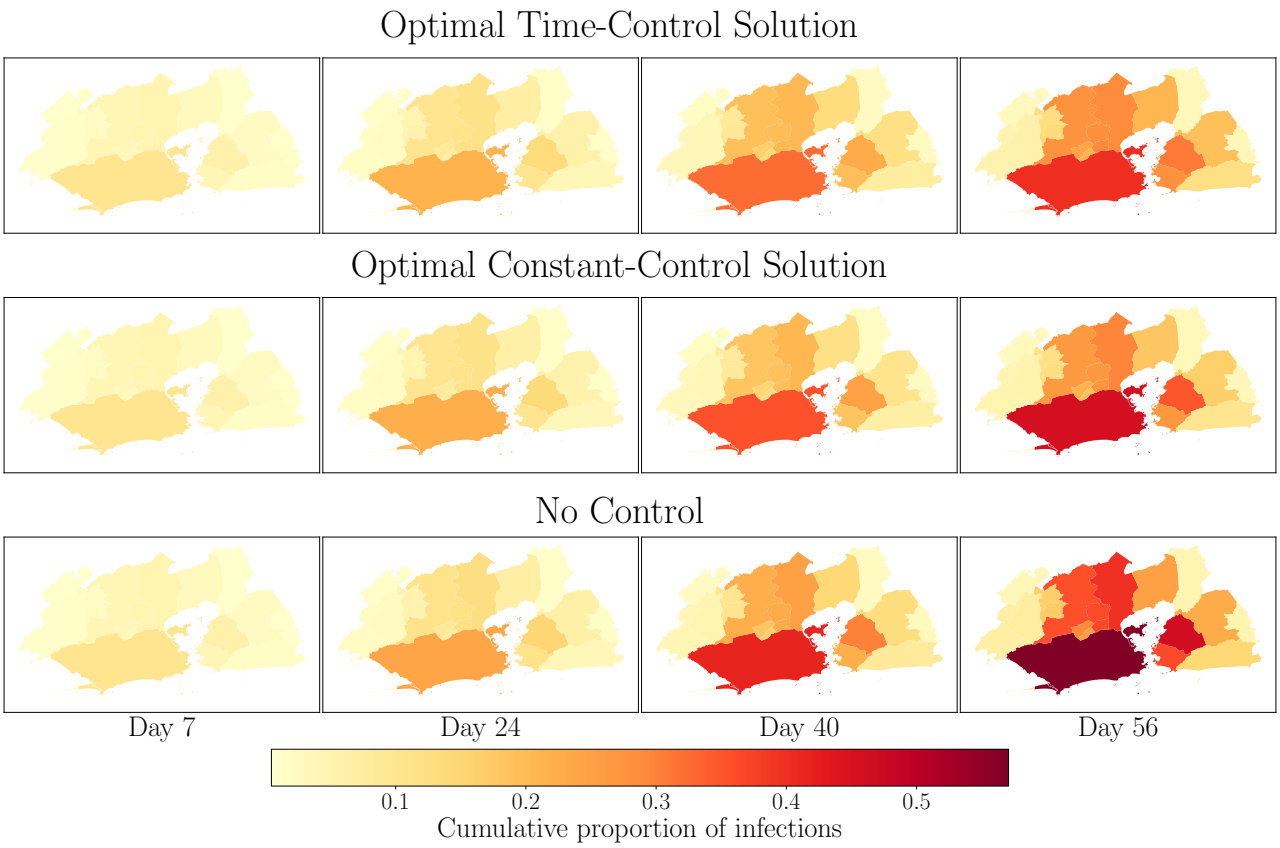


Figure 20: **Disease Progression in the Rio de Janeiro Metropolitan Area under Various Vaccination Scenarios.**
Cumulative proportion of infections via a heatmap, with the framework established in Figure 18. Data sourced from Brazilian Institute of Geography and Statistics (IBGE).

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

control to mimic and manage the propagation of infectious diseases. The study here presented aims to offer a realistic representation of disease transmission within densely populated urban regions.

As mentioned in the text, the application of optimal control in epidemiological models allows us to determine the most effective strategies for mitigating the impact of disease outbreaks while satisfying resource restrictions. By incorporating commuting patterns into the model, we could capture the complex dynamics of disease spread in metropolitan areas. Ignoring them may lead to inaccurate predictions and ineffective control strategies. Moreover, a strategy that does not consider several cities in its plan is doomed to failure due to mobility.

The literature review highlights the importance of integrating spatial heterogeneity and human/social behavior into epidemiological models. It also evidences the need for more realistic and detailed models tailored to specific research questions.

One of the key findings of this work regards the upper and lower bounds for the basic reproduction number, $\mathcal{R}_0$, in particular for the case of a metropolitan region. However, a closed-form expression for it is still an open question since calculating the spectral radius of a diagonal plus one-rank matrix remains an open question in linear algebra. We showed that $\mathcal{R}_0$ is close to $\mathcal{R}_0^1$ and most influenced by the infection rate in the capital city, β_1. Thus, the upper bound may serve as an estimate for $\mathcal{R}_0$. Efficient numerical methods for calculating $\mathcal{R}_0$ were discussed, but they are also open research in the field of perturbation theory.

The optimal control model possesses mixed control-state and pure-state constraints, in addition to control-affine dynamics and cost functional. This complex structure leads to many theoretical and numerical challenges, such as the occurrence of *singular arcs*, the presence of a multiplier that belongs to the space of measures, and others. We are developing the theoretical aspects related to the optimal control model (9) in another article (some results are already present in [23])

Our numerical simulations have provided valuable insights into the dynamics of the disease and the effectiveness of control measures. In particular, they showed that higher vaccination rates in the capital can significantly reduce the number of infections and the overall impact of the epidemic in the whole metropolitan area. This underscores the importance of targeted vaccination strategies in controlling disease outbreaks. We also highlighted how the infection rate of the capital drives most of the dynamics.

In conclusion, this study has contributed to our understanding of epidemic dynamics in metropolitan areas and the role of optimal control in mitigating disease outbreaks. It has also opened up several interesting directions for future research. As we continue to grapple with the challenges posed by infectious diseases, studies like this one will be crucial in guiding our response and ensuring the health and well-being of our communities.

5 Acknowledgments

This research was mainly developed while the first author was a graduate student at the School of Applied Mathematics (FGV EMAp). The authors kindly thank Guilherme Tegoni Goedert from FGV EMAp for providing the virtual machine used in this work; and to Luis Briceño Arias (Universidad Técnica Federico Santa María, Chile), Luis Almeida (Laboratoire Jacques-Louis Lions, France), and Maria do Rosário de Pinho (Universidade do Porto, Portugal) for their insightful discussions and valuable suggestions. Both authors acknowledge the financial support from the School of Applied Mathematics (FGV EMAp), and Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) for the funding through process E-26/203.223/2017. The second author further thanks the financial support of CNPq (Brazil) through process 310452/2019-8.

References

[1] J. Arino and P. Van den Driessche. A multi-city epidemic model. *Mathematical Population Studies*, 10(3): 175–193, 2003.

[2] E. Asano, L. J. Gross, S. Lenhart, and L. A. Real. Optimal control of vaccine distribution in a rabies metapopulation model. *Mathematical Biosciences and Engineering*, 5(2):219–238, 2008.

[3] F. Avram, L. Freddi, and D. Goreac. Optimal control of a sir epidemic with icu constraints and target objectives. *Applied Mathematics and Computation*, 418:126816, 2022.

[4] L. D. Beal, D. C. Hill, R. A. Martin, and J. D. Hedengren. Gekko optimization suite. *Processes*, 6(8):106, 2018.

[5] H. Behncke. Optimal control of deterministic epidemics. *Optimal control applications and methods*, 21(6): 269–285, 2000.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

[6] M. H. A. Biswas, L. T. Paiva, M. D. R. de Pinho, et al. A seir model for control of infectious diseases with constraints. *Mathematical Biosciences and Engineering*, 11(4):761–784, 2014.

[7] F. Brauer, C. Castillo-Chavez, and C. Castillo-Chavez. *Mathematical models in population biology and epidemiology*, volume 2. Springer, 2012.

[8] A. Bressan and B. Piccoli. *Introduction to the mathematical theory of control*, volume 1. American institute of mathematical sciences Springfield, 2007.

[9] V. Colizza and A. Vespignani. Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations. *Journal of theoretical biology*, 251(3):450–467, 2008.

[10] V. Colizza, R. Pastor-Satorras, and A. Vespignani. Reaction–diffusion processes and metapopulation models in heterogeneous networks. *Nature Physics*, 3(4):276–282, 2007.

[11] B. de Lapasse and É. Prévost. La mobilité des retraités à tous les âges. *Retraite et société*, 86(2):211–228, 2021.

[12] M. D. R. De Pinho, I. Kornienko, and H. Maurer. Optimal control of a seir model with mixed constraints and l 1 cost. In *CONTROLO2014–Proceedings of the 11th Portuguese Conference on Automatic Control*, pages 135–145. Springer, 2015.

[13] H. Gaff and E. Schaefer. Optimal control applied to vaccination and treatment strategies for various epidemiological models. *Math. Biosci. Eng*, 6(3):469–492, 2009.

[14] G. H. Golub. Some modified matrix eigenvalue problems. *SIAM review*, 15(2):318–334, 1973.

[15] A. Gutiérrez, L. Pereyra, and M. Bruschi. Encuesta enmodo 2009-2010, 2021. URL https://transporteterritorio.wixsite.com/pttuba/cartrografias-de-movilidad. Cartografía base del Instituto Geográfico Nacional.

[16] E. Hansen and T. Day. Optimal control of epidemics with limited resources. *Journal of mathematical biology*, 62:423–451, 2011.

[17] J. C. Lemaitre, D. Pasetto, M. Zanon, E. Bertuzzo, L. Mari, S. Miccoli, R. Casagrandi, M. Gatto, and A. Rinaldo. Optimal control of the spatial allocation of covid-19 vaccines: Italy as a case study. *PLoS computational biology*, 18(7):e1010237, 2022.

[18] S. Lenhart and J. T. Workman. *Optimal control applied to biological models*. CRC press, 2007.

[19] S. Levin, T. Powell, and J. Steele. Patch dynamics. Technical report, Springer-Verlag, New York, NY (United States), 1993.

[20] R.-C. Li. Matrix perturbation theory. In *Handbook of linear algebra*, pages 15–1. Chapman and Hall/CRC, 2006.

[21] J. Liu and Y. Zhou. Global stability of an sirs epidemic model with transport-related infection. *Chaos, Solitons & Fractals*, 40(1):145–158, 2009.

[22] X. Liu and P. Stechlinski. Transmission dynamics of a switched multi-city model with transport-related infections. *Nonlinear Analysis: Real World Applications*, 14(1):264–279, 2013.

[23] L. M. Moschen. Optimal vaccination strategies for epidemics in metropolitan areas. Master's dissertation, Fundação Getulio Vargas, Aug 2023. URL https://hdl.handle.net/10438/34154.

[24] R. M. Neilan and S. Lenhart. An introduction to optimal control with an application in disease modeling. In *Modeling paradigms and analysis of disease transmission models*, pages 67–81, 2010.

[25] L. G. Nonato, P. Peixoto, T. Pereira, C. Sagastizábal, and P. J. Silva. Robot Dance: A mathematical optimization platform for intervention against COVID-19 in a complex network. *EURO Journal on Computational Optimization*, 10:100025, 2022. ISSN 21924406. doi: 10.1016/j.ejco.2022.100025. URL https://linkinghub.elsevier.com/retrieve/pii/S2192440622000016.

[26] P. Ögren and C. F. Martin. Vaccination strategies for epidemics in highly mobile populations. *Applied mathematics and computation*, 127(2-3):261–276, 2002.

[27] R. Pastor-Satorras and A. Vespignani. Epidemic dynamics and endemic states in complex networks. *Physical Review E*, 63(6):066117, 2001.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

[28] R. Pastor-Satorras and A. Vespignani. Epidemic dynamics in finite size scale-free networks. *Physical Review E*, 65(3):035108, 2002.

[29] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani. Epidemic processes in complex networks. *Reviews of modern physics*, 87(3):925, 2015.

[30] R. E. Rowthorn, R. Laxminarayan, and C. A. Gilligan. Optimal control of epidemics in metapopulations. *Journal of the Royal Society Interface*, 6(41):1135–1144, 2009.

[31] L. A. Rvachev and I. M. Longini Jr. A mathematical model for the global spread of influenza. *Mathematical biosciences*, 75(1):3–22, 1985.

[32] L. Sattenspiel and K. Dietz. A structured epidemic model incorporating geographic mobility among regions. *Mathematical biosciences*, 128(1-2):71–91, 1995.

[33] H. Schättler and U. Ledzewicz. *Geometric optimal control: theory, methods and examples*, volume 38. Springer, 2012.

[34] Sebrae. Mobilidade urbana e mercado de trabalho no rio de janeiro. [Internet]. Sebrae, 2013. Available at: https://sebrae.com.br/Sebrae/Portal%20Sebrae/UFs/RJ/Menu%20Institucional/SEBRAE_EPG_set13_mob_urb_merc_trab_rj.pdf.

[35] O. Sharomi and T. Malik. Optimal control in epidemiology. *Annals of Operations Research*, 251:55–71, 2017.

[36] L. M. Stolerman, D. Coombs, and S. Boatto. Sir-network model and its application to dengue fever. *SIAM Journal on Applied Mathematics*, 75(6):2581–2609, 2015.

[37] N. J. Stor, I. Slapničar, and J. L. Barlow. Forward stable eigenvalue decomposition of rank-one modifications of diagonal matrices. *Linear Algebra and its Applications*, 487:301–315, 2015.

[38] Y. Takeuchi, Y. Saito, et al. Spreading disease with transport-related infection. *Journal of theoretical biology*, 239(3):376–390, 2006.

[39] 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, 2002.

[40] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, et al. Scipy 1.0: fundamental algorithms for scientific computing in python. *Nature methods*, 17(3):261–272, 2020.

[41] Q. Yin, Z. Wang, C. Xia, M. Dehmer, F. Emmert-Streib, and Z. Jin. A novel epidemic model considering demographics and intercity commuting on complex dynamical networks. *Applied Mathematics and Computation*, 386:125517, Dec. 2020. ISSN 0096-3003. doi: 10.1016/j.amc.2020.125517. URL https://www.sciencedirect.com/science/article/pii/S0096300320304756.

A The basic reproduction number for general compartmental models

We present the calculus of $\mathcal{R}_0$ within the framework of a general compartmental model, following the seminal work of Van den Driessche and Watmough (2002) [39]. They focus on autonomous systems, namely, those whose right-hand side does not explicitly depend on t. Let $x \in \mathbb{R}^{m+n}_{\geq 0}$ (real vectors with non-negative entries) be the number (or proportion) of individuals in $m+n$ compartments, where the first m are infected states and the remaining n are non-infected states. The rate of new infections in the ith infected compartment is denoted by $\mathcal{F}_i$, the rate of transfer into the ith compartment (except new infections) by $\mathcal{V}_i^+$, and the rate of transfer out of the ith compartment by $\mathcal{V}_i^-$. The net transfer is then given by $\mathcal{V}_i = \mathcal{V}_i^- - \mathcal{V}_i^+$. The disease transmission model consists of the equations

$$x_i' = f_i(x) = \mathcal{F}_i(x) - \mathcal{V}_i(x), \text{ for } i = 1, \ldots, m+n.$$

The choice of the infected and non-infected compartments depends on the model's interpretation.

A *disease-free state* is a state $x \in \mathbb{R}^{m+n}_{\geq 0}$ such that $x_i = 0$ for $i = 1, \ldots, m$, and a *disease-free equilibrium* (DFE) is a disease-free state that is an asymptotically stable equilibrium. We ensure the model's well-posedness and the existence of such equilibrium, we make the following assumptions:

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

(A1) The transfer of individuals between compartments is non-negative: if $x \in \mathbb{R}^{m+n}_{\geq 0}$, then

$$\mathcal{F}_i(x), \mathcal{V}_i^+(x), \mathcal{V}_i^-(x) \geq 0.$$

(A2) There is no transfer out of an empty compartment: if $x = 0$, then $\mathcal{V}_i^-(x) = 0$.

(A3) Non-infected compartments do not receive new infections: $\mathcal{F}_i(x) = 0$ for $i > m$.

(A4) The set of the disease-free states is invariant: if $x_i = 0$ for $i = 1, \ldots, m$, then $\mathcal{F}_i(x) = \mathcal{V}_i^+(x) = 0$ for $i = 1, \ldots, m$.

(A5) The DFE is stable in the absence of new infections: if $\mathcal{F}(x) = 0$ and x_0 is a DFE, $Df(x_0)$ is a Hurwitz matrix.

From these assumptions and if x_0 is a DFE, we obtain that

$$D\mathcal{F}(x_0) = \begin{bmatrix} F & 0 \\ 0 & 0 \end{bmatrix}, \quad D\mathcal{V}(x_0) = \begin{bmatrix} V & 0 \\ J_3 & J_4 \end{bmatrix},$$

where F is non-negative, V is a non-singular M-matrix and the eigenvalues of J_4 have positive real part [39, Lemma 1]. A square matrix B is an *M-matrix* if it can be expressed as $B = sI - P$, where P is a matrix with non-negative elements, and s is a real scalar such that $s \geq \rho(P)$. If $s > \rho(P)$, then B is a *non-singular M-matrix*. If $s = \rho(P)$, it is a *singular M-matrix*. The matrix FV^{-1} is named the *next generation matrix* for the model and

$$\mathcal{R}_0 = \rho(FV^{-1}), \tag{11}$$

in which $\rho(A)$ is the spectral radius of the matrix A. The (i, j) entry of matrix FV^{-1} estimates how many new infections are expected to arise in compartment i if an infected individual is introduced into compartment j. This definition implies that x_0 is asymptotically stable if $\mathcal{R}_0 < 1$, but unstable if $\mathcal{R}_0 > 1$ [39, Theorem 2].

A.1 Computing the basic reproduction number for our model

Following the notation of Appendix A, we reorder the compartments as follows:

$$x = (I_1, \ldots, I_K, S_1, \ldots, S_K, R_1, \ldots, R_K).$$

The rate of new infections is

$$\mathcal{F}_i = \alpha\beta_i S_i I_i + (1-\alpha) S_i \sum_{j=1}^{K} \beta_j p_{ij} I_j^{\text{eff}}, \text{ for } i = 1, \ldots, K,$$

and zero for non-infected compartments. The rate of transfer between compartments is defined as

$$\mathcal{V}_k = \begin{cases} (\gamma+\mu) I_i, & 1 \leq k \leq K,\ i = k, \\ \alpha\beta_i S_i I_i + (1-\alpha) S_i \sum_{j=1}^{K} \beta_j p_{ij} I_j^{\text{eff}} + \mu S_i - \mu, & K+1 \leq k \leq 2K,\ i = k-K, \\ \mu R_i - \gamma I_i, & 2K+1 \leq k \leq 3K,\ i = k-2K. \end{cases}$$

Utilizing the auxiliary result

$$\frac{dI_k^{\text{eff}}}{dI_j} = \frac{p_{jk} n_j}{P_k^{\text{eff}}},$$

we proceed to calculate matrices F and V. The elements of these matrices are given by

$$F_{ij} = \alpha\beta_i S_i I_i \delta_{ij} + (1-\alpha) S_i \sum_{k=1}^{K} \beta_k p_{ik} p_{jk} \frac{n_j}{P_k^{\text{eff}}},$$

$$V_{ij} = \gamma + \mu,$$

where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. We can represent the matrix F as

$$F = \alpha\mathcal{S}B + (1-\alpha)\mathcal{S}PBE^{-1}P^T N = \mathcal{S}\left[\alpha B + (1-\alpha)PBE^{-1}P^T N\right], \tag{12}$$

where $\mathcal{S}, N, B$ and E are diagonal matrices, such that $\mathcal{S}_{ii} = S_i$, $N_{ii} = n_i$, $B_{ii} = \beta_i$ and $E_{ii} = P_i^{\text{eff}}$. Intuitively, the entry (i, j) of F is the rate at which infected individuals in city j contribute to new infections in city i. Therefore, the sum $\sum_{j=1}^{K} A_{ij}$ quantifies the total rate of new infections in city i. Then

$$FV^{-1} = \frac{1}{\gamma+\mu} F,$$

the eigenvalues of FV^{-1} are the eigenvalues of F divided by $\gamma + \mu$, and $\mathcal{R}_0 = \rho(F)/(\gamma+\mu)$. We can readily see that matrix F is equivalent to a symmetric matrix, which implies the following lemma.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Lemma A.1. *The eigenvalues of FV^{-1} are real.*

As far as we know, there is no closed-form expression for the spectral radius of F as a function of the parameters of the model. However, we can simplify equation (12) by applying Assumption 2.1 which defines a specific format for P. Consequently, we calculate that

$$PBE^{-1} = \begin{bmatrix} \frac{\beta_1}{P_1^{\text{eff}}} & 0 & 0 & \cdots & 0 \\ \frac{\beta_1}{P_1^{\text{eff}}} p_{21} & \frac{\beta_2}{P_2^{\text{eff}}} p_{22} & 0 & \cdots & 0 \\ \frac{\beta_1}{P_1^{\text{eff}}} p_{31} & 0 & \frac{\beta_3}{P_3^{\text{eff}}} p_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\beta_1}{P_1^{\text{eff}}} p_{K1} & 0 & 0 & \cdots & \frac{\beta_K}{P_K^{\text{eff}}} p_{KK} \end{bmatrix}$$

and, subsequently,

$$PBE^{-1}P^T N = \frac{\beta_1}{P_1^{\text{eff}}} \bar{p}\bar{p}^T N + \Lambda,$$

where $\bar{p} = (1, p_{21}, \ldots, p_{K1})$ and $\Lambda = \text{diag}\left(\beta_k p_{kk}(1-\delta_{k1})\right)_{k=1,\ldots,K}$, with $P_k^{\text{eff}} = p_{kk} n_k$ for $k > 1$. Therefore, we can express the matrix F from (12) as

$$F = \mathcal{S}\left[\alpha B + (1-\alpha)\left(\frac{\beta_1}{P_1^{\text{eff}}} \bar{p}\bar{p}^T N + \Lambda\right)\right] = \mathcal{S}[\alpha B + (1-\alpha)\Lambda] + (1-\alpha)\frac{\beta_1}{P_1^{\text{eff}}}(\mathcal{S}\bar{p})(\bar{p}^T N), \tag{13}$$

which is a sum of a diagonal matrix and a product of vectors. The problem of finding the eigenvalues of matrices with this structure, referred to as *diagonal plus rank-one matrix*, such as matrix F, is well-known in the literature, and there are specific algorithms to solve it, such as [14, 37]. An inequality for $\mathcal{R}_0$ can be computed through Weyl's inequality by considering this structure.

The unique disease-free equilibrium of system (1) sets $S_i = 1$ and $I_i = R_i = 0$ for all cities $i = 1, \ldots, K$, which simplifies $\mathcal{S}$ to the identity matrix. In the modified model (5) with a constant vaccination rate in the population, the dynamics of compartments I_i are unchanged, which leads to the same calculation of the basic reproduction number. However, the disease-free equilibrium is redefined as

$$\mu - \mu S_i - u_i S_i = 0 \implies S_i = \frac{\mu}{\mu + u_i} \quad \text{and} \quad R_i = \frac{u_i}{\mu + u_i},$$

besides $I_i = 0$. Therefore the only change is on the diagonal matrix $\mathcal{S}$.

B Proofs

This section includes all the proofs of results given throughout the text.

Proposition B.1 (Positive Invariance)**.** *The set*

$$\mathcal{C} := \{X \in \mathbb{R}_{\geq 0}^{K\times 3} : X_{i1} + X_{i2} + X_{i3} = 1, \ \textit{for } i = 1, \ldots, K\}$$

is positively invariant under the flow of system (1)*.*

Proof. Firstly, by the smoothness of the system, existence and uniqueness of solution $[\boldsymbol{S}(t), \boldsymbol{I}(t), \boldsymbol{R}(t)]$ in $[0, T]$ such that $[\boldsymbol{S}(0), \boldsymbol{I}(0), \boldsymbol{R}(0)] \in \mathcal{C}$ is guaranteed. By the uniqueness, since $\tilde{\boldsymbol{I}}(t) = \tilde{\boldsymbol{R}}(t) = 0$ is also a solution over $[0, T]$, we have $I_i(t), R_i(t) \geq 0$ for all $i = 1, \ldots, K$. Moreover, notice that

$$S_i'(t) \geq -S_i(t)\left[\alpha\beta_i I_i + (1-\alpha)\sum_{j=1}^{K} \beta_j p_{ij} I_j^{\text{eff}} + \mu\right],$$

which implies, by Gronwall's Inequality, that

$$-S_i(t) \leq -S_i(0)\exp\left\{-\int_0^T \alpha\beta_i I_i(s) + (1-\alpha)\sum_{j=1}^{K} \beta_j p_{ij} I_j^{\text{eff}}(s)\, ds - \mu\right\} \leq 0.$$

This proves the non-negativity of the solution. Finally it is straightforward that for each i, $S_i'(t) + I_i'(t) + R_i'(t) = 0$, which implies that $S_i(t) + I_i(t) + R_i(t) = S_i(0) + I_i(0) + R_i(0) = 1$ for all $t \in [0, T]$. If we consider the parameter α as a constant by parts function, notice that the result is proven by induction in each interval in which $\alpha(\cdot)$ is constant. □

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Proof of Theorem 2.1. Formula (2) is derived from the basic reproduction number from a simple SIR model. By considering the expression (12) for matrix F, for each $1 \leq i, j \leq K$,

$$\begin{aligned} F_{ij} &= S_i \left[\alpha\beta_i\delta_{ij} + (1-\alpha)\big(PBE^{-1}P^TN\big)_{ij}\right] \\ &= S_i \left[\alpha\beta_i\delta_{ij} + (1-\alpha)\sum_{k=1}^{K}\big(PBE^{-1}\big)_{ik}\big(P^TN\big)_{kj}\right] \\ &= S_i \left[\alpha\beta_i\delta_{ij} + (1-\alpha)\sum_{k=1}^{K} p_{ik}\beta_k p_{jk} n_j / P_k^{\text{eff}}\right], \end{aligned}$$

where $\delta_{ij} = 1 \iff i = j$ and $\delta_{ij} = 0$ otherwise. Assuming $\alpha > 0$ and that for all cities $S_i > 0$ and $\beta_i > 0$, we infer that $F_{ij} > 0$. This means that F is a positive matrix and satisfies Perron-Frobenius's theorem. We conclude that there is $r > 0$ such that r is an eigenvalue of F and $\rho(F) = r$. Furthermore,

$$\min_i \sum_{j=1}^{K} F_{ij} \leq r \leq \max_i \sum_{j=1}^{K} F_{ij}.$$

Calculating,

$$\begin{aligned} \sum_{j=1}^{K} F_{ij} &= S_i \left[\alpha\beta_i + (1-\alpha)\sum_{j=1}^{K}\sum_{k=1}^{K} p_{ik}\beta_k / P_k^{\text{eff}} p_{jk} n_j\right] \\ &= S_i \left[\alpha\beta_i + (1-\alpha)\sum_{k=1}^{K} p_{ik}\beta_k / P_k^{\text{eff}} \sum_{j=1}^{K} p_{jk} n_j\right] \\ &= S_i \left[\alpha\beta_i + (1-\alpha)\sum_{k=1}^{K} p_{ik}\beta_k\right], \end{aligned}$$

which implies that

$$\min_i S_i \left[\alpha\beta_i + (1-\alpha)\sum_{k=1}^{K} p_{ik}\beta_k\right] \leq \rho(F) \leq \max_i S_i \left[\alpha\beta_i + (1-\alpha)\sum_{k=1}^{K} p_{ik}\beta_k\right].$$

At the DFE, where $S_i = 1$ for all cities i, we conclude that

$$\frac{\min_i \alpha\beta_i + (1-\alpha)\sum_{k=1}^{K} p_{ik}\beta_k}{\gamma + \mu} \leq \mathcal{R}_0 \leq \frac{\max_i \alpha\beta_i + (1-\alpha)\sum_{k=1}^{K} p_{ik}\beta_k}{\gamma + \mu}$$

or, in terms of the basic reproduction numbers of the isolated cities,

$$\min_i \alpha\mathcal{R}_0^i + (1-\alpha)\sum_{k=1}^{K} p_{ik}\mathcal{R}_0^k \leq \mathcal{R}_0 \leq \max_i \alpha\mathcal{R}_0^i + (1-\alpha)\sum_{k=1}^{K} p_{ik}\mathcal{R}_0^k.$$

□

Proof of Corollary 2.1. Immediate since $\alpha + (1-\alpha)\sum_{k=1}^{K} p_{ik} = 1$ and, therefore, w_i is a convex combination of $\mathcal{R}_0^1, \ldots, \mathcal{R}_0^K$. □

Proof of Theorem 2.2. Following the expression (13), rewrite the matrix F as

$$F = N^{-1/2}\mathcal{S}^{1/2}[\alpha\mathcal{S}B + (1-\alpha)\mathcal{S}\Lambda + (1-\alpha)\frac{\beta_1}{P_1^{\text{eff}}}\mathcal{S}^{1/2}N^{1/2}\bar{p}\bar{p}^TN^{1/2}\mathcal{S}^{1/2}]\mathcal{S}^{-1/2}N^{1/2},$$

which is similar, and therefore has the same eigenvalues, to

$$\bar{F} = \alpha\mathcal{S}B + (1-\alpha)\mathcal{S}\Lambda + (1-\alpha)\frac{\beta_1}{P_1^{\text{eff}}}\mathcal{S}^{1/2}N^{1/2}\bar{p}\Big(\mathcal{S}^{1/2}N^{1/2}\bar{p}\Big)^T.$$

By setting $\mathcal{S}$ as the identity matrix (DFE) and utilizing the symmetry of the matrices, we can apply Weyl's inequality to obtain

$$\rho(\alpha B + (1-\alpha)\Lambda) \leq \rho(F) \leq \rho(\alpha B + (1-\alpha)\Lambda) + (1-\alpha)\frac{\beta_1}{P_1^{\text{eff}}}\big(N^{1/2}\bar{p}\big)^TN^{1/2}\bar{p},$$

since the only non-zero eigenvalue of a matrix vv^T is v^Tv. We finally set $\xi = \rho(\alpha B + (1-\alpha)\Lambda)$ and calculate $\big(N^{1/2}\bar{p}\big)^TN^{1/2}\bar{p} = \sum_{i=1}^{K} n_i p_{i1}^2$. □

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Proposition B.2 (Positive Invariance with control)**.** *The region*

$$\mathcal{C} = \{X \in \mathbb{R}_{\geq 0}^{K\times 4} : X_{i1} + X_{i2} + X_{i3} = 1, \ \textit{for } i = 1, \dots, K\}$$

is positively invariant under the flow of system (5)*, for each measurable function* u*.*

Proof. Let u be a measurable function in $[0, T]$ and suppose it assumes values on a compact set U. Since the dynamic is smooth, there is a unique solution $[\boldsymbol{S}(t), \boldsymbol{I}(t), \boldsymbol{R}(t), \boldsymbol{V}(t)]$ for $t \in [0, T]$ [8]. As similarly proven for Proposition B.1,

$$-S_i(t) \leq -S_i(0) \exp\left\{ -\int_0^T \alpha\beta_i I_i(s) + (1-\alpha)\sum_{j=1}^K \beta_j p_{ij} I_j^{\text{eff}}(s) + u_i(s)\, ds - \mu \right\} \leq 0, \tag{14}$$

for each $i = 1, \dots, K$. By uniqueness, $I_i(t) \geq 0$ for each $i = 1, \dots, K$ since the solution cannot cross the solution with no infections. Moreover, $R_i'(t) \geq -\mu R_i(t)$, which results, by Gronwall's Inequality, in

$$-R_i(t) \leq -R_i(0)e^{-\mu t} \leq 0.$$

Finally $V_i'(t) \geq 0$ for every $t \in [0, T]$, which implies $V_i(t) \geq V_i(0) = 0$ for every $t \in [0, T]$. We conclude by observing that

$$\frac{d(S_i + I_i + R_i)}{dt} = 0 \implies (S_i + I_i + R_i)(t) = (S_i + I_i + R_i)(0) = 1,$$

for all $t > 0$ and $i = 1, \dots, K$. Supposing that U is compact was without loss of generality because we impose that $u_i(t)S_i(t) \leq v_i^{\max}$ for each i and each t. Therefore,

$$v_i^{\max} \geq u_i(t)S_i(t) \geq u_i(t) \min_{t\in[0,T]} S_i(t) \implies 0 \leq u_i(t) \leq v_i^{\max} / \min_{t\in[0,T]} S_i(t)$$

and $\min_{t\in[0,T]} S_i(t) > 0 \iff S_i(0) > 0$ by inequality (14). □

Proof of Lemma A.1. The matrix F can be rewritten as

$$F = N^{-1/2}\mathcal{S}^{1/2}[\alpha B\mathcal{S} + (1-\alpha)\mathcal{S}^{1/2}N^{1/2}PBE^{-1}P^TN^{1/2}\mathcal{S}^{1/2}]\mathcal{S}^{-1/2}N^{1/2},$$

which is similar, in the sense of matrices, to $\alpha B\mathcal{S} + (1-\alpha)\mathcal{S}^{1/2}N^{1/2}PBE^{-1}P^TN^{1/2}\mathcal{S}^{1/2}$. The matrix $B\mathcal{S}$ is diagonal, while $\mathcal{S}^{1/2}N^{1/2}PBE^{-1}P^TN^{1/2}\mathcal{S}^{1/2}$ is symmetric. Therefore F is similar to a symmetric matrix and, consequently, it is symmetric. Consequently, the eigenvalues of F are real and so are the eigenvalues of FV^{-1}. □

C Additional simulations and figures

In this section, we present additional results that are less central to the overall work but may be of interest to the audience.

C.1 On the effect of some parameters in $\mathcal{R}_0$

Throughout the text, we have observed that the transmission rate primarily influences the spread of the disease in the population. Considering a scenario involving a metropolitan area with two cities, Figure C.21 illustrates the impact of parameters α and p_{21} on $\mathcal{R}_0$. The histogram displays possible values of $\mathcal{R}_0$ as a function of random combinations of these parameters. Notably, it reveals that higher values of the ratio n_1/n_2 shrink the distribution, as a byproduct of the fact that as n_1/n_2 approaches infinity, $\mathcal{R}_0$ converges to $\mathcal{R}_0^1$. By setting β_1 close to γ, then $\mathcal{R}_0 \approx 1$, we observe that for smaller values of the ratio n_1/n_2, the mobility parameters may affect the threshold $\mathcal{R}_0 > 1$, changing the behavior of the disease.

C.2 Epidemic values for different structures of the transition matrix

In the context of multiple cities, we consider a scenario of $K = 5$ cities. We first examine five distinct mobility scenarios: (I) metropolitan area where 10% of the residents of each city work in the capital; (II) metropolitan area where 40% of each city works in the capital; (III) all cities are interconnected, where 60% of the residents work in their city, while the remaining are distributed evenly across the other four cities; (IV) a variant of a metropolitan area where 30% of each city's residents work in the capital and 10% of the capital's workforce is employed in the other cities; (V) no commuting between cities. Our findings are summarized in Table C.2.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

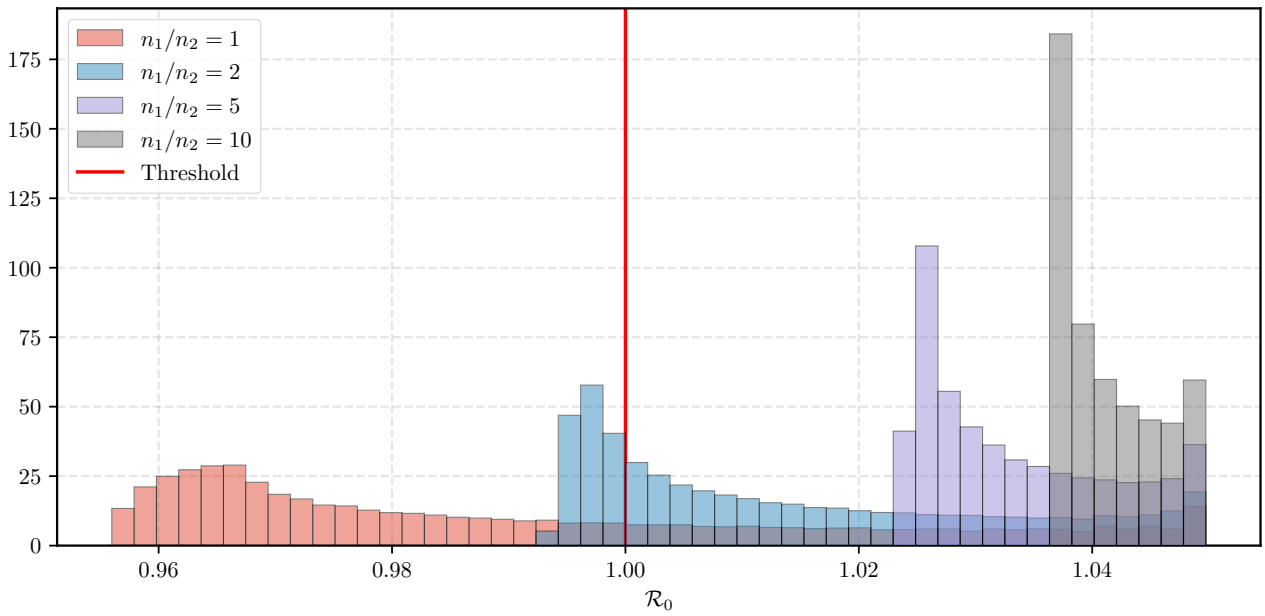


Figure C.21: **Histogram of the possible values of $\mathcal{R}_0$ as function of α and p_{21} for two cities**: histogram of the values of $\mathcal{R}_0$ generated by randomly selecting $\alpha \in (0, 1)$ and $p_{21} \in (0, 0.9)$. We consider four scenarios for the ratio n_1/n_2, $(1, 2, 5, 10)$. The other parameters are $\beta_1 = 0.15$, $\gamma = 1/7$ and $\mu = 3.6 \cdot 10^{-5}$, the same as in Figure 6.

Metrics	Scenarios	Cities 1	2	3	4	5	Aggregated
Peak size (%)	(I)	27.0	14.1	9.3	5.7	3.7	20.5
	(II)	26.3	18.4	15.3	12.5	10.2	22.5
	(III)	24.2	17.3	14.7	12.5	10.2	21.0
	(IV)	25.5	17.7	14.3	12.7	9.8	21.7
	(V)	27.5	10.9	4.5	0.0	0.0	19.2
Peak day	(I)	39	53	55	52	49	40
	(II)	40	46	47	47	47	41
	(III)	44	49	50	50	50	45
	(IV)	41	47	49	50	50	45
	(V)	38	81	134	—	—	38
Duration (days)	(I)	146	188	218	234	191	191
	(II)	145	164	170	172	166	157
	(III)	151	168	172	171	168	161
	(IV)	147	167	175	167	164	160
	(V)	137	228	350	—	—	306
Attack rate (%)	(I)	91.4	72.9	58.1	37.4	20.4	82.5
	(II)	91.2	77.6	69.0	58.3	47.0	85.2
	(III)	89.8	75.5	67.4	58.3	47.9	83.7
	(IV)	90.7	76.6	66.9	57.5	44.9	84.3
	(V)	91.5	70.6	50.8	0.0	0.0	80.4

Table C.2: **Simulation results for different structures of the transition matrix**: presents the comparison of Peak Size (maximum proportion of infectious individuals), Peak Day (day of the peak size), Duration (time from epidemic onset to the day when the proportion of infectious individuals achieves 10^{-5}), and Attack Rate (proportion of individuals who contract the disease during 350 days of the epidemic) across different mobility scenarios. The aggregated column represents the metrics considering all cities together. For this experiment we $\beta = (0.4, 0.25, 0.2, 0.15, 0.1)$, $\alpha = 0.64$ and population sizes $10^5 \cdot (50, 10, 10, 1, 1)$. Lastly, the initial conditions are $I_1(0) = I_2(0) = I_3(0) = 10^{-4}$ and no recovered individuals.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

C.3 Evaluating the sharpness of the bound for $\mathcal{R}_0$

Considering a scenario with two cities, we perform a simulation over $N = 100,000$ iterations which randomly assigns values for $\beta_2 \in (0.01, 0.2)$, $\beta_1 \in (\beta_2, 0.6)$, $\alpha \in (0,1)$, $p_{21} \in (0, 0.9)$, and $n_1 \in (1, 10) \cap \mathbb{Z}$. The simulation then computes the upper and lower bounds using both general (Theorem 2.1) and metropolitan (Theorem 2.2) bounds, recording whether the latter provides a tighter bound than the former. The results reveal that the metropolitan method offers better bounds in 50.0% of the cases. In 18.1% of the cases, the metropolitan method provides a better upper bound but not a better lower bound, while in 31.9% of the cases, the general method provides better bounds. In the remaining 18.0% of cases, the bounds are equal. When we assume $\alpha \geq 0.5$, the bounds from Theorem 2.2 are better in 87.1% of the simulations and better only for the upper bounds in 4.9% of the simulations. The general method provides better bounds in only 8.1% of the simulations.

In the five-city scenario, we perform a simulation that runs for $N = 500,000$ iterations, with each iteration randomly generating values for $\beta \in (0.01, 0.6)$, $\alpha \in (0,1)$, $p \in (0, 0.9)^4$ being the transition probabilities from each of the other four cities to the capital, and $n \in (1, 50)^5 \cap \mathbb{Z}$ the population sizes. We sort the vector n in descending order. The code then calculates the upper and lower bounds, the same as before. The results of the simulation indicate that the metropolitan method provides better bounds in 48.9% of the cases. In 46% of the cases, the metropolitan method provides a better lower bound but not a better upper bound, while in 3.8% of the cases, the general method provides better bounds. When we assume $\alpha \geq 0.5$, the bounds of Theorem 2.2 are better in 80.6% of the simulations and better only for the lower bounds in 19.4% of the simulations. No recorded simulation had better bounds for the general method, arguing in favor of tighter bounds for $K = 5$ cities.

C.4 Impact of vaccination in the peak size

Although $\mathcal{R}_0$ provides a good understanding of the impact of vaccination on epidemic dynamics, it only informs about the asymptotic behavior of the dynamics. For a finite time horizon, analyzing the peak size of the epidemic provides crucial insights into the optimal vaccination strategy. This is the goal of Figure C.22a, which presents four different scenarios under a uniform vaccination policy. We have chosen $s = 42$ days for this experiment. The peak size of the baseline case, when no vaccine is introduced, is about 21% of the population. To reduce it to around 15%, as shown in orange, over 60% of the population needs to be vaccinated. Early vaccination can significantly reduce the peak size of the epidemic, as illustrated in Figure C.22b, which shows the impact of the vaccination starting date on the peak size.

C.5 Evaluating the impact of the nighttime proportion α

The parameter α, representing the average proportion of the time spent in their home city, was previously shown to not have a significant impact on $\mathcal{R}_0$ in a two-city model. However, in the multi-city framework, its effect on $\mathcal{R}_0$ is more pronounced, as depicted in Figure C.23. Interestingly, an increase in α generates higher values for $\mathcal{R}_0$ but does not lead to a higher peak size or attack rate. Another conclusion we draw is that the intensity of the epidemic increases with a higher amount of time spent in the capital, which occurs when α is closer to 0.

D Further remarks and possible extensions

D.1 Generalizing commuting patters

It is possible to consider different commuting times for each pair of cities α_{ij}, which is a set of data that can be easily retrieved. One can in fact generalize this feature by taking time-dependent parameters α_{ij}, leading to the following modified model

$$P_i^{\text{eff}}(t) = \sum_{j=1}^{K} (1 - \alpha_{ji}(t)) p_{ji} n_j,$$

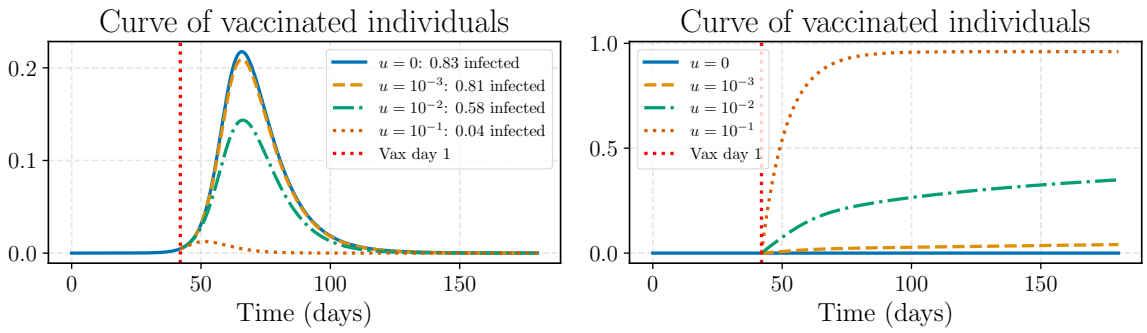
$$I_j^{\text{eff}}(t) = \frac{1}{P_j^{\text{eff}}(t)} \sum_{k=1}^{K} (1 - \alpha_{kj}(t)) p_{kj} I_k n_k.$$

Then, we get

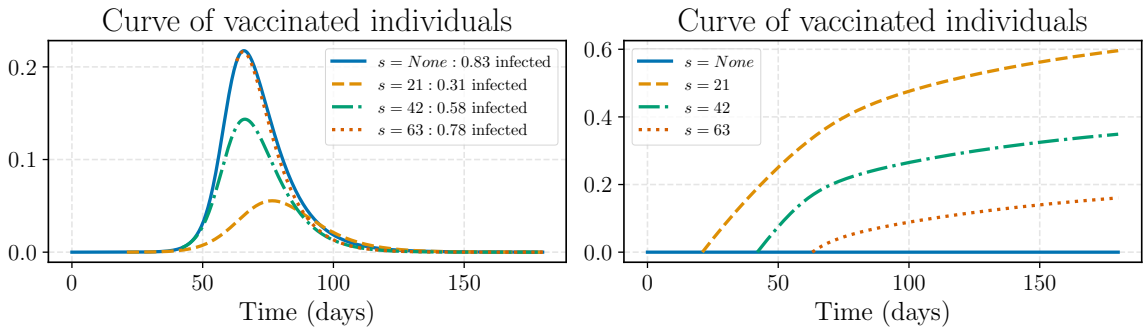
$$\frac{dS_i}{dt}(t) = \mu - \beta_i S_i \sum_{j=1}^{K} \alpha_{ji}(t) p_{ij} I_i - S_i \sum_{j=1}^{K} \beta_j (1 - \alpha_{ij}(t)) p_{ij} I_j^{\text{eff}} - \mu S_i,$$

This version also opens the door to considering infections during commuting.

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .



(a) Dynamics of infectious and vaccinated individuals in the metropolitan area



(b) Comparing different vaccination starting days

Figure C.22: **Epidemic behavior under vaccination strategies**: (a) This subfigure illustrates four epidemic scenarios under a constant vaccination strategy applied uniformly across all cities with the red line indicating the beginning of the vaccination campaign on day $s = 42$. The curves represent the proportion of infectious individuals (left) and the proportion of vaccinated individuals (right) across the entire metropolitan area. (b) Using the same settings, this subfigure shows different starting days of the vaccine campaign for $s = 21, 42, 63$ as the first day of vaccination. The parameters are: $\beta = (0.4, 0.3, 0.15, 0.15, 0.1)$, $\alpha = 0.64$, $p_{k1} = 0.2$ for each city $k > 1$, and $n = 10^5(50, 10, 10, 1, 1)$ is the vector of population sizes, starting with only one exposed individual in the capital.

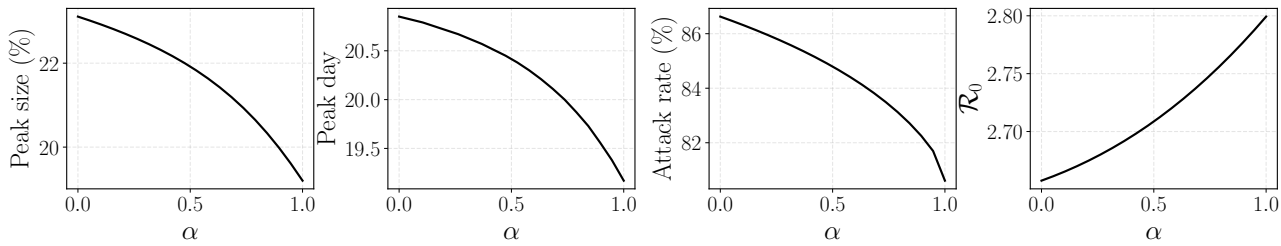


Figure C.23: **Impact of α on the epidemic**: We analyze four different epidemiological quantities considering the aggregated population, the peak size (the highest value in the infectious curve), the peak day (the day when the peak size occurs), the attack rate (the number of individuals who contract the disease after $T = 200$ days), and $\mathcal{R}_0$, as functions of α. The parameters used are based on the setting in Table C.2, except for matrix P, where we assume that 20% of each city's population goes to the capital while the remaining population stays in their home city.

D.2 Modeling possible vaccination in the workplace

The model presented in this paper, in particular the system (5), assumes that individuals are vaccinated in their city of residence, leading to a vaccination rate of $u_i(t)S_i(t)$ per unit of time in each city. However, this assumption may not hold in all cases, specifically in cities like Rio de Janeiro where the health system is universal and people can be vaccinated in any city of preference. A particular case is when the individuals are vaccinated close to their workplace, which can be addressed as a simple modification.

To accommodate this fact, we propose an extension to our model: a proportion of individuals, denoted by q, get vaccinated in their city of work instead of their city of residence. Therefore, the quantity $qp_{ij}S_i$ represents

medRxiv preprint doi: https://doi.org/10.1101/2024.01.31.24302083; this version posted April 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

the proportion of people who live in the city i, work in the city j, and choose to be vaccinated at their workplace. This leads to a revised effective vaccination rate, expressed as:

$$(1-q)u_iS_i + q\sum_{j=1}^{K} u_j p_{ij} S_i.$$

This adjustment yields a modified SIR model with an altered counter variable V_i represented as follows:

$$\begin{aligned}
\frac{dS_i}{dt} &= \mu - \alpha\beta_i S_i I_i - (1-\alpha)S_i \sum_{j=1}^{K} \beta_j p_{ij} I_j^{\text{eff}} - (1-q)u_i S_i - q\sum_{j=1}^{K} u_j p_{ij} S_i - \mu S_i, \\
\frac{dI_i}{dt} &= \alpha\beta_i S_i I_i + (1-\alpha)S_i \sum_{j=1}^{K} \beta_j p_{ij} I_j^{\text{eff}} - \gamma I_i - \mu I_i, \\
\frac{dR_i}{dt} &= (1-q)u_i S_i + q\sum_{j=1}^{K} u_j p_{ij} S_i + \gamma I_i - \mu R_i, \\
\frac{dV_i}{dt} &= (1-q)u_i S_i + q\sum_{j=1}^{K} u_j p_{ij} S_i.
\end{aligned}$$

This change does not alter the paper's results regarding the overall behavior of the solution which focuses on the capital. Considering the role of the capital as the economic hub of the metropolitan area, this model reinforces the importance of prioritizing the vaccine delivery to the capital. This is because more people are there during the day and a higher vaccination rate effectively increases the vaccinated population in the capital and, indirectly, in the other cities.