LEMMA – Document de travail

DT 2026-03

The social value of information in times of epidemics

Chantal Marlats

Université de Tours, LEO and LEMMA

Lucie Ménager

Université Paris Panthéon-Assas, LEMMA

Dominique Baril-Tremblay

LEMMA EA 4442 - 4 rue Blaise Desgoffe 75006 Paris

The social value of information in times of epidemics*

Chantal Marlats†, Lucie Ménager‡, and Dominique Baril-Tremblay.

March 25, 2026

Abstract

We study an epidemic model in which individuals optimally choose social activity while facing uncertainty about the characteristics of the epidemic. Individual behavior affects infection dynamics, and beliefs evolve endogenously over time. We characterize the symmetric equilibrium and show that, even when equilibrium prevalence would decrease monotonically under full information, uncertainty can generate non-monotonic epidemic dynamics and multiple infection peaks. The mechanism operates through belief-driven shifts in the intertemporal allocation of social activity, leading to a premature relaxation of distancing when the epidemic is perceived to be severe. We analyze the welfare and mortality effects of public disclosure of the epidemic state. While transparency allows individuals to flatten the curve in severe epidemics, it may reduce welfare or increase mortality, depending on the timing of vaccine arrival. Overall, the analysis highlights a tension between welfare maximization and mortality minimization in epidemic information policies.

Keywords: SIR model; Social distancing; Uncertainty.

JEL codes: C73, D84, I12.

*This research was supported by the French National Research Agency (ANR-17-CE38-0005-01) and by the Labex MME-DII (ANR11-LBX-0023-01). We have greatly benefited from discussions with Raouf Boucekkine, Glenn Ellison, Johannes Hörner, Nenad Kos, Thomas Mariotti, Sven Rady, Régis Renault, Flavio Toxvaerd, Jean-Robert Tyran, Joel Sobel and with seminar participants at City London University, Dynamic games and Application 2022, ENS-Paris Saclay, ESEM 2023 and Frontiers in Epidemiological Economics 2023.

†Université de Tours, LEO and LEMMA, chantal.marlats@univ-tours.fr

‡Université Paris-Panthéon-Assas, LEMMA, lucie.menager@assas-universite.fr

1 Introduction

> *Hitherto neither the newspapers nor the Ransdoc Information Bureau had been given any official statistics relating to the epidemic. Now the Prefect supplied them daily to the bureau, with the request that they should be broadcast once a week.*
>
> Albert Camus
>
> The Plague, 1947

The dilemma faced by the prefect of Oran in The Plague is the same one that confronts any public-health authority during an epidemic of a severe infectious disease: what information should be revealed, when, and how often? During the COVID-19 pandemic, governments around the world adopted remarkably different disclosure strategies. In a comparative analysis of official communications from 324 government leaders, health ministers, and ministries in 139 countries, Weder and Courtois (2022) document substantial cross-country variation in the frequency, granularity, and transparency of public health reporting. To illustrate, in France the Ministry of Health publicly announced the number of new COVID-19 cases on national television every evening during the first lockdown, whereas in the United States, two years after the onset of the pandemic, the Centers for Disease Control and Prevention (C.D.C.) had released only a small fraction of the hospitalization data it had collected.

While several mechanisms[1] may account for such differences in disclosure patterns, limited transparency by public authorities often provoked strong public reactions. In the United States, The New York Times reported that the C.D.C. withheld critical data, prompting accusations of excessive secrecy and eroding public trust (Mandavilli, February 25, 2022). In Australia, journalists and researchers similarly noted that much of the modelling and data informing government decisions remained unpublished, thereby restricting external scrutiny of policy choices and epidemic projections (ABC News, January 7, 2023). Across these cases,

[1] Fear of misinterpretation by anti-vaccine groups is one reason for withholding data. This has been reported by Kristen Nordlund, a spokeswoman for the C.D.C. (The New York Times, February 25, 2022), and by public health officials in Scotland who stopped releasing hospitalization and death data by vaccination status in 2022 (The Scotsman, February 17, 2022). Moreover, the state of the health-system capacity might also explain differences in communication. Weder and Courtois (2022) find that countries with lower degrees of access to universal healthcare compensated their structural vulnerabilities by highlighting individual and community responsibilities over government measures, while countries with widely accessible and strongly equipped health care systems were more transparent.

the perceived opacity of official communication raised concerns about accountability and the credibility of public health institutions. The main argument advanced by journalists and scientists in favor of greater transparency is that detailed and timely information —such as data on hospitalizations, booster uptake, or wastewater analyses— can help protect the most vulnerable populations. Yet whether this relationship between transparency and welfare is as direct as commonly assumed remains an open question. Does increased information disclosure by public authorities necessarily lead to better health outcomes? More broadly, how does the information available to the population shape individual behavior and, ultimately, the dynamics of an epidemic?

The reason information matters during an epidemic is that individuals adjust their social interactions in response to the perceived epidemiological environment, trading off the expected cost of infection against the utility loss from reducing social activity. Yet the expected cost of infection depends on characteristics related to the epidemic that individuals largely do not know. First, they are uncertain whether they would develop symptoms if infected, thus some may remain asymptomatic and unknowingly spread the disease. Second, they lack precise knowledge of medical parameters such as symptom severity, contagiousness, and fatality rates. Finally, upon learning about the existence of an epidemic, they do not know its prevalence, which in turn prevents them from inferring the current state of the epidemic even conditional on individuals' behaviors. In the case of COVID-19, it became clear early on that asymptomatic individuals contributed to transmission, but the proportion of asymptomatic cases and parameters such as contagiousness, incubation period, and fatality rate remained uncertain for months. Moreover, the prevalence of the virus was initially unknown: retrospective evidence suggests that SARS-CoV-2 was already circulating in Europe as early as September 2019, thus well before official recognition.[2]

This paper explores how uncertainty about the characteristics of the epidemic shapes individuals' endogenous distancing behavior and how this affects the dynamics of an epidemic.

[2] Amendola et al. (2021) report viral RNA in a throat swab collected in Italy in December 2019 from a child whose symptoms began in November 2019. Similarly, Reis et al. (2022) found antibodies in samples collected as early as September-October 2019. In the USA, the first autochthonous case was diagnosed on February 26, 2020 in California. An earlier presence of SARS-CoV-2 in the United States was suspected based on the results of testing of sera obtained from routine blood donations to the American Red Cross. Of 7,389 samples analyzed, which had been collected between December 13 and 16, 2019, 106 (1.44%) were positive.

To address this question, we analyse a Susceptible-Infected-Recovered (SIR) epidemiological model in which myopic individuals choose their level of social activity at each point in time while being uncertain about their type (asymptomatic vs symptomatic) and the characteristics of the epidemic, i.e., the *epidemic state*. All individuals observe is whether they have experienced symptoms of the disease or not. They form subjective beliefs about both the epidemic state and their own type, which they continuously update based on their interactions, and make social-activity decisions accordingly. We provide analytical results on equilibrium distancing behavior dynamics and numerical simulations illustrating the effects of different information disclosure regimes.

In equilibrium, individuals adjust their level of social activity in response to infection dynamics, which are themselves shaped by aggregate social activity. This two-way dependence between behavior and prevalence renders equilibrium dynamics non-trivial. We identify a *severity condition* relative to epidemic states under which, when the epidemic state is known, equilibrium prevalence is monotonic and decreases from the onset of the epidemic. As prevalence falls, equilibrium social activity gradually increases until herd immunity is reached, after which individuals resume normal social activity. Our main result is that, even when this severity condition holds, introducing uncertainty about the epidemic state can overturn monotonic prevalence dynamics and generate multiple epidemic peaks. This occurs when individuals mistakenly believe the epidemic to be more severe than it actually is. Believing the epidemic to be severe, individuals self-isolate strongly at the onset and later resume normal social activity once they perceive that herd immunity has been reached, that is, when they believe the susceptible pool has fallen below the threshold at which prevalence would decline even without further self-isolation. If the true epidemic is milder, early self-isolation slows down contagion more than individuals anticipate, so that the susceptible population remains above the herd-immunity threshold when social activity resumes. This premature relaxation of distancing causes prevalence to increase again, generating a second epidemic wave.

We then study the welfare and mortality implications of uncertainty about the epidemic state through numerical simulations in a two-state environment that differs only in severity (mild versus severe). Although knowledge of the realized state enables individuals to flatten the curve in each case, transparency need not be optimal. A higher prior probability of the mild state shifts equilibrium activity from later to earlier stages of the epidemic,

generating an intertemporal reallocation whose effects on mortality and welfare depend on both the realized state and the epidemic horizon. The simulations reveal a sharp contrast. Transparency increases ex-post mortality and reduces ex-post payoffs in two configurations: when the epidemic is severe and vaccination arrives late, and when it is mild and vaccination arrives early. Ex ante, however, transparency strictly dominates secrecy in terms of welfare for all prior beliefs and vaccination horizons, while it may fail to do so in terms of mortality. These findings uncover a tension between welfare maximization and mortality minimization as policy objectives and provide a potential explanation for the strong public backlash against information withholding during the COVID-19 pandemic.

Related Literature

This paper belongs to the economic literature on infectious diseases that incorporates individual behavior into Susceptible-Infected-Recovered (SIR) models (see Auld et al. (2025) for a recent overview). A central insight of this literature is that epidemics may be self-limiting once individual behavior is taken into account. Most existing contributions rely on numerical simulations to characterize equilibrium dynamics and policy outcomes (see, among others, Reluga (2010), Fenichel et al. (2011), Fenichel (2013), Chen (2012), Toxvaerd (2020, 2022), Makris and Toxvaerd (2020), Baril-Tremblay et al (2021), Farboodi et al. (2021), Toxvaerd and Rowthorn (2022), Phelan and Toda (2022), Gans (2023) and Avery et al. (2024)). A smaller but growing set of papers derive analytical results on equilibrium behavior and epidemic dynamics (see, for instance, Dasaratha (2023), Carnehl et al. (2023), McAdams et al. (2023) and Rachel (2025)).

Our contribution is twofold. First, we introduce uncertainty about the characteristics of the epidemic into a behavioral SIR framework and study how it affects individual protective behavior. Second, and more importantly, we identify a novel mechanism through which uncertainty can generate multiple infection peaks, and we characterize this mechanism analytically. To the best of our knowledge, this is the first paper to provide analytical conditions under which belief-driven behavior alone can lead to multiple epidemic waves.

In the existing literature, multiple infection waves arise from mechanisms that are conceptually distinct from the one we identify. In a SIRS model with temporary immunity, Giannitsarou et al. (2021) show numerically that the disease may become endemic, with mitigation effort exhibiting dampened oscillations. Similar patterns emerge when births and

deaths are introduced into SIR models, as the susceptible population is continuously replenished (see e.g., Horan and Fenichel (2007)). A distinct strand focuses on policy-induced rebounds. Fenichel et al. (2011) document through simulations that lifting lockdown policies before herd immunity is reached may trigger a resurgence of infections. Rachel (2025) provides analytical results showing that the stronger the reduction in the basic reproduction number induced by the lockdown, the more severe the subsequent rebound. Carnehl et al. (2025) study distancing fatigue in a model with non-stationary isolation costs and show that interactions between fatigue and policy-induced discontinuities in isolation costs can generate second waves.

Finally, we establish that uncertainty may decrease mortality while the population may prefer transparency. In a closely related model, Toxvaerd (2022) shows through simulations that uncertainty about individual types may increase total infections but reduce symptomatic cases, resulting in higher welfare. Less closely related work examines uncertainty in models without endogenous distancing. Gollier (2020) shows that uncertainty about post-lockdown transmission reduces the optimal initial intensity of lockdown. Barnett et al. (2025) show that uncertainty about disease severity leads to stricter and longer quarantines, whereas uncertainty about economic costs leads to less stringent policies. Gubar et al. (2023) analyze vaccination decisions when part of the population is exposed to misinformation. Our results complement this literature by showing that, with endogenous distancing, uncertainty generates a tension for public authorities between welfare and mortality objectives.

The remainder of the paper is organized as follows. Section 2 presents the model. Section 3 characterizes the unique symmetric equilibrium. Section 4 studies how uncertainty shapes equilibrium prevalence dynamics and shows that infection rebounds may arise under uncertainty. Section 5 examines the value of information. Section 6 concludes and discusses the assumptions. All proofs are relegated to the Appendix.

2 An epidemiological model with uncertainty

An epidemic is characterized by both the medical parameters of the disease (severity of the symptoms, contagious rate, etc) and the penetration rate of the disease in the population at time 0, which will be formally defined later on. All these characteristics are described

by the *epidemic state* $\omega \in \Omega$, where Ω is a countable set of possible states. To model the spread of an epidemic, we use the Susceptible-Infected-Recovered (SIR) model by Kermack and McKendrick (1927), which we amend to introduce individual distancing behaviors and uncertainty about the characteristics of the epidemic.

The population. The population consists of a continuum of infinitely lived individuals, indexed by $j \in [0,1]$, who discount time $t \in [0,+\infty)$ at a common rate $r > 0$. At time 0, individuals are informed that a contagious disease is spreading in the population. They do not know the epidemic state and share a common prior belief μ^0 over Ω. Individuals are also informed that the entire population will be vaccinated at time $T > 0$ with a vaccine that is 100% effective. Infection may be totally asymptomatic, and ends by either recovery or death. Individuals are contagious as long as they are infected, and are immune to the disease after recovering. Whether an individual j develops symptoms or not when she is infected is an idiosyncratic characteristic described by her *type* $\theta_j \in \{\theta^a, \theta^s\}$. Individuals of type θ^s –the *symptomatic type*– experience the symptoms of the disease immediately after being infected, and individuals of type θ^a –the *asymptomatic type*– do not have symptoms in case of infection, thus never realize when they have been infected. Individuals do not know their type unless they are of type θ^s and catch the disease. The proportion of asymptomatic types in the population depends on the epidemic state and is denoted $\alpha^\omega \in (0,1)$.

At each time t and state $\omega \in \Omega$, the population is divided into five groups: the group of *susceptible* individuals who have never been infected by the disease, denoted by $S(t \mid \omega)$ and of size $s(t \mid \omega) := \int_{j \in S(t|\omega)} dj$; the group of *symptomatic infected* individuals who are infected with symptoms at time t, denoted by $I(t \mid \omega)$ and of size $i(t \mid \omega) := \int_{j \in I(t|\omega)} dj$; the group of *asymptomatic infected* individuals who are infected without symptoms at time t, denoted by $A(t \mid \omega)$ and of size $a(t \mid \omega) := \int_{j \in A(t|\omega)} dj$; the group of *recovered* individuals who already healed from the disease (with or without symptoms), denoted by $R(t \mid \omega)$ and of size $r(t \mid \omega) := \int_{j \in R(t|\omega)} dj$ and the group of *dead* individuals, denoted by $D(t \mid \omega)$ and of size $d(t \mid \omega) = 1 - s(t \mid \omega) - i(t \mid \omega) - a(t \mid \omega) - r(t \mid \omega)$. The *prevalence rate* of the disease at time t and state ω is $a(t \mid \omega) + i(t \mid \omega)$.

The composition of the population at time 0 defines the *initial penetration rate of the disease*. Without loss of generality for our purpose, we assume that individuals did not die or recover yet from the disease at time 0, i.e., $r(0 \mid \omega) = d(0 \mid \omega) = 0$ for each ω. Moreover,

by the law of large numbers, the initial proportion of infected individuals with symptoms represents a share $1-\alpha^\omega$ of the prevalence rate at time 0, hence $i(0 \mid \omega) = \frac{1-\alpha^\omega}{\alpha^\omega} a(0 \mid \omega)$. As, by definition, $s(0 \mid \omega) + a(0 \mid \omega) + i(0 \mid \omega) + r(0 \mid \omega) + d(0 \mid \omega) = 1$, the proportion of susceptible individuals at time 0 is $s(0 \mid \omega) = 1 - \frac{a(0|\omega)}{\alpha^\omega}$. Therefore, the initial penetration rate of the disease in state ω is determined by $a(0 \mid \omega)$.

Individuals can be contaminated by contact with an infected individual, hence respond to the threat of infection by choosing at each point in time a level of social activity. The dynamics of the epidemic thus depend on the behavior of the population. We assume that an individual who gets symptoms immediately stops social activity until the end of the symptoms, either to protect others, or simply because she is too sick to go out.[3] As a consequence, the disease is spread in the population by asymptomatic, infected, individuals, who still go on with meeting people without knowing that they are contagious. A strategy for player j is thus a measurable function $k_j : \mathbb{R}_+ \to [0,1]$, with the interpretation that $k_j(t)$ is the level of social activity of individual j at time t, absent symptoms by time t. Conversely, $1 - k_j(t)$ is the level of *self isolation* of individual j at time t.

Dynamics of the epidemic. In state $\omega \in \Omega$, a susceptible individual meeting an infected individual is infected at rate β^ω. Infection ends by recovery at rate γ^ω for an asymptomatic type, while, for a symptomatic type, it ends by recovery at rate γ_s^ω and by death at rate ν^ω. We assume $\gamma^\omega = \gamma_s^\omega + \nu^\omega$, i.e., the infection periods ends at the same rate for both types of individuals. This implies that the prevalence rate is $a(t \mid \omega)/\alpha^\omega$ for all t and ω.[4] The fraction of susceptible individuals evolves as follows: [5]

$$\dot{s}(t \mid \omega) = -\bar{k}_S(t \mid \omega) s(t \mid \omega) \bar{k}_A(t \mid \omega) a(t \mid \omega) \beta^\omega, \tag{1}$$

[3] The usual assumption in the literature is that infected individuals chose a constant social activity level $\bar{k} \in [0,1]$ during the symptomatic period. Assuming $\bar{k} = 0$ lightens analytic expressions.

[4] Indeed, the fact that $i(0 \mid \omega) = \frac{1-\alpha^\omega}{\alpha^\omega} a(0 \mid \omega)$ and the assumption $\gamma^\omega = \gamma_s^\omega + \nu^\omega$ together imply that $i(t \mid \omega) = \frac{1-\alpha^\omega}{\alpha^\omega} a(t \mid \omega)$ for all t. Therefore, the prevalence rate is $a(t \mid \omega) + i(t \mid \omega) = a(t \mid \omega)/\alpha^\omega$.

[5] To see why, consider the expected mass of individuals who become infected in the interval $[t, t+dt)$. The probability that a susceptible individual $s \in S(t \mid \omega)$ meets and is infected by some infected asymptomatic individual $a \in A(t \mid \omega)$ in the interval $[t, t+dt)$ is $k_s(t) k_a(t) \beta^\omega dt$. Therefore, the probability that individual s becomes infected in $[t, t+dt)$ is $k_s(t) \left(\int_{j \in A(t|\omega)} k_j(t) dj \right) \beta^\omega dt$, and the expected mass of newly infected individuals is $-(s(t+dt \mid \omega) - s(t \mid \omega)) = \int_{s \in S(t|\omega)} \left(k_s(t) (\int_{a \in A(t|\omega)} k_a(t) da) \beta^\omega dt \right) ds \equiv s(t \mid \omega) \bar{k}_S(t \mid \omega) a(t \mid \omega) \bar{k}_A(t \mid \omega) \beta^\omega dt$. The result follows from $\dot{s}(t \mid \omega) = \lim_{dt \to 0} \frac{s(t+dt|\omega) - s(t|\omega)}{dt}$.

where $\bar{k}_S(t \mid \omega) := \frac{1}{s(t|\omega)} \int_{j \in S(t|\omega)} k_j(t) dj$ and $\bar{k}_A(t \mid \omega) := \frac{1}{a(t|\omega)} \int_{j \in A(t|\omega)} k_j(t) dj$ denote the average behavior of susceptible and asymptomatic infected individuals, respectively. At each time t, the fraction of newly infected $-\dot{s}(t \mid \omega)$ is split between $A(t \mid \omega)$ and $I(t \mid \omega)$ in proportions α^ω and $1 - \alpha^\omega$, respectively. The other groups of the population thus evolve as follows:

$$\dot{a}(t \mid \omega) = -\alpha^\omega \dot{s}(t \mid \omega) - \gamma^\omega a(t \mid \omega), \tag{2}$$

$$\dot{i}(t \mid \omega) = -(1 - \alpha^\omega)\dot{s}(t \mid \omega) - \gamma^\omega i(t \mid \omega), \tag{3}$$

$$\dot{r}(t \mid \omega) = \gamma^\omega a(t \mid \omega) + \gamma_s^\omega i(t \mid \omega), \tag{4}$$

$$\dot{d}(t \mid \omega) = \nu^\omega i(t \mid \omega). \tag{5}$$

Uncertainty and beliefs. At time 0, individuals learn the existence of an epidemic, but do not know their own type nor the epidemic state. Moreover, they never observe the current value of the penetration of the disease, hence the only additional information they have at time t is whether or not they had symptoms before t. We denote by $p_j(t \mid \omega)$ the probability that player j is of type θ^s, conditionally on the state being ω and j having experienced no symptom by time t. A susceptible, type θ^s, individual j develops symptoms in $[s, s + ds)$ if she meets and is infected by an asymptomatic individual, which occurs with instantaneous probability $k_j(s)\bar{k}_A(s \mid \omega)a(s \mid \omega)\beta^\omega ds$ in state ω. By Bayes' rule, the probability that j is of type θ^s in state ω is therefore:

$$p_j(t \mid \omega) = \frac{(1 - \alpha^\omega)e^{-\int_0^t k_j(s)\bar{k}_A(s|\omega)a(s|\omega)\beta^\omega ds}}{\alpha^\omega + (1 - \alpha^\omega)e^{-\int_0^t k_j(s)\bar{k}_A(s|\omega)a(s|\omega)\beta^\omega ds}}. \tag{6}$$

After time 0, individuals continuously update their belief about the epidemic state upon the absence of symptoms, conditionally on the strategy profile of the population and the dynamic system formed by (1) and (2). The subjective belief of individual j at time t over Ω is denoted $\mu_j(t) : \Omega \to [0, 1]$, with the interpretation that $\mu_j(t)(\omega)$ is j's probability at time t that the state is ω, conditional on having experienced no symptom by time t. By Bayes' rule,

$$\mu_j(t)(\omega) = \frac{\mu^0(\omega)/(1 - p_j(t \mid \omega))}{\sum_{\omega' \in \Omega} \mu^0(\omega')/(1 - p_j(t \mid \omega'))}. \tag{7}$$

Payoffs. The flow utility of having social activity and being healthy is normalized to 0. As individuals self-isolate only to protect themselves from the disease, those who are vaccinated

or have already had symptoms do not self-isolate and get payoff 0. For other individuals, the level of social activity $k_j(t)$ costs $c_S(1-k_j(t))$ and might yield to the state-dependant desutility $-v_I^\omega < 0$ for the symptomatic types who get infected, which occurs with subjective probability $p_j(t \mid \omega)k_j(t)\bar{k}_A(t \mid \omega)a(t \mid \omega)\beta^\omega$. Therefore, if player j never experienced symptoms at time t, her instantaneous expected payoff in state ω is:

$$v_j(t \mid \omega) = -p_j(t \mid \omega)k_j(t)\bar{k}_A(t \mid \omega)a(t \mid \omega)\beta^\omega v_I^\omega - (1 - k_j(t))c_S. \tag{8}$$

The total discounted payoff to individual j in state ω is[6]

$$V_j(\omega) = \int_0^T e^{-rt} \underbrace{\frac{\alpha^\omega}{1 - p_j(t \mid \omega)}}_{\substack{\text{Probability that } j \text{ has no} \\ \text{symptom by time } t \text{ in } \omega.}} v_j(t \mid \omega)dt. \tag{9}$$

From individual j's perspective, the composition of the population, the characteristics of the disease, and the probability of being the symptomatic types are random variables that we denote by $s(t) : \omega \mapsto s(t \mid \omega)$, $a(t) : \omega \mapsto a(t \mid \omega)$, $\alpha : \omega \mapsto \alpha^\omega$, $\beta : \omega \mapsto \beta^\omega$, $\gamma : \omega \mapsto \gamma^\omega$, $v_I : \omega \mapsto v_I^\omega$ and $p_j(t) : \omega \mapsto p_j(t \mid \omega)$.

As a benchmark, we characterize the prevalence rate dynamics in the non-behavioral SIR model, where agents never self-isolate. Plugging $\bar{k}_S(t) = \bar{k}_A(t) = 1$ into (1) and (2) yields

$$\dot{a}(t \mid \omega) = a(t \mid \omega)\left(\alpha^\omega \beta^\omega s(t \mid \omega) - \gamma^\omega\right).$$

Since the susceptible population is monotonically decreasing by (1), prevalence either never increases (if $s(0 \mid \omega) < \gamma^\omega/(\alpha^\omega\beta^\omega)$) or increases until the pool of susceptible reaches

$$\frac{\gamma^\omega}{\alpha^\omega \beta^\omega},$$

called *herd immunity* threshold of epidemic ω. Once this threshold is reached, the susceptible pool is too small for the epidemic to sustain itself, even in the absence of self-isolation, and prevalence subsequently converges to zero. A key implication is that the non-behavioral SIR model generates a *single epidemic peak*, which is reached at time

$$\tau_\omega := \min\{t,\ s(t \mid \omega) \leq \frac{\gamma^\omega}{\alpha^\omega \beta^\omega} \text{ in the non-behavioral SIR model}\}.$$

[6] Only symptomatic individuals have symptoms, hence the probability in state ω that j never gets symptom in $[0, t)$ is $\alpha^\omega + (1 - \alpha^\omega)e^{-\int_0^t k_j(s)\bar{k}_A(s|\omega)a(s|\omega)\beta^\omega ds}$, which equals $\alpha^\omega/(1 - p_j(t \mid \omega))$ by (6).

We now turn to the analysis of prevalence dynamics when self-isolation is endogenous.

3 Equilibrium

We analyze the behavior of Bayesian and myopic individuals. Agents have perfect memory and continuously update their beliefs according to Bayes' rule, but do not internalize how their current actions affect future beliefs and payoffs. As a result, individuals choose their actions by maximizing their expected instantaneous payoff at each point in time. Formally, individual j's best response at time t to the strategy profile $\mathbf{k}$ solves the static optimization problem

$$\max_{k_j(t)\in[0,1]} \mathrm{E}_{\mu_j(t)}\left[v_j(t \mid \omega)\right] = \left(-k_j(t)\mathrm{E}_{\mu_j(t)}[p_j(t)\bar{k}_A(t)a(t)\beta v_I] - c_S(1-k_j(t))\right),$$

where the first term captures the expected instantaneous cost of infection associated with social activity, and the second term reflects the direct cost of self-isolation. By linearity of the objective function, the best response takes a threshold form:

$$k_j(t)\begin{cases} = 1 & \text{if } c_S > \mathrm{E}_{\mu_j(t)}[p_j(t)\bar{k}_A(t)a(t)\beta v_I], \\ = 0 & \text{if } c_S < \mathrm{E}_{\mu_j(t)}[p_j(t)\bar{k}_A(t)a(t)\beta v_I], \\ \in [0,1] & \text{if } c_S = \mathrm{E}_{\mu_j(t)}[p_j(t)\bar{k}_A(t)a(t)\beta v_I]. \end{cases} \tag{10}$$

Thus, myopic individuals compare the direct marginal cost of self-isolation c_S to the expected instantaneous cost of engaging in social activity. This yields a unique symmetric equilibrium, characterized as follows.

Proposition 1. *In the unique symmetric equilibrium, agents play the strategy $\hat{k}$ defined by*

$$\hat{k}(t) = \min\left\{1, \frac{c_S}{\mathrm{E}_{\hat{\mu}(t)}\left[\hat{p}(t)\hat{a}(t)\beta v_I\right]}\right\} \ \forall\, t \le T,$$

where $\hat{a}$, $\hat{p}$ and $\hat{\mu}$ are determined for each $\omega \in \Omega$ by the system:

$$\hat{a}(t \mid \omega) = \alpha^\omega - \alpha^\omega s(0 \mid \omega)e^{-\beta^\omega \int_0^t \hat{k}(u)^2\hat{a}(u|\omega)du} - \gamma^\omega \int_0^t \hat{a}(u \mid \omega)du, \tag{11}$$

$$\hat{p}(t \mid \omega) = \frac{1-\alpha^\omega}{1-\alpha^\omega + \alpha^\omega e^{\beta^\omega \int_0^t \hat{k}(u)^2\hat{a}(u|\omega)du}}, \tag{12}$$

$$\hat{\mu}(t)(\omega) = \mu^0(\omega)\frac{\alpha^\omega + (1-\alpha^\omega)e^{-\beta^\omega \int_0^t \hat{k}(u)^2\hat{a}(u|\omega)du}}{\alpha^\omega + (1-\alpha^\omega)\mathrm{E}_{\mu^0}[e^{-\beta \int_0^t \hat{k}(u)^2\hat{a}(u)du}]}. \tag{13}$$

Moreover, the size of the susceptible population in state ω is given by :

$$\hat{s}(t \mid \omega) = s(0 \mid \omega)e^{-\beta^\omega \int_0^t \hat{k}(u)^2\hat{a}(u|\omega)du}. \tag{14}$$

Proof. See Section A in the Appendix. □

Proposition 1 highlights the two-way dependence between behavior and prevalence: individuals' social activity affects prevalence dynamics, while evolving prevalence feeds back into behavior through perceived infection risk. Unlike in non-behavioral SIR models, this feedback between endogenous behavior and infection risk renders equilibrium prevalence dynamics non-trivial. In equilibrium, prevalence evolves according to

$$\dot{\hat{a}}(t \mid \omega) = \hat{a}(t \mid \omega) \left(\alpha^{\omega} \beta^{\omega} \hat{k}^2(t) \hat{s}(t \mid \omega) - \gamma^{\omega} \right). \tag{15}$$

Equation (15) implies that prevalence decreases if and only if

$$\hat{k}(t)^2 \hat{s}(t \mid \omega) < \frac{\gamma^{\omega}}{\alpha^{\omega} \beta^{\omega}},$$

that is, when individuals self-isolate sufficiently and/or when the susceptible pool has become sufficiently small. Consequently, even when $s(0 \mid \omega)$ is large, agents can "flatten the curve" by sharply reducing social activity at the onset of the epidemic and maintaining $\hat{k}^2(t)\hat{s}(t \mid \omega)$ below the herd-immunity threshold. Conversely, if $\hat{k}^2(t)\hat{s}(t \mid \omega)$ is non-monotonic and exceeds this threshold at some dates, prevalence may increase again, generating multiple epidemic peaks in equilibrium.

4 Uncertainty and prevalence

In this section, we explore how uncertainty shapes equilibrium prevalence dynamics. We begin by characterizing a *severity condition* under which, when the epidemic state is known, equilibrium prevalence is monotonic and decreases from the onset of the epidemic. We then show that, even when this condition is satisfied, introducing uncertainty about the epidemic state can overturn this monotonicity and generate multiple epidemic peaks.

Definition. The severity of epidemic ω is defined as

$$\eta^{\omega} := a(0 \mid \omega) \beta^{\omega} v_I^{\omega} (1 - \alpha^{\omega}).$$

The severity index aggregates the epidemiological and medical parameters that governed the perceived danger of the epidemic: it increases with the initial penetration rate, the transmission rate, the desutility from developing symptoms and the prior probability of being symptomatic.

4.1 Decreasing prevalence under certainty

We show that when individuals are certain that the epidemic is sufficiently severe, they respond by sharply reducing social activity at the onset of the outbreak. This strong initial behavioral response prevents the occurrence of multiple epidemic peaks. This result is related to earlier findings in Dasaratha (2023) and Carnehl et al. (2023), showing that sufficiently strong behavioral responses can make epidemics self-limiting.

For notational clarity, let $\hat{k}_\omega(t)$ and $\hat{a}_\omega(t \mid \omega)$ denote the equilibrium levels of social activity and asymptomatic prevalence when the epidemic state ω is common knowledge, that is, when $\mu^0(\omega) = 1$.

Theorem 1. *Fix $\omega \in \Omega$ and suppose $\mu^0(\omega) = 1$. There exists a positive cutoff $\underline{n}(\omega)$ such that, if the epidemic severity satisfies $\eta^\omega \geq \underline{n}(\omega)$, then $\hat{a}_\omega(t \mid \omega)$ is strictly decreasing with t on $[0, T]$. Moreover, there exists a finite time $\hat{\tau}_\omega > 0$ such that:*

(i) $\hat{k}_\omega(t)$ is increasing on $[0, \min\{\hat{\tau}_\omega, T\}]$ and $\hat{k}_\omega(t) = 1$ for $t \geq \min\{\hat{\tau}_\omega, T\}$;

(ii) the susceptible share at time $\hat{\tau}_\omega$ satisfies $\hat{s}(\hat{\tau}_\omega \mid \omega) < \frac{\gamma^\omega}{\alpha^\omega \beta^\omega}$.

Proof. See Section B in the Appendix. □

The intuition behind Theorem 1 is as follows. Upon learning that an epidemic of type ω has started, individuals choose their initial level of social activity as

$$\hat{k}_\omega(0) = \min\left\{1, \frac{c_S}{a(0 \mid \omega)\beta^\omega v_I^\omega (1 - \alpha^\omega)}\right\} = \min\left\{1, \frac{c_S}{\eta^\omega}\right\}.$$

Hence, the more severe the epidemic, the stronger the initial reduction in social activity. Substituting $\hat{k}_\omega(0)$ into the prevalence dynamics (15), we obtain that prevalence decreases at time 0 if and only if η^ω is sufficiently large.[7] Thus, when the epidemic is sufficiently severe, the endogenous reaction to the onset of the epidemic triggers an immediate decrease in infections. As prevalence falls over time, the perceived infection risk decreases, which induces a gradual relaxation of self-isolation in equilibrium: social activity $\hat{k}_\omega(t)$ increases and may eventually reach 1 before the arrival of the vaccine. When this occurs, the population has reached herd immunity because $\hat{s}(\hat{\tau}_\omega \mid \omega) < \gamma^\omega/(\alpha^\omega \beta^\omega)$, thus prevalence continues to decrease even in the absence of further self-isolation.

[7] Indeed, if $\eta^\omega > c_S$, $\hat{k}_\omega(0) = \frac{c_S}{\eta^\omega}$ and $\alpha^\omega \beta^\omega \hat{k}_\omega(0)^2 s(0 \mid \omega) - \gamma^\omega < 0 \Leftrightarrow (\eta^\omega)^2 + \frac{c_S^2}{v_I^\omega (1-\alpha^\omega)\gamma^\omega}\eta^\omega - \frac{\alpha^\omega \beta^\omega}{\gamma^\omega} c_S^2 > 0$.

The explicit expression of the severity cutoff $\underline{n}(\omega)$ is cumbersome and therefore relegated to the Appendix. Nevertheless, we can evaluate the impact of several parameters of interest on the severity condition.

Proposition 2. *Let ω satisfying the severity condition $\eta^\omega \geq \underline{n}(\omega)$. The state ω still satisfies the severity condition if, ceteris paribus,*

- *$a(0 \mid \omega)$, γ^ω, or v_I^ω increases;*
- *α^ω or c_S decreases.*

The effect of an increase in β^ω is ambiguous.

Proof. See Section B in the Appendix. □

Several of these comparative statics are intuitive. A lower self-isolation cost c_S or a higher disutility of symptoms v_I^ω strengthens incentives to reduce social activity at the onset of the epidemic, making it more likely that prevalence decreases immediately. Similarly, a higher initial penetration rate $a(0 \mid \omega)$ or probability of being symptomatic $(1 - \alpha^\omega)$ increases the perceived infection risk, which amplifies the initial behavioral response and favors the decrease in prevalence. By shortening the expected duration of infectiousness, a higher recovery rate γ^ω increases the herd-immunity threshold. As a result, a given level of self-isolation is more likely to suffice to put prevalence on a declining path. Finally, the ambiguous effect of the transmission rate β^ω reflects two countervailing forces. On the one hand, a higher transmission rate raises infection risk and therefore induces stronger initial self-isolation. On the other hand, it also decreases the herd-immunity threshold, making it more difficult to keep prevalence below the threshold required for monotonic decline. The net effect depends on which of these two forces dominates.

4.2 Multiple epidemic peaks under uncertainty

When individuals are uncertain about the epidemic state, their behavior depends on their prior belief μ^0 over Ω and may differ substantially from the behavior they would adopt under full information. In this section, we show that such belief-driven distortions in behavior can generate multiple epidemic peaks. Specifically, we prove that there exist epidemic states in

which prevalence is monotonic and decreasing under certainty, yet becomes non-monotonic under uncertainty.

Theorem 2. *There exist ω_L and ω_H such that, if T is large enough,*

(i) when the state is ω_L and individuals know the state, equilibrium prevalence $\hat{a}_{\omega_L}(t \mid \omega_L)$ decreases on $[0, T]$;

(ii) when the state is ω_L and $\mu^0(\omega_H)$ is sufficiently large, equilibrium prevalence $\hat{a}(t \mid \omega_L)$ decreases on some interval $[0, \underline{t})$ and increases on some interval $[\underline{t}, \bar{t})$, with $\bar{t} < T$.

To establish the result, we construct sufficient conditions on a pair of epidemic states (ω_L, ω_H) under which prevalence in state ω_L is monotonic under certainty, yet becomes non-monotonic under uncertainty. The argument proceeds in three steps. First, we ensure that prevalence in state ω_L is decreasing when individuals know the true state. This is guaranteed by assuming that the severity condition holds in ω_L:

$$\textbf{Condition (a):}\ \eta^{\omega_L} \geq \underline{n}(\omega_L).$$

Under this condition, Theorem 1 implies that, when $\mu^0(\omega_L) = 1$, individuals self-isolate sufficiently at the onset of the epidemic to place prevalence on a strictly decreasing path. Second, we ensure that, when state ω_H is sufficiently likely, equilibrium self-isolation at the onset of the epidemic is higher than under certainty of state ω_L. This is achieved by assuming that epidemic ω_H is more severe than ω_L:

$$\textbf{Condition (b):}\ \eta^{\omega_H} \geq \eta^{\omega_L}.$$

Indeed, condition (b) is equivalent to $\hat{k}_{\omega_H}(0) < \hat{k}_{\omega_L}(0)$. Therefore, when beliefs place substantial weight on ω_H, individuals self-isolate more at the onset than they would under certainty in ω_L. Since prevalence is decreasing in ω_L when individuals play $\hat{k}_{\omega_L}(0)$, it decreases *a fortiori* at time 0 when they play $\hat{k}_{\omega_H}(0)$. Third, we ensure that uncertainty leads to a premature relaxation of self-isolation in state ω_L. To this end, we assume that epidemic ω_H is sufficiently severe for individuals to stop self-isolating in finite time under certainty:

$$\textbf{Condition (c):}\ \eta^{\omega_H} \geq \underline{n}(\omega_H).$$

By Theorem 1, under condition (c) individuals cease self-isolation at time $\hat{\tau}_{\omega_H}$ when they know that the state is ω_H. We further assume that this time precedes the herd-immunity threshold in the non-behavioral SIR model in state ω_L:

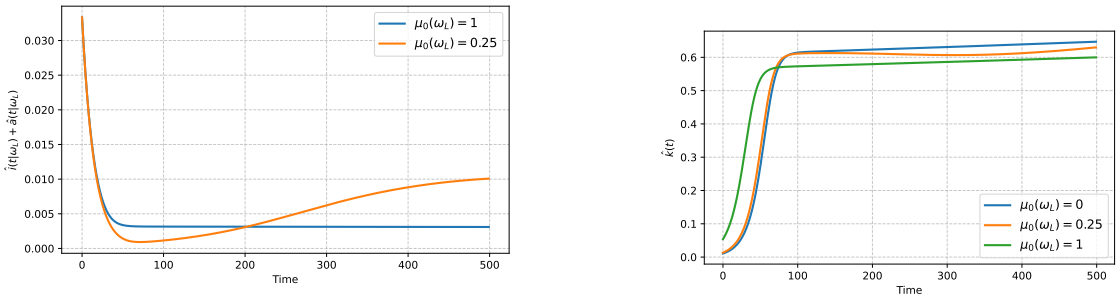
Condition (d)**:** $\hat{\tau}_{\omega_H} < \tau_{\omega_L}$.

Condition (d) implies that, when the true state is ω_L, resuming normal social activity at time $\hat{\tau}_{\omega_H}$ occurs before herd immunity would be reached without any self-isolation. Since self-isolation delays the depletion of the susceptible pool relative to the non-behavioral SIR benchmark, this relaxation is premature in state ω_L and causes prevalence to increase again, generating a second infection peak.

Finally, we establish that the set of values of $a(0 \mid \omega_L), \alpha^L, \beta^L, \gamma^L, v_I^L, a(0 \mid \omega_H), \alpha^H, \beta^H, \gamma^H$ and v_I^H that satisfy (a), (b), (c) and (d) is an open subset of $\mathbb{R}^{10}$.

4.3 Illustration

To illustrate our results, we simulate equilibrium dynamics under two possible epidemic states, ω_L and ω_H, which differ only by the initial penetration rate, with $a(0 \mid \omega_L) < a(0 \mid \omega_H)$. All other parameters ($\alpha^\omega$, β^ω, γ^ω, ν^ω and v_I^ω), are identical across states, and calibrated to COVID-19.[8] Thus, we refer to ω_H as the severe epidemic, and to ω_L as the mild epidemic.



(a) Prevalence rate

(b) Social interaction

Figure 1: Dynamics for $T = 500$.

Figure 1a displays the prevalence dynamics in state ω_L under two beliefs scenarios: when individuals know that the state is ω_L ($\mu^0(\omega_L) = 1$) and, under uncertainty, when they assign a high probability to ω_H ($\mu^0(\omega_L) = 0.25$). When the state is known, prevalence decreases on the entire interval $[0, T]$. When the state ω_H is likely, however, prevalence becomes non-monotonic, and exhibits two infection peaks over horizon $T = 500$.

[8] Specifically, we set $a(0 \mid \omega_L) = 0.01$ and $a(0 \mid \omega_H) = 0.05$, $\alpha^\omega = 0.3$, $\beta^\omega = 0.71$, $\gamma^\omega = 1/15$, $\nu^\omega = 0.033\%$, and arbitrarily set the cost parameters to $r = 0.014\%$, $c_S = 1$ and $v_I^\omega = -37$. See Section D for a detailed discussion of the calibration strategy.

Figure 1b illustrates the underlying mechanism by comparing the equilibrium paths of social activity under the three beliefs scenarios $\mu^0(\omega_L) = 0$, $\mu^0(\omega_L) = 0.25$ and $\mu^0(\omega_L) = 1$. Individuals self-isolate more at the onset and resume normal social activity earlier when they know that the state is ω_H rather than when they know the state is ω_L. The reason is that the initial susceptible pool is smaller in ω_H ($s(0 \mid \omega_H) < s(0 \mid \omega_L)$), thus herd immunity is reached faster. When the prior belief is $\mu^0(\omega_L) = 0.25$, the equilibrium social-activity path lies between $\hat{k}_{\omega_H}(t)$ and $\hat{k}_{\omega_L}(t)$, but remains closer to $\hat{k}_{\omega_H}(t)$ given the relatively high probability assigned to state ω_H. Consequently, in state ω_L, individuals engage in more social activity from around $t = 70$ onward under the belief $\mu^0(\omega_L) = 0.25$ than under certainty. Since herd immunity has not yet been reached in state ω_L at that time, this premature relaxation causes prevalence to increase again around the same date, generating the second peak shown in Figure 1a.

5 The value of information

Figure 1a illustrates a configuration in which knowledge that the epidemic is mild enables individuals to flatten the curve through self-isolation, whereas a second wave arises when individuals assign a sufficiently high probability to the severe state. The figure also shows that, prior to the second wave, self-isolation is higher when the severe state is perceived as more likely. Since self-isolation is privately costly yet reduces infection risk, transparency affects behavior through different channels. As a result, the mortality and welfare effects of disclosing the epidemic state are a priori ambiguous, as they depend both on the realized state and on the epidemic horizon.

In this section, we study the value of information using numerical simulations in the two epidemic states ω_L and ω_H. We compare two polar communication policies to which the public authority commits ex ante, under both early ($T = 100$) and late ($T = 500$) vaccine arrival. Under the *transparency* policy, the authority discloses the realized epidemic state in each state, whereas under the *secrecy* policy, they never disclose the realized state. Although these simulations are only illustrative, they highlight how transparency interacts with belief

dynamics and vaccine timing to shape both welfare and mortality outcomes.[9]

5.1 Death toll

Symptomatic individuals infected before the arrival of the vaccine may die after time T. Therefore, the ex-post share of the population that ultimately dies from the disease in state ω under secrecy is[10]

$$\hat{d}(\omega) := \lim_{t\to+\infty} \hat{d}(t \mid \omega) = \nu^\omega \frac{1-\alpha^\omega}{\alpha^\omega} \left(\int_0^T \hat{a}(t \mid \omega) dt + \frac{1}{\gamma^\omega} \hat{a}(T \mid \omega) \right).$$

Under transparency, authorities commit ex ante to disclose the realized state in each state. The prevalence rate in ω is thus given by $\hat{a}_\omega(t \mid \omega)$, which yields the ex-post death toll under transparency:

$$\hat{d}^{tr}(\omega) := \nu^\omega \frac{1-\alpha^\omega}{\alpha^\omega} \left(\int_0^T \hat{a}_\omega(t \mid \omega) dt + \frac{1}{\gamma^\omega} \hat{a}_\omega(T \mid \omega) \right).$$

In our simulations, when the vaccine arrives early ($T = 100$), $\hat{d}(\omega)$ increases with $\mu^0(\omega_L)$ in both epidemic states (see Figure 2a). When the vaccine arrives late ($T = 500$), it decreases with $\mu^0(\omega_L)$ in state ω_L, whereas it displays a U-shaped relationship in state ω_H, attaining its maximum at $\mu^0(\omega_L) = 0$ (see Figure 2b).

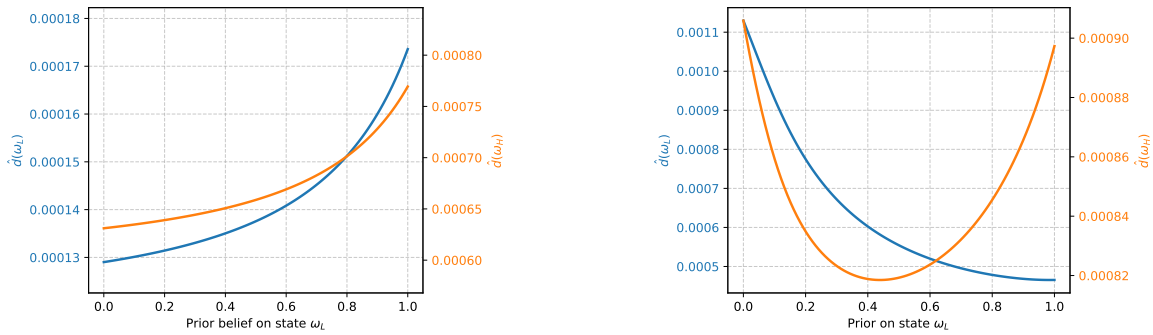
To interpret these patterns, it is useful to recall that under certainty individuals self-isolate more intensively in the early phase of a severe epidemic ω_H than in a mild epidemic ω_L, and less intensively in the late phase (see Figure 1b). As a result, increasing the prior probability of ω_L shifts self-isolation effort from the early to the late phase of the epidemic.

When the vaccine arrives early, mortality is driven by early infections. As a higher $\mu^0(\omega_L)$ increases early infections by reducing early self-isolation, $\hat{d}(\omega)$ increases with $\mu^0(\omega_L)$ in both epidemic states when $T = 100$.

When the vaccine arrives late, mortality reflects both early and late infections. For low values of $\mu^0(\omega_L)$, individuals self-isolate little in the late phase, leading to substantial late

[9] In a different setting, Phelan and Toda (2022) also show that optimal lockdown policies depend on vaccine timing: when vaccination arrives far in the future, a planner with perfect enforcement may optimally encourage activity before herd immunity is reached.

[10] By definition, $\hat{d}(\omega) = \lim_{t\to\infty} \hat{d}(t \mid \omega) = \int_0^\infty \dot{\hat{d}}(t \mid \omega) dt$. Yet, by equation (5), $\dot{\hat{d}}(t \mid \omega) = \nu^\omega \frac{1-\alpha^\omega}{\alpha^\omega} \hat{a}(t \mid \omega)$, thus the equilibrium death toll in state ω is: $\lim_{t\to\infty} \hat{d}(t \mid \omega) = \nu^\omega \frac{1-\alpha^\omega}{\alpha^\omega} \int_0^\infty \hat{a}(t \mid \omega) dt$. Moreover, for all $t \geq T$, $\dot{\hat{s}}(t \mid \omega) = 0$, hence $\dot{\hat{a}}(t \mid \omega) = -\gamma^\omega \hat{a}(t \mid \omega)$ and $\hat{a}(t \mid \omega) = \hat{a}(T \mid \omega) e^{-\gamma^\omega (t-T)}$. It follows that $\int_0^\infty \hat{a}(t \mid \omega) dt = \int_0^T \hat{a}(t \mid \omega) dt + \frac{1}{\gamma^\omega} \hat{a}(T \mid \omega)$, thus the result.



(a) $T = 100$

(b) $T = 500$

Figure 2: Death toll

infections. A marginal increase in $\mu^0(\omega_L)$ raises early infections but reduces late infections, and the latter effect dominates. As a result, $\hat{d}(\omega)$ decreases with $\mu^0(\omega_L)$ for low prior beliefs. In the mild state ω_L, this mechanism continues to operate even for high values of $\mu^0(\omega_L)$, because the increase in early infections remains limited. In contrast, in the severe state ω_H, once $\mu^0(\omega_L)$ exceeds a threshold, the relaxation of early self-isolation becomes sufficiently strong that the resulting increase in early infections outweighs the reduction in late mortality. Consequently, $\hat{d}(\omega_H)$ increases with $\mu^0(\omega_L)$ for high values of $\mu^0(\omega_L)$.

The policy implications of these findings are the following. When public authorities anticipate an early arrival of the vaccine, disclosure is desirable in the severe epidemic but not in the mild one, for all prior beliefs. In sharp contrast, when vaccine development is slow, disclosure is desirable only in the mild state and not in the severe one, for all prior beliefs.

Whether the expected gain from disclosure in one state outweighs the expected loss in the other depends on the timing of vaccine arrival (see Figure 3). Intuitively, a higher prior probability of ω_L, amplifies the benefits of secrecy when the vaccine arrives early, but amplifies the benefits of transparency when vaccination is delayed. In our simulations, secrecy strictly dominates transparency for all prior beliefs when $T = 100$. By contrast, when $T = 500$, secrecy dominates transparency only when the mild state is sufficiently unlikely. Formally,

$$\textbf{Early vaccine}: \quad \hat{d}^{tr}(\omega_L) > \hat{d}(\omega_L), \quad \hat{d}^{tr}(\omega_H) < \hat{d}(\omega_H), \quad \mathrm{E}_{\mu^0}[\hat{d}(\omega) - \hat{d}^{tr}(\omega)] < 0,$$

$$\textbf{Late vaccine}: \quad \hat{d}^{tr}(\omega_L) < \hat{d}(\omega_L), \quad \hat{d}^{tr}(\omega_H) > \hat{d}(\omega_H), \quad \mathrm{E}_{\mu^0}[\hat{d}(\omega) - \hat{d}^{tr}(\omega)] < 0 \text{ iff } \mu^0(\omega_L) > 0.57.$$

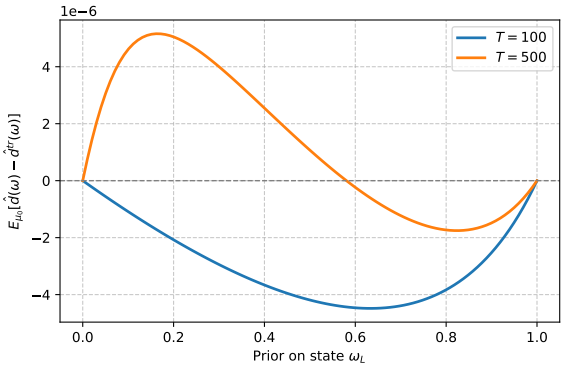


Figure 3: Value of information for mortality

5.2 Payoffs

Under secrecy, the total discounted payoff in state ω is[11]

$$\hat{V}(\omega) = -\int_0^T e^{-rt} \frac{\alpha^\omega}{1-\hat{p}(t \mid \omega)} (\hat{p}(t \mid \omega)\hat{k}(t)^2 \hat{a}(t \mid \omega)\beta^\omega v_I^\omega + (1-\hat{k}(t))c_S)dt.$$

Under transparency, equilibrium social activity in state ω is $\hat{k}_\omega(t)$, and the probability of being the symptomatic type is $\hat{p}_\omega(t \mid \omega) := (1-\alpha^\omega)/(1-\alpha^\omega + \alpha^\omega e^{\beta^\omega \int_0^t \hat{k}_\omega(u)^2 \hat{a}_\omega(u|\omega)du})$, which yields the total discounted payoff

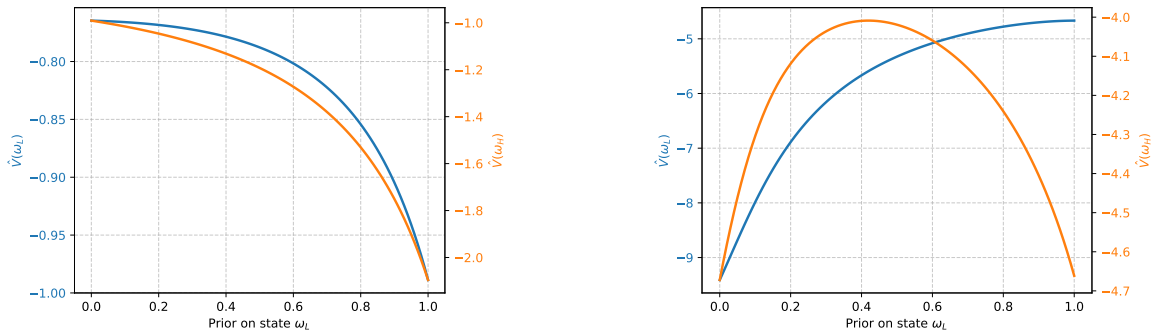
$$\hat{V}^{tr}(\omega) = -\int_0^T e^{-rt} \frac{\alpha^\omega}{1-\hat{p}_\omega(t \mid \omega)} (\hat{p}_\omega(t \mid \omega)\hat{k}_\omega(t)^2 \hat{a}_\omega(t \mid \omega)\beta^\omega v_I^\omega + (1-\hat{k}_\omega(t))c_S)dt.$$

In our simulations, when the vaccine arrives early, total equilibrium payoffs $\hat{V}(\omega)$ decrease with $\mu^0(\omega_L)$ in both epidemic states (see Figure 4a). When the vaccine arrives late, $\hat{V}(\omega_L)$ increases with $\mu^0(\omega_L)$, while $\hat{V}(\omega_H)$ exhibits a non-monotonic, bell-shaped relationship, attaining its minimum at $\mu^0(\omega_L) = 0$ (see Figure 4b).

These patterns can be understood through the reallocation of social activity induced by beliefs. An increase in $\mu^0(\omega_L)$ shifts equilibrium activity from the late to the early phase of the epidemic. This intertemporal reallocation generates opposing effects on payoffs. In the short run, higher activity raises infection risk but reduces self-isolation costs. In the long run, lower activity reduces infection risk but increases self-isolation costs. The net effect of this trade-off depends on the epidemic horizon.

When the vaccine arrives early ($T = 100$), short-run outcomes are decisive. The simulations show that $\hat{V}(\omega)$ decreases with $\mu^0(\omega_L)$ in both states. Hence, when the horizon is

[11] This expression is obtained by plugging $k_j(t) = \bar{k}_A(t \mid \omega) = \hat{k}(t)$ and $p_j(t \mid \omega) = \hat{p}(t \mid \omega)$ into (8) and (9).



(a) $T = 100$

(b) $T = 500$

Figure 4: Total discounted payoff

short, the increase in early infection risk dominates the savings in self-isolation costs. When the vaccine arrives late ($T = 500$), cumulative self-isolation costs become quantitatively important and may reverse the trade-off. In this case, $\hat{V}(\omega_L)$ increases with $\mu^0(\omega_L)$, indicating that in the mild state the reduction in long-run infection risk outweighs the increase in distancing costs. In the severe state ω_H, the effect is non-monotonic. For low values of $\mu^0(\omega_L)$, $\hat{V}(\omega_H)$ increases with the prior belief, whereas for sufficiently high values it decreases, reaching a maximum at an interior belief. Figure 1b clarifies this mechanism. An increase from $\mu^0(\omega_L) = 0$ to $\mu^0(\omega_L) = 0.25$ primarily affects long-run behavior, with little impact on short-run activity. By contrast, increasing $\mu^0(\omega_L)$ from 0.25 to 1 substantially alters also short-run behavior, generating a pronounced intertemporal reallocation of activity. This pattern reflects a general property: for low prior beliefs, marginal changes in $\mu^0(\omega_L)$ mainly affect long-run activity, whereas for high prior beliefs they also distort short-run behavior. As in state ω_L, the reduction in long-run infection risk outweighs the increase in distancing costs when $\mu^0(\omega_L)$ is small. For high values of $\mu^0(\omega_L)$, by contrast, the induced shift toward early activity substantially raises short-run infection risk and necessitates stronger distancing later on, so that the overall welfare effect becomes negative.

These findings yield the following policy implications. When the vaccine arrives early, welfare is maximized by disclosing the epidemic state in ω_H and withholding information in ω_L, regardless of the prior belief. When vaccine development is slow, the opposite holds: individuals prefer disclosure in ω_L and uncertainty in ω_H, for all μ^0. Despite these state-contingent effects, the expected welfare gain from disclosure in one state outweighs the expected welfare

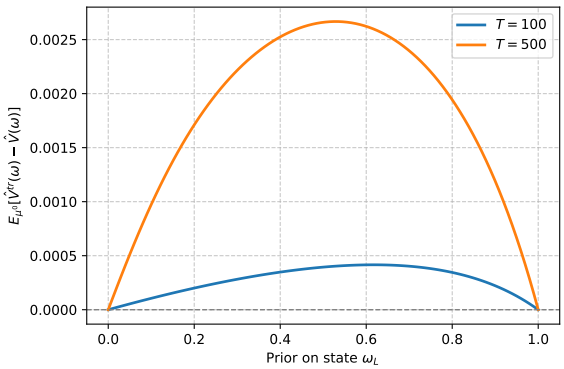


Figure 5: Value of information for payoffs

loss from disclosure in the other for all prior beliefs and both vaccine arrival dates (see Figure 5). Consequently, transparency dominates secrecy in terms of payoffs, with a larger value of information when the vaccine arrives late. Formally, for all prior beliefs,

$$\textbf{Early vaccine}: \quad \hat{V}^{tr}(\omega_L) < \hat{V}(\omega_L), \quad \hat{V}^{tr}(\omega_H) > \hat{V}(\omega_H), \quad E_{\mu^0}[\hat{V}^{tr}(\omega) - \hat{V}(\omega)] > 0,$$
$$\textbf{Late vaccine}: \quad \hat{V}^{tr}(\omega_L) > \hat{V}(\omega_L), \quad \hat{V}^{tr}(\omega_H) < \hat{V}(\omega_H), \quad E_{\mu^0}[\hat{V}^{tr}(\omega) - \hat{V}(\omega)] > 0.$$

6 Conclusion

This paper studies how uncertainty about epidemic characteristics shapes individual distancing behavior and, in turn, epidemic dynamics. We show that uncertainty can generate non-monotonic behavioral responses and infection rebounds: when agents overestimate epidemic severity, they initially self-isolate excessively and relax distancing before herd immunity is reached, thereby triggering a second infection peak. Overall, our analysis highlights a fundamental tension between mortality minimization and welfare maximization in epidemic management. By reshaping the intertemporal allocation of risk-taking, communication policies can either amplify or dampen epidemic waves, sometimes in counterintuitive ways. These findings suggest that transparency is not unambiguously welfare-improving and that its desirability depends critically on epidemic severity, prior beliefs, and the expected duration of the crisis.

Our analysis relies on a parsimonious model that abstracts from several features of real-world epidemics. We briefly discuss the robustness of our results to relaxing some of these

assumptions.

Horizon length. We assume a finite horizon, interpreted as the arrival date of a vaccine. In the case of new diseases, such as COVID-19, the arrival time of the vaccine is uncertain. However, under the myopia assumption, the analytical results of Sections 3 and 4 extend to infinite or stochastic horizon, as behavior depends only on current incentives.

Pre-symptomatic individuals. Many diseases, including COVID-19, feature a pre-symptomatic phase during which individuals are infectious but not yet symptomatic. Under the myopia assumption, asymptomatic agents can be reinterpreted as pre-symptomatic individuals who self-isolate upon symptom onset, leaving equilibrium behavior unchanged.

Waning immunity. For most viruses, immunity is not permanent. Introducing waning immunity would affect agents who have previously experienced symptoms. When immunity loss is sufficiently slow relative to the epidemic horizon, the share of agents who have lost immunity is relatively small, and the equilibrium dynamics remain close to those we identify. By continuity, uncertainty-driven infection rebounds are likely to persist.

Information about prevalence. In our model, individuals observe their own symptoms but not aggregate prevalence. In reality, individuals form beliefs about prevalence also on the basis of their social environment. Our model could be extended to include a noisy public signal about prevalence. Whether uncertainty-driven second peaks persist in equilibrium would then depend on the speed of learning in the population, thus on the informativeness of this signal.

References

[1] Amendola A., et al, (2021), Evidence of SARS-CoV-2 RNA in an oropharyngeal swab specimen, Milan, Italy, early december 2019, *Emerging Infectious Diseases*, 27(2): 648-650.

[2] Auld M., Fenichel E., and Toxvaerd F., (2025), The economics of infectious diseases, *Journal of Economic Literature*, 63(4): 12811330.

[3] Avery C., Chen F., and McAdams D., (2024), Steady-state social distancing and vaccination, *American Economic Review: Insights*, 6(1): 1-19.

[4] Baril-Tremblay D., Marlats C., and Ménager L., (2021), Self-isloation, *Journal of Mathematical Economics*, 93: 102483.

[5] Barnett M., Buchak G., and Yannelis C., (2025), Epidemic responses under uncertainty, NBER Working Paper 27289.

[6] Camus, A., (1947), *La peste*, Paris, Gallimard.

[7] Carnehl C., Fukuda S., and Kos N., (2023), Epidemics with behavior, *Journal of Economic Theory*, 207(C).

[8] Carnehl C., Fukuda S., and Kos N., (2025), Time-varying cost of distancing: distancing fatigue, holidays and lockdowns, mimeo.

[9] Chen F., (2012), A mathematical analysis of public avoidance behavior during epidemics using game theory, *Journal of Theoretical Biology*, 302: 18-28.

[10] Dasaratha K., (2023), Virus Dynamics with Behavioral Responses, *Journal of Economic Theory*, 214(C).

[11] Farboodi M., Jarosch G., and Shimer R., (2021), Internal and external effects of social distancing in a pandemic, *Journal of Economic Theory*, 196: 105293.

[12] Fenichel E., (2013), Economic considerations for social distancing and behavioral based policies during an epidemic, *Journal of Health Economics*, 32(2): 440-451.

[13] Fenichel E., et al, (2011), Adaptive human behavior in epidemiological models, *Proceedings of the National Academy of Sciences*, 108(15): 6306-6311.

[14] Gans J., (2023), Vaccine hesitancy, passports, and the demand for vaccination, *International Economic Review*, 64(2): 641-652.

[15] Giannitsarou C., Kissler S., and Toxvaerd F., (2021), Waning immunity and the second wave: some projections for SARS-CoV-2, *American Economic Review: Insights*, 3(3): 321-338.

[16] Gollier C., (2020), Pandemic economics: optimal dynamic confinement under uncertainty and learning, *Geneva Risk Insurance Review*, 45(2): 80-93.

[17] Gubar E., Taynitskiy V., and Dahmouni Y., (2023), The impact of fake news on infection dynamics in pandemic control: an evolutionary SIR model, *IFAC PapersOnLine*, 5(2): 1778-1783.

[18] Gudbjartsson D., et al, (2020), Humoral immune response to SARS-CoV-2 in Iceland, *New England Journal of Medicine*, 383(18): 1724-1734.

[19] Horan R., and Fenichel E., (2007), Economics and ecology of managing emerging infectious animal diseases, *American Journal of Agricultural Economics*, 89: 1232-1238.

[20] Kermack W., and McKendrick A., (1927), A contribution to the mathematical theory of epidemics, *Proceedings of the Royal Society of London.*, 115(772): 700-721.

[21] Makris R., and Toxvaerd F., (2020), Great expectations: social distancing in anticipation of pharmaceutical innovations, Faculty of Economics, University of Cambridge.

[22] McAdams D., Song Y., and Zou D., (2023), Equilibrium social activity during an epidemic, *Journal of Economic Theory*, 207: 105591.

[23] Phelan T., and Toda A., (2022), Optimal epidemic control in equilibrium with imperfect testing and enforcement, *Journal of Economic Theory*, 206(C).

[24] Pollán M., et al, (2020), Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study, *The Lancet*, 396(10250): 535-544.

[25] Rachel L., (2025), The second wave, *Review of Economic Design*, 29: 87113.

[26] Reis J., et al, (2022), Covid-19: early cases and disease spread, *Annals of Global Health*, 88(1): 113.

[27] Reluga T., (2010), Game theory of social distancing in response to an epidemic, *PLoS Compututational Biology*, 6(5): e1000793.

[28] Remuzzi A., and Remuzzi G., (2020), COVID-19 and Italy: what next?, *The Lancet*, 395: 1225-1228.

[29] Toxvaerd F., (2020), Equilibrium social distancing, Faculty of Economics, University of Cambridge.

[30] Toxvaerd F., (2022), Silent spreaders: behavior and equilibrium under asymptomatic infection, CEPR Discussion Paper No. 17553.

[31] Toxvaerd F., and Rowthorn R., (2022), On the management of population immunity, *Journal of Economic Theory*, 204(C).

[32] Verity R., et al., (2020), Estimates of the severity of COVID-19 disease, *Lancet Infectious Diseases*, 20(6): 669-677.

[33] Weder F., and Courtois C., (2022), Differences in universal health coverage and governments' COVID-19 communication: a global comparative analysis., *Front. Commun.* 7:1080948.

Appendix

Let us recall here the main equations of the model:

$$\dot{s}(t \mid \omega) = -\bar{k}_S(t \mid \omega)s(t \mid \omega)\bar{k}_A(t \mid \omega)a(t \mid \omega)\beta^\omega \tag{1}$$

$$\dot{a}(t \mid \omega) = -\alpha^\omega \dot{s}(t \mid \omega) - \gamma^\omega a(t \mid \omega) \tag{2}$$

$$p_j(t \mid \omega) = \frac{(1-\alpha^\omega)e^{-\int_0^t k_j(s)\bar{k}_A(s|\omega)a(s|\omega)\beta^\omega ds}}{\alpha^\omega + (1-\alpha^\omega)e^{-\int_0^t k_j(s)\bar{k}_A(s|\omega)a(s|\omega)\beta^\omega ds}} \tag{6}$$

$$\mu_j(t)(\omega) = \frac{\mu^0(\omega)/(1-p_j(t \mid \omega))}{\sum_{\omega' \in \Omega} \mu^0(\omega')/(1-p_j(t \mid \omega'))} \tag{7}$$

Differentiating (6), we obtain

$$\dot{p}_j(t \mid \omega) = -p_j(t \mid \omega)(1-p_j(t \mid \omega))k_j(t)\bar{k}_A(t \mid \omega)a(t \mid \omega)\beta^\omega \tag{16}$$

A Proof of Proposition 1

Let us first express the values of $\hat{s}$, $\hat{a}$, and $\hat{\mu}$ obtained when the population plays some symmetric profile $\hat{k}$. If $k_j(t) = \hat{k}(t)$ for all $j \in S(t \mid \omega) \cup A(t \mid \omega)$, then $\bar{k}_S(t \mid \omega) = \bar{k}_A(t \mid \omega) = \hat{k}(t)$, and there is $\hat{p}$ such that $p_j = \hat{p}$ for all j. Plugging this into (1), (2), and (6), we obtain:

$$\dot{\hat{s}}(t \mid \omega) = -\hat{k}(t)^2 \hat{s}(t \mid \omega)\hat{a}(t \mid \omega)\beta^\omega \tag{17}$$

$$\dot{\hat{a}}(t \mid \omega) = -\alpha^\omega \dot{\hat{s}}(t \mid \omega) - \gamma^\omega \hat{a}(t \mid \omega) \tag{18}$$

$$\hat{p}(t \mid \omega) = \frac{(1-\alpha^\omega)e^{-\int_0^t \hat{k}(t)^2 \hat{a}(t|\omega)\beta^\omega}}{\alpha^\omega + (1-\alpha^\omega)e^{-\int_0^t \hat{k}(t)^2 \hat{a}(t|\omega)\beta^\omega}} \tag{19}$$

Rewriting (17) as follows:

$$\frac{\dot{\hat{s}}(t \mid \omega)}{\hat{s}(t \mid \omega)} = -\hat{k}(t)^2 \hat{a}(t \mid \omega)\beta^\omega,$$

and integrating between 0 and t, we obtain:

$$\hat{s}(t \mid \omega) = s(0 \mid \omega) e^{-\beta^\omega \int_0^t \hat{k}(u)^2 \hat{a}(u|\omega) du}. \tag{20}$$

Moreover, integrating both sides of (18) between 0 and t yields:

$$\hat{a}(t \mid \omega) = a(0 \mid \omega) - \alpha^\omega(\hat{s}(t \mid \omega) - s(0 \mid \omega)) - \gamma^\omega \int_0^t \hat{a}(u \mid \omega) du.$$

Plugging (20) and the identity $a(0 \mid \omega) = \alpha^\omega(1 - s(0 \mid \omega))$ into the latter expression, we obtain:

$$\hat{a}(t \mid \omega) = \alpha^\omega - \alpha^\omega s(0 \mid \omega) e^{-\beta^\omega \int_0^t \hat{k}(u)^2 \hat{a}(u|\omega) du} - \gamma^\omega \int_0^t \hat{a}(u \mid \omega) du.$$

Finally, plugging (19) into (7), we obtain for all j:

$$\mu_j(t)(\omega) = \hat{\mu}(t)(\omega) = \mu^0(\omega) \frac{\alpha^\omega + (1-\alpha^\omega) e^{-\beta^\omega \int_0^t \hat{k}(u)^2 \hat{a}(u|\omega) du}}{\alpha^\omega + (1-\alpha^\omega) \mathrm{E}_{\mu^0}[e^{-\beta \int_0^t \hat{k}(u)^2 \hat{a}(u) du}]} \ \forall\ j,\ t.$$

Let us now prove that the strategy profile $\hat{\mathbf{k}}$ is an equilibrium. Suppose that the population plays

$$\hat{k}(t) = \min\left\{1, \frac{c_S}{\mathrm{E}_{\hat{\mu}(t)}\left[\hat{p}(t)\hat{a}(t)\beta v_I\right]}\right\} \ \forall\, t \leq T,$$

thus that the dynamics of epidemics and beliefs are given by (11),(12) and (13). Let us prove that no individual has a profitable deviation from $\hat{k}$. Fix some player j. Plugging $\bar{k}_A(t \mid \omega) = \hat{k}(t)$ into (10), we obtain that player j's best response to $\hat{\mathbf{k}}$ is:

$$k_j(t) \begin{cases} = 1 & \text{if } c_S > \hat{k}(t)\mathrm{E}_{\mu_j(t)}[p_j(t)\hat{a}(t)\beta v_I], \\ = 0 & \text{if } c_S < \hat{k}(t)\mathrm{E}_{\mu_j(t)}[p_j(t)\hat{a}(t)\beta v_I], \\ \in [0,1] & \text{if } c_S = \hat{k}(t)\mathrm{E}_{\mu_j(t)}[p_j(t)\hat{a}(t)\beta v_I]. \end{cases}$$

Suppose that j has a profitable deviation $\tilde{k}_j$ from $\hat{k}$ and let $\tau \geq 0$ be the first time player j deviates, i.e., $\tau := \min\{t : \tilde{k}_j(t) \neq \hat{k}(t)\}$.

• At time 0, player j has the same beliefs as the rest of the population, hence $\mathrm{E}_{\mu_j(0)}[p_j(0)\hat{a}(0)\beta v_I] = \mathrm{E}_{\hat{\mu}(0)}[\hat{p}(0)\hat{a}(0)\beta v_I]$.

- If $\hat{k}(0) = 1$, then by definition of $\hat{k}$, $c_S > \mathrm{E}_{\hat{\mu}(0)}\left[\hat{p}(0)\hat{a}(0)\beta v_I\right]$, thus player j's best response at time 0 is $k_j(0) = 1 = \hat{k}(0)$.

- If $\hat{k}(0) = \frac{c_S}{\mathrm{E}_{\hat{\mu}(0)}[\hat{p}(0)\hat{a}(0)\beta v_I]} < 1$, then any $k_j(0) \in [0,1]$ best responds to $\hat{k}(0)$.

This proves that player j has no profitable deviation from playing $\hat{k}(0)$ at time 0, thus $\tau > 0$.

• For all $t < \tau$, $\tilde{k}_j(t) = \hat{k}(t)$, thus player j holds the same beliefs as the population at time τ. It follows that $\mathrm{E}_{\mu_j(\tau)}[p_j(\tau)\hat{a}(\tau)\beta v_I] = \mathrm{E}_{\hat{\mu}(\tau)}[\hat{p}(\tau)\hat{a}(\tau)\beta v_I]$. We can thus apply the previous argument to prove that player j cannot improve her payoff by playing $\tilde{k}_j(\tau) \neq \hat{k}(\tau)$, which is a contradiction.

Finally, let us prove that $\hat{k}$ is the unique symmetric equilibrium. Suppose on the contrary that there exists another symmetric equilibrium strategy profile $\tilde{k} \neq \hat{k}$ and let $\tilde{p}$, $\tilde{\mu}$ and $\tilde{a}$ denote the associated equilibrium functions. According to the best response analysis, $\tilde{k}$ satisfies the following system at each time t:

$$\tilde{k}(t) \begin{cases} = 1 & \text{if } c_S > \mathrm{E}_{\tilde{\mu}(t)}[\tilde{p}(t)\tilde{a}(t)\beta v_I], \\ \in [0,1] & \text{if } c_S = \tilde{k}(t)\mathrm{E}_{\tilde{\mu}(t)}[\tilde{p}(t)\tilde{a}(t)\beta v_I]. \end{cases}$$

Let $\tau := \min\{t : \tilde{k}(t) \neq \hat{k}(t)\}$. At time 0, beliefs are the same in both equilibria, thus $\hat{k}(0) = \tilde{k}(0)$, which implies $\tau > 0$. Because $\tilde{k}(t) = \hat{k}(t)$ for all $t < \tau$, $\tilde{\mu}(t) = \hat{\mu}(t)$, $\tilde{p}(t) = \hat{p}(t)$, $\tilde{a}(t) = \hat{a}(t)$ for all $t < \tau$, hence, by continuity, $\mathrm{E}_{\tilde{\mu}(\tau)}[\tilde{p}(\tau)\tilde{a}(\tau)\beta v_I] = \mathrm{E}_{\hat{\mu}(\tau)}[\hat{p}(\tau)\hat{a}(\tau)\beta v_I]$. This entails $\tilde{k}(\tau) = \hat{k}(\tau)$, hence a contradiction.

B Proofs of Theorem 1 and Proposition 2

In what follows, we use the notation

$$\chi^\omega := \frac{c_S}{\beta^\omega v_I^\omega}.$$

B.1 Proof of Theorem 1

Fix some $\omega \in \Omega$ and suppose that $\mu^0(\omega) = 1$. Applying Proposition 1, in the unique symmetric equilibrium, the level of social activity at time t is

$$\hat{k}_\omega(t) = \min\left\{1, \Phi_\omega(t)\right\},$$

where

$$\Phi_\omega(t) := \frac{\chi^\omega}{\hat{p}(t \mid \omega)\hat{a}(t \mid \omega)}. \tag{21}$$

Moreover, equilibrium dynamics are

$$\dot{\hat{s}}(t \mid \omega) = -\hat{k}_\omega(t)^2 \hat{s}(t \mid \omega)\hat{a}(t \mid \omega)\beta^\omega \tag{22}$$

$$\dot{\hat{a}}(t \mid \omega) = \hat{a}(t \mid \omega)\left(\alpha^\omega \beta^\omega \hat{k}_\omega(t)^2 \hat{s}(t \mid \omega) - \gamma^\omega\right) \tag{23}$$

$$\dot{\hat{p}}(t \mid \omega) = -\hat{p}(t \mid \omega)(1 - \hat{p}(t \mid \omega))\hat{k}_\omega(t)^2 \hat{a}(t \mid \omega)\beta^\omega \tag{24}$$

Finally, recall that

$$\hat{s}(t \mid \omega) = \frac{\alpha^\omega}{1-\alpha^\omega}\frac{\hat{p}(t \mid \omega)}{1-\hat{p}(t \mid \omega)} s(0 \mid \omega) \tag{25}$$

We start by proving what follows.

Lemma 1. *If* (a) $\hat{k}_\omega(0) < 1$, (b) $s(0 \mid \omega) < \frac{\gamma^\omega}{\alpha^\omega \beta^\omega}\frac{1-\alpha^\omega}{\alpha^\omega}$ *and* (c) $\dot{\hat{a}}(0 \mid \omega) < 0$, *then* $\dot{\hat{a}}(t \mid \omega) < 0$ *for all* $t \leq T$.

Proof. It is immediate from (22) than $\hat{s}(t \mid \omega)$ strictly decreases since $\hat{s}(t \mid \omega)$ and $\hat{a}(t \mid \omega)$ are positive. Define

$$t_s := \inf\left\{t : \hat{s}(t \mid \omega) < \frac{\gamma^\omega}{\alpha^\omega \beta^\omega}\right\}.$$

As $\hat{k}(t) \leq 1$, $\dot{\hat{a}}(t \mid \omega) < 0$ for all $t > \tau_s$ by (23). The aim of the proof is to show that conditions (a), (b) and (c) together imply that $\dot{\hat{a}}(t \mid \omega) \leq 0$ for all $t \leq \tau_s$.

By Condition (a), $\hat{k}_\omega(0) < 1$, thus we can define

$$t_k := \inf\{t : \hat{k}(t) = 1\}.$$

- If $t_s \leq t_k$, then $\hat{k}(t) = \Phi_\omega(t) < 1$ for all $t < t_s$. Plugging (21) and (25) into (23) yields:

$$\dot{\hat{a}}(t \mid \omega) = \hat{a}(t \mid \omega)\left(\frac{(\alpha^\omega \chi^\omega)^2 \beta^\omega s(0 \mid \omega)}{1-\alpha^\omega}\frac{1}{\hat{p}(t \mid \omega)(1-\hat{p}(t \mid \omega))\hat{a}(t \mid \omega)^2} - \gamma^\omega\right).$$

  Differentiating this expression with respect to t, we obtain:

$$\ddot{\hat{a}}(t \mid \omega) = \frac{\dot{\hat{a}}(t \mid \omega)^2}{\hat{a}(t \mid \omega)} - \frac{(\alpha^\omega \chi^\omega)^2 \beta^\omega s(0 \mid \omega)}{1-\alpha^\omega}\frac{\dot{\hat{p}}(t \mid \omega)}{\hat{p}(t \mid \omega)^2(1-\hat{p}(t \mid \omega))^2 \hat{a}(t \mid \omega)}(1 - 2\hat{p}(t \mid \omega)).$$

  If $\dot{\hat{a}}(\bar{t} \mid \omega) = 0$ for some $\bar{t} < t_s$, then the latter expression evaluated in $\bar{t}$ is:

$$\ddot{\hat{a}}(\bar{t} \mid \omega) = -\frac{(\alpha^\omega \chi^\omega)^2 \beta^\omega s(0 \mid \omega)}{1-\alpha^\omega}\frac{\dot{\hat{p}}(\bar{t} \mid \omega)}{\hat{p}(\bar{t} \mid \omega)^2(1-\hat{p}(\bar{t} \mid \omega))^2 \hat{a}(\bar{t} \mid \omega)}(1 - 2\hat{p}(\bar{t} \mid \omega)).$$

However, by definition of t_s, $\hat{s}(\bar{t} \mid \omega) > \frac{\gamma^\omega}{\alpha^\omega \beta^\omega}$, thus, under condition (b),

$$\hat{s}(\bar{t} \mid \omega) > s(0 \mid \omega) \frac{\alpha^\omega}{1 - \alpha^\omega}. \tag{26}$$

Using (25), inequality (26) can be rewritten as:

$$\hat{p}(\bar{t} \mid \omega) > \frac{1}{2},$$

which implies that $\ddot{\hat{a}}(\bar{t} \mid \omega) < 0$. Moreover, $\dot{\hat{a}}(0 \mid \omega) < 0$ under condition (c). Therefore, $\dot{\hat{a}}(t \mid \omega) < 0$ for all $t < t_s$ if $t_s \leq t_k$.

- Let us now prove that $t_s \leq t_k$. Suppose on the contrary that $t_k < t_s$. Then, for all $t < t_k$, the former argument applies and proves that $\dot{\hat{a}}(t \mid \omega) < 0$ for all $t < t_k$, thus that $\hat{s}(t \mid \omega)\hat{k}_\omega(t)^2 < \gamma^\omega/(\alpha^\omega \beta^\omega)$ for all $t < t_k$. As $\hat{s}(t \mid \omega)$ is continuous in t,

$$\lim_{t \to t_k} \hat{s}(t \mid \omega)\hat{k}(t)^2 \leq \frac{\gamma^\omega}{\alpha^\omega \beta^\omega},$$

i.e., $\hat{s}(t_k \mid \omega) \leq \gamma^\omega/(\alpha^\omega \beta^\omega)$, which implies $t_s \geq t_k$, thus a contradiction.

□

We now turn to the proof of (i) and (ii). It is immediate to see that $\Phi_\omega(t)$ is increasing as $\hat{p}(t \mid \omega)$ and $\hat{a}(t \mid \omega)$ are decreasing by (24) and Lemma 1. Let us now prove that $\lim_{t \to +\infty} \Phi^\omega(t) = +\infty$. As $\dot{\hat{a}}(t \mid \omega) < 0$ for all t, $\hat{a}(u \mid \omega) \geq \hat{a}(t \mid \omega)$ for all $u \leq t$, thus $\int_0^t \hat{a}(u \mid \omega) du \geq t \times \hat{a}(t \mid \omega)$. Plugging this inequality into (23), we obtain that, for all t,

$$0 \leq \hat{a}(t \mid \omega) \leq \frac{\alpha^\omega}{1 + \gamma^\omega t}(1 - \hat{s}(t \mid \omega)), \tag{27}$$

which proves that $\lim_{t \to +\infty} \hat{a}(t \mid \omega) = 0$ since the right-hand side of (27) converges to 0. This and the fact that $\hat{p}(t \mid \omega)$ is bounded proves that $\lim_{t \to +\infty} \Phi^\omega(t) = +\infty$. Finally, by condition (a), it holds that $\Phi_\omega(0) < 1$. Therefore, there exists a unique $\hat{\tau}_\omega \in (0, +\infty)$ such that $\Phi_\omega(\hat{\tau}_\omega) = 1$, and $\hat{k}_\omega(t)$ is increasing for all $t \leq \hat{\tau}_\omega$. This proves (i).

By definition of $\hat{\tau}_\omega$, $\hat{\tau}_\omega > t_k$ and we have proven that $t_k \geq t_s$. This implies that $\hat{\tau}_\omega > t_s$, which further implies

$$\hat{s}(\hat{\tau}_\omega \mid \omega) < \frac{\gamma^\omega}{\alpha^\omega \beta^\omega}.$$

This proves (ii).

Finally, we express conditions (a), (b) and (c) as a unique condition pertaining on η^ω. Using the identity $a(0 \mid \omega) = \alpha^\omega(1 - s(0 \mid \omega))$, these conditions are rewritten as follows:

(a) $\eta^\omega > c_S$,

(b) $\eta^\omega > \left(\alpha^\omega - \frac{\gamma^\omega}{\alpha^\omega \beta^\omega}(1-\alpha^\omega)\right)(1-\alpha^\omega)\beta^\omega v_I^\omega := n_b$,

(c) $\eta^\omega > 2\alpha^\omega(1-\alpha^\omega)\beta^\omega v_I^\omega / \left(1 + \sqrt{1 + 4\frac{\alpha^\omega \gamma^\omega}{\beta^\omega}\left(\frac{1-\alpha^\omega}{\chi^\omega}\right)^2}\right) := n_c$.

The result of Theorem 1 is obtained by letting

$$\underline{n}(\omega) := \max\{c_S, n_b, n_c\}.$$

Lemma 2. *The value of $\underline{n}(\omega)$ is as follows.*

- *If* $\left(\alpha^\omega - \frac{\gamma^\omega}{\alpha^\omega \beta^\omega}(1-\alpha^\omega)\right) < 0$, *then*

$$\underline{n}(\omega) = \begin{cases} n_c & \text{if } \chi^\omega \leq (\alpha^\omega - \frac{\gamma^\omega}{\beta^\omega})(1-\alpha^\omega), \\ c_S & \text{if } \chi^\omega \geq (\alpha^\omega - \frac{\gamma^\omega}{\beta^\omega})(1-\alpha^\omega). \end{cases}$$

- *If* $\left(\alpha^\omega - \frac{\gamma^\omega}{\alpha^\omega \beta^\omega}(1-\alpha^\omega)\right) > 0$ *and* $\alpha^\omega > 1/2$, *then*

$$\underline{n}(\omega) = \begin{cases} n_b & \text{if } \chi^\omega \leq \left(\alpha^\omega - \frac{\gamma^\omega}{\alpha^\omega \beta^\omega}(1-\alpha^\omega)\right)(1-\alpha^\omega), \\ c_S & \text{if } \chi^\omega \geq \left(\alpha^\omega - \frac{\gamma^\omega}{\alpha^\omega \beta^\omega}(1-\alpha^\omega)\right)(1-\alpha^\omega). \end{cases}$$

- *If* $\left(\alpha^\omega - \frac{\gamma^\omega}{\alpha^\omega \beta^\omega}(1-\alpha^\omega)\right) > 0$ *and* $\alpha^\omega < 1/2$, *then*

$$\underline{n}(\omega) = \begin{cases} n_b & \text{if } \chi^\omega \leq \left(\alpha^\omega - \frac{\gamma^\omega}{\alpha^\omega \beta^\omega}(1-\alpha^\omega)\right)\sqrt{\alpha^\omega(1-\alpha^\omega)}, \\ n_c & \text{if } \chi^\omega \in [\left(\alpha^\omega - \frac{\gamma^\omega}{\alpha^\omega \beta^\omega}(1-\alpha^\omega)\right)\sqrt{\alpha^\omega(1-\alpha^\omega)}, (\alpha^\omega - \frac{\gamma^\omega}{\beta^\omega})(1-\alpha^\omega)], \\ c_S & \text{if } \chi^\omega \geq (\alpha^\omega - \frac{\gamma^\omega}{\beta^\omega})(1-\alpha^\omega). \end{cases}$$

Proof. Basic calculation. □

B.2 Proof of Proposition 2

The severity condition holds if the following constraints are satisfied:

1. $\eta^\omega \geq c_S$;

2. $\eta^\omega \geq n_b$;

3. $\eta^\omega \geq n_c$;

We first analyze the impact of the parameters on each constraint.

- $\eta^\omega \geq c_S \Leftrightarrow a(0 \mid \omega)(1-\alpha^\omega)\beta^\omega v_I^\omega \geq c_S$

  The left-hand term increases with $a(0 \mid \omega)$, β^ω and v_I^ω and decreases with α^ω. Therefore, an increase in $a(0 \mid \omega)$, β^ω or v_I^ω relaxes the constraint 1, while an increase in α^ω strengthens it. An increase in γ^ω has no impact on constraint 1.

- $\eta^\omega \geq n_b \Leftrightarrow a(0 \mid \omega) \geq \alpha^\omega - \frac{\gamma^\omega}{\alpha^\omega \beta^\omega}(1-\alpha^\omega)$.

  The left-hand term increases with $a(0 \mid \omega)$. The right-hand term increases with α^ω and β^ω, and decreases with γ^ω.

  Therefore, an increase in $a(0 \mid \omega)$ or γ^ω relaxes the constraint 2, while an increase in α^ω or β^ω strengthens it. An increase in v_I^ω has no impact on constraint 2.

- $\eta^\omega \geq n_c \Leftrightarrow a(0 \mid \omega) \geq \frac{2\alpha^\omega}{1+\sqrt{1+4\alpha^\omega\gamma^\omega((1-\alpha^\omega)v_I^\omega)^2\beta^\omega/c_S^2}}$.

  The left-hand term increases with $a(0 \mid \omega)$. The right-hand term decreases with β^ω, γ^ω and v_I^ω, and increases with α^ω.

  Therefore, an increase in $a(0 \mid \omega)$, β^ω, γ^ω or v_I^ω relaxes constraint 3, while an increase in α^ω strengthens it.

Summarizing, constraints 1, 2 and 3 are (weakly) relaxed by an increase in $a(0 \mid \omega)$, γ^ω, v_I^ω and $(1-\alpha^\omega)$ and a decrease in c_S. The impact of an increase in β^ω is ambiguous.

C Proof of Theorem 2

To lighten the notation, we write α^i, β^i, γ^i, and χ^i instead of α^{ω_i}, β^{ω_i}, γ^{ω_i}, and χ^{ω_i} for $i = L, H$.

The proof strategy is to compare the equilibrium dynamics in state ω_L with those of the non-behavioral SIR model in the same state, i.e., when all individuals choose $k(t) = 1$ for all t. It consists in three intermediary claims. We first prove that, in the non-behavioral SIR model, the infection rate increases up to some time $\tau_{\omega_L} > 0$ and decreases afterwards, provided there

is no herd immunity at time 0 (Claim 1). Next, we prove that if individuals stop self-isolating before τ_{ω_L} when they know that the state is ω_H, then, in equilibrium they also stop self-isolating before time τ_{ω_L} provided the prior probability of state ω_H is large enough (Claim 2). Then, we prove that if individuals stop self-isolating before τ_{ω_L} in equilibrium, then, in state ω_L, there is a non-empty interval of time during which the infection rate increases (Claim 3).

Claim 1. *Let $\check{s}(t \mid \omega)$ denote the proportion of the susceptible pool at time t and state ω in the non-behavioral SIR model. Let $\tau_{\omega_L} := \min\{t,\ \check{s}(t \mid \omega_L) \leq \frac{\gamma^L}{\alpha^L \beta^L}\}$. If $s(0 \mid \omega_L) > \frac{\gamma^L}{\alpha^L \beta^L}$, τ_{ω_L} is positive and defined by*

$$\tau_{\omega_L} = \int_{\frac{\gamma^L}{\alpha^L \beta^L}}^{s(0|\omega_L)} \frac{1}{u\left(\alpha^L \beta^L (1-u) + \gamma^L \ln\left(\frac{u}{s(0|\omega_L)}\right)\right)} du.$$

Proof. See section C.1. □

Claim 2. *If $0 < \hat{\tau}_H < \tau_{\omega_L}$, then there exist $\mu_1^* < 1$ and $\tau \in (0, \tau_{\omega_L})$ such that, for all $\mu^0(\omega_H) \geq \mu_1^*$, $\hat{k}(\tau) = 1$.*

Proof. See section C.2. □

Claim 3. *If $\hat{k}(\tau) = 1$ for some $\tau < \tau_{\omega_L}$, then there are $t_1 < t_2 \leq \tau$ such that $\dot{\hat{a}}(t \mid \omega_L) > 0$ for all $t \in (t_1, t_2]$.*

Proof. See section C.3. □

We can now turn to the proof of Theorem 2. Consider two epidemic states ω_H and ω_L such that $\eta^{\omega_H} > \underline{n}(\omega_H)$, $\eta^{\omega_L} > \underline{n}(\omega_L)$ and the following conditions holds:

$$\mathbf{A}: \quad a(0 \mid \omega_L) < \frac{\alpha^L(\alpha^L \beta^L - \gamma^L)}{\alpha^L \beta^L - \gamma^L + \gamma^L e^{\frac{\alpha^L \beta^L}{\gamma^H}\left(\frac{\alpha^H(1-\alpha^H)}{\chi^H} - 1\right)}}$$

$$\mathbf{B}: \quad \eta^H > \eta^L$$

By definition, $\underline{n}(\omega_H) \geq c_S$. Condition $\eta^{\omega_H} > \underline{n}(\omega_H)$ thus implies $a(0 \mid \omega_H)\beta^H v_I^H(1-\alpha^H) > c_S$, which can be rewritten as $a(0 \mid \omega_H)(1-\alpha^H) > \chi^H$. As $a(0 \mid \omega_H) \leq \alpha^H$ by assumption, it follows that

$$(i)\ \chi^H < \alpha^H(1-\alpha^H).$$

Moreover, applying Theorem 1, we know that there exists a positive, finite cutoff $\hat{\tau}_H$ such that $k_H(\hat{\tau}_H) = 1$. Let us bound the cutoff $\hat{\tau}_H$. As inequality (27) holds for all t and $\hat{s}(\hat{\tau}_H \mid \omega_H) \geq 0$, we can write:

$$\hat{\tau}_H \leq \frac{1}{\gamma^H}\left(\frac{\alpha^H}{\hat{a}(\hat{\tau}_H \mid \omega_H)} - 1\right).$$

Yet, by definition of $\hat{\tau}_H$, it holds that $\Phi_H(\hat{\tau}_H) = 1$, which implies

$$\hat{a}(\tau_H \mid \omega_H) = \frac{\chi^H}{\hat{p}(\hat{\tau}_H \mid \omega_H)}.$$

Because $\hat{p}(\hat{\tau}_H \mid \omega_H) \leq \hat{p}(0 \mid \omega_H) = 1 - \alpha^H$, it follows from the latter expression that $\hat{a}(\hat{\tau}_H \mid \omega_H) > \chi^H/(1-\alpha^H)$, which further yields:

$$\hat{\tau}_H < \frac{1}{\gamma^H}\left(\frac{\alpha^H(1-\alpha^H)}{\chi^H} - 1\right).$$

The inequality (i) $\chi^H < \alpha^H(1-\alpha^H)$ has two direct consequences. First, the upper bound on $\hat{\tau}_H$ is strictly positive. Second, the following inequality holds:

$$\frac{\alpha^L(\alpha^L\beta^L - \gamma^L)}{\alpha^L\beta^L - \gamma^L + \gamma^L e^{\frac{\alpha^L\beta^L}{\gamma^H}\left(\frac{\alpha^H(1-\alpha^H)}{\chi^H}-1\right)}} < \alpha^L - \frac{\gamma^L}{\beta^L}.$$

Therefore, under condition **A** it holds that $a(0 \mid \omega_L) < \alpha^L - \frac{\gamma^L}{\beta^L}$, which is equivalent to

$$s(0 \mid \omega_L) > \frac{\gamma^L}{\alpha^L\beta^L}.$$

We can therefore apply Claim 1 and state that there exists $\tau_{\omega_L} > 0$ such that $\dot{\hat{a}}(t \mid \omega_L) > 0$ if and only if $t < \tau_{\omega_L}$. Basic integration techniques allow to prove that the cutoff τ_{ω_L} satisfies :

$$\tau_{\omega_L} > \frac{1}{\alpha^L\beta^L}(\ln\left(\frac{\alpha^L\beta^L - \gamma^L}{\gamma^L}\right) - \ln\left(\frac{1 - s(0 \mid \omega_L)}{s(0 \mid \omega_L)}\right)).$$

Moreover, condition **A** is equivalent to the right hand term in the above expression being larger than $\frac{1}{\gamma^H}\left(\frac{\alpha^H(1-\alpha^H)}{\chi^H} - 1\right)$. Therefore, condition **A** implies:

$$\hat{\tau}_H < \tau_{\omega_L}.$$

We can thus apply Claim 2 and Claim 3, and state that there exists $\mu_1^* < 1$ and $t_1 < t_2 < \hat{\tau}_H$ such that $\dot{\hat{a}}(t \mid \omega_L) > 0$ on (t_1, t_2) provided $\mu^0(\omega_H) > \mu_1^*$.

Let us now prove that $\dot{\hat{a}}(0 \mid \omega_L) < 0$. Since $\eta^{\omega_L} > \underline{n}(\omega_L)$, we know by Theorem 1 that prevalence decreases in state ω_L under certainty. In particular, it decreases in time 0, which implies:

$$s(0 \mid \omega_L) < \frac{\gamma^L}{\alpha^L\beta^L}\frac{1}{\hat{k}_L(0)^2}.$$

Condition **B** is equivalent to $\hat{k}_H(0) < \hat{k}_L(0)$. Therefore, under **B** it holds that

$$s(0 \mid \omega_L) < \frac{\gamma^L}{\alpha^L \beta^L} \frac{1}{\hat{k}_H(0)^2},$$

which is equivalent to

$$\dot{\hat{a}}_H(0 \mid \omega_L) < 0.$$

It is straightforward to prove that $\lim_{\mu^0(\omega_H)\to 1} \dot{\hat{a}}(0 \mid \omega_L) = \dot{\hat{a}}_H(0 \mid \omega_L)$. Therefore, there exists $\mu_2^* < 1$ such that, if $\mu^0(\omega_H) > \mu_2^*$, equilibrium prevalence decrease at time 0 in state ω_L, i.e.,

$$\forall\ \mu^0(\omega_H) > \mu_2^*,\ \dot{\hat{a}}(0 \mid \omega_L) < 0.$$

Taking $\mu^* = \max\{\mu_1^*, \mu_2^*\}$, we can state that for all $\mu^0(\omega_H) > \mu^*$, $\hat{a}(t \mid \omega_L)$ decreases in time 0 and increases on (t_1, t_2) with $0 < t_1 < t_2 < \hat{\tau}_H$.

We conclude by proving that one can find ω_L and ω_H that satisfy the conditions of Theorem 2.

Lemma 3. *There exist* $\omega_L : (a(0 \mid \omega_L), \alpha^L, \beta^L, \gamma^L, v_I^L)$ *and* $\omega_H : (a(0 \mid \omega_H), \alpha^H, \beta^H, \gamma^H, v_I^H)$ *such that* $a(0 \mid \omega) < \alpha^\omega$ *and*

$$\begin{aligned}
&(a) \quad \eta^H > \underline{n}(\omega_H) \\
&(b) \quad \eta^L > \underline{n}(\omega_L) \\
&(c) \quad \eta^H > \eta^L \\
&(d) \quad a(0 \mid \omega_L) < \frac{\alpha^L(\alpha^L\beta^L - \gamma^L)}{\alpha^L\beta^L - \gamma^L + \gamma^L e^{\frac{\alpha^L\beta^L}{\gamma^H}(\frac{\alpha^H(1-\alpha^H)}{\chi^H} - 1)}}
\end{aligned}$$

Proof.
Let us first prove that, if $\alpha^H < 1/2$, $\chi^H \in \left((\alpha^H - \frac{\gamma^H}{\beta^H})(1-\alpha^H), a(0 \mid \omega_H)(1-\alpha^H)\sqrt{\frac{\alpha^L}{1-\alpha^L}}\right)$ and $\alpha^L < 1/2$, then condition (a) is satisfied. By Lemma 2, when $\alpha^H < 1/2$ and $\chi^H > (\alpha^H - \frac{\gamma^H}{\beta^H})(1-\alpha^H)$, condition (a) is rewritten

$$(a) \Leftrightarrow a(0 \mid \omega_H) > \frac{\chi^H}{1-\alpha^H}.$$

Yet, as $\alpha^L < 1/2$, $\sqrt{\frac{\alpha^L}{1-\alpha^L}} < 1$, hence inequality $\chi^H < a(0 \mid \omega_H)(1-\alpha^H)\sqrt{\frac{\alpha^L}{1-\alpha^L}}$ implies $\chi^H < a(0 \mid \omega_H)(1-\alpha^H)$, which proves (a) is satisfied.

Next, let us prove that if $\chi^H \in \left((\alpha^H - \frac{\gamma^H}{\beta^H})(1-\alpha^H), a(0 \mid \omega_H)(1-\alpha^H)\sqrt{\frac{\alpha^L}{1-\alpha^L}}\right)$, $\alpha^L < 1/2$, $\alpha^L - \frac{\gamma^L}{\alpha^L\beta^L}(1-\alpha^L) > 0$ and

$$(*)\ \alpha^H\beta^H - \gamma^H > \frac{\alpha^L\beta^L}{\ln((1-\alpha^L)(\alpha^L\beta^L - \gamma^L)) - \ln((\alpha^L)^2\beta^L - \gamma^L(1-\alpha^L))} := A(\alpha^L, \beta^L, \gamma^L),$$

then the set

$$\mathcal{C} := \{(a(0 \mid \omega_L), v_I^L) \text{ such that } (b),\ (c), \text{ and } (d) \text{ hold}\}$$

is not empty.

As $\alpha^L < 1/2$ and $\alpha^L - \frac{\gamma^L}{\alpha^L\beta^L}(1-\alpha^L) > 0$, if $\chi^L \leq \left(\alpha^L - \frac{\gamma^L}{\alpha^L\beta^L}(1-\alpha^L)\right)\sqrt{\alpha^L(1-\alpha^L)}$, then by Lemma 2, condition (b) is rewritten

$$(b) \Leftrightarrow a(0 \mid \omega_L) > \alpha^L - \frac{\gamma^L}{\alpha^L\beta^L}(1-\alpha^L).$$

Moreover, condition (c) is rewritten

$$(c) \Leftrightarrow a(0 \mid \omega_L) < \frac{\chi^L}{1-\alpha^L}\frac{a(0 \mid \omega_H)(1-\alpha^H)}{\chi^H}.$$

Observe that the right-hand term in the above inequality is larger than $\alpha^L - \frac{\gamma^L}{\alpha^L\beta^L}(1-\alpha^L)$ if and only if $\chi^L > \left(\alpha^L - \frac{\gamma^L}{\alpha^L\beta^L}(1-\alpha^L)\right)\frac{(1-\alpha^L)\chi^H}{a(0|\omega^H)(1-\alpha^H)}$. As $\chi^H < a(0 \mid \omega_H)(1-\alpha^H)\sqrt{\frac{\alpha^L}{1-\alpha^L}}$, the interval

$$\left(\left(\alpha^L - \frac{\gamma^L}{\alpha^L\beta^L}(1-\alpha^L)\right)\frac{(1-\alpha^L)\chi^H}{a(0 \mid \omega_H)(1-\alpha^H)}, \left(\alpha^L - \frac{\gamma^L}{\alpha^L\beta^L}(1-\alpha^L)\right)\sqrt{\alpha^L(1-\alpha^L)}\right)$$

is not empty, thus for all χ^L in this interval, conditions (b) and (c) are satisfied for all values

$$a(0 \mid \omega^L) \in \left(\alpha^L - \frac{\gamma^L}{\alpha^L\beta^L}(1-\alpha^L), \frac{\chi^L}{1-\alpha^L}\frac{a(0 \mid \omega_H)(1-\alpha^H)}{\chi^H}\right).$$

The set

$$\mathcal{C}' := \{(a(0 \mid \omega_L), v_I^L) \mid (b) \text{ and } (c) \text{ hold}\}$$

is thus non empty. Moreover, as $\chi^H > (\alpha^H - \frac{\gamma^H}{\beta^H})(1-\alpha^H)$, it holds that

$$\frac{\alpha^L(\alpha^L\beta^L - \gamma^L)}{\alpha^L\beta^L - \gamma^L + \gamma^L e^{\frac{\alpha^L\beta^L}{\gamma^H}(\frac{\alpha^H(1-\alpha^H)}{\chi^H} - 1)}} > \frac{\alpha^L(\alpha^L\beta^L - \gamma^L)}{\alpha^L\beta^L - \gamma^L + \gamma^L e^{\frac{\alpha^L\beta^L}{\alpha^H\beta^H - \gamma^H}}} := z(\alpha^H\beta^H - \gamma^H).$$

As a consequence, $a(0 \mid \omega_L) < z(\alpha^H\beta^H - \gamma^H) \Rightarrow (d)$. Finally, it is straightforward to see that the set

$$\mathcal{C}'' := \mathcal{C}' \cap \left\{(a(0 \mid \omega_L), v_I^L),\ a(0 \mid \omega_L) < z(\alpha^H\beta^H - \gamma^H)\right\}$$

is not empty if and only if $z(\alpha^H\beta^H - \gamma^H) > \alpha^L - \frac{\gamma^L}{\alpha^L\beta^L}(1-\alpha^L)$, which is equivalent to condition $(*)$. As we have proven that $\mathcal{C}'' \subseteq \mathcal{C}$, the result follows.

Let us now conclude the proof. It is possible to choose $\alpha^L, \beta^L, \gamma^L, \alpha^H, \beta^H, \gamma^H$ and v_I^H as follows:

1. $\alpha^L < 1/2$ and β^L, γ^L such that $\alpha^L - \frac{\gamma^L}{\alpha^L\beta^L}(1-\alpha^L) > 0$ (which implies $\alpha^L\beta^L - \gamma^L > 0$).
2. $\alpha^H < 1/2$ and β^H such that $\alpha^H\beta^H > A(\alpha^L, \beta^L, \gamma^L)\sqrt{\frac{1-\alpha^L}{\alpha^L}}$.

As $\alpha^L < 1/2$, $A(\alpha^L, \beta^L, \gamma^L) > 0$. Moreover, as $\alpha^H\beta^H > A(\alpha^L, \beta^L, \gamma^L)\sqrt{\frac{1-\alpha^L}{\alpha^L}}$, the interval $(\alpha^H\beta^H(1-\sqrt{\frac{\alpha^L}{1-\alpha^L}}), \alpha^H\beta^H - A(\alpha^L, \beta^L, \gamma^L))$ is not empty. This allows to chose

3. $\gamma^H \in (\alpha^H\beta^H(1-\sqrt{\frac{\alpha^L}{1-\alpha^L}}), \alpha^H\beta^H - A(\alpha^L, \beta^L, \gamma^L))$.

As $\gamma^H < \alpha^H\beta^H - A(\alpha^L, \beta^L, \gamma^L)$, condition (*) is satisfied and $\alpha^H\beta^H - \gamma^H > 0$. As $\gamma^H > \alpha^H\beta^H(1-\sqrt{\frac{\alpha^L}{1-\alpha^L}})$, the interval $\left(\frac{c_S\sqrt{1-\alpha^L}}{\alpha^H\beta^H(1-\alpha^H)\sqrt{\alpha^L}}, \frac{c_S}{(\alpha^H\beta^H-\gamma^H)(1-\alpha^H)}\right)$ is not empty, and we can chose

4. $v_I^H \in \left(\frac{c_S\sqrt{1-\alpha^L}}{\alpha^H\beta^H(1-\alpha^H)\sqrt{\alpha^L}}, \frac{c_S}{(\alpha^H\beta^H-\gamma^H)(1-\alpha^H)}\right)$.

The latter point implies $\frac{c_S\sqrt{1-\alpha^L}}{\beta^H v_I^H(1-\alpha^H)\sqrt{\alpha^L}} < \alpha^H$, hence we can chose $a(0 \mid \omega_H)$ as follows:

5. $a(0 \mid \omega_H) \in \left(\frac{c_S\sqrt{1-\alpha^L}}{\beta^H v_I^H(1-\alpha^H)\sqrt{\alpha^L}}, \alpha^H\right)$.

Point 5. directly implies that $a(0 \mid \omega_H) < \alpha^H$.
Points 4. and 5. imply $(\alpha^H - \frac{\gamma^H}{\beta^H})(1-\alpha^H) < \chi^H < a(0 \mid \omega_H)(1-\alpha_H)\sqrt{\frac{\alpha^L}{1-\alpha^L}}$, which, together with inequalities $\alpha^L < 1/2$ (point 1.) and $\alpha^H < 1/2$ (point 2.), imply that (a) is satisfied. Moreover, points 1., 2., 3., 4. and 5. guarantee that the set $\mathcal{C}$ is not empty. Therefore, we can choose $a(0 \mid \omega_L)$ and v_I^L such that

5. $(a(0 \mid \omega_L), v_I^L) \in \mathcal{C}$.

As $(a(0 \mid \omega_L), v_I^L) \in \mathcal{C}$, conditions (b), (c) and (d) are satisfied. Finally, as the right-hand side in inequality (d) is smaller than α^L, it holds that $a(0 \mid \omega_L) < \alpha^L$.

□

C.1 Proof of Claim 1

Let us denote by $\check{s}(t \mid \omega_L)$ and $\check{a}(t \mid \omega_L)$ the fraction of the population susceptible and infected without symptoms at time t when the epidemic is ω_L in the non-behavioral SIR model. Plugging $k(t) = 1$ into (1) and (2), we obtain the dynamics:

$$\begin{cases} \dot{\check{s}}(t \mid \omega_L) = -\beta^L \check{s}(t \mid \omega_L)\check{a}(t \mid \omega_L), \\ \dot{\check{a}}(t \mid \omega_L) = \check{a}(t \mid \omega_L)(\alpha^L \beta^L \check{s}(t \mid \omega_L) - \gamma^L), \\ \text{with } \check{s}(0 \mid \omega_L) = s(0 \mid \omega_L) \text{ and } \check{a}(0 \mid \omega_L) = a(0 \mid \omega_L). \end{cases} \tag{28}$$

We work towards a contradiction. Suppose in contrast that $\check{s}(t \mid \omega_L) > \gamma^L/(\alpha^L\beta^L)$ for all t. This implies $\dot{\check{a}}(t \mid \omega_L) > 0$ for all t, thus $\check{a}(t \mid \omega_L) > a(0 \mid \omega_L)$ for all t. As $\check{s}(t \mid \omega_L) = s(0 \mid \omega_L)e^{-\beta^L \int_0^t \check{a}(u|\omega_L)du}$ by (28), this further implies that $\check{s}(t \mid \omega_L) < s(0 \mid \omega_L)e^{-\beta^L a(0|\omega_L)\times t}$ for all t, hence that $\lim_{t\to+\infty} \check{s}(t \mid \omega_L) = 0$, which contradicts $\check{s}(t \mid \omega_L) > \gamma^L/(\alpha^L\beta^L)$ for all t. This, together with $s(0 \mid \omega_L) > \gamma^L/(\alpha^L\beta^L)$ and the fact that $\dot{\check{s}}(t \mid \omega_L) \leq 0$ for all t, proves that there is a unique τ_{ω_L} such that $\check{s}(\tau_{\omega_L} \mid \omega_L) = \gamma^L/(\alpha^L\beta^L)$. Let us now give the analytical expression of τ_{ω_L}. Combining the first two equations of (28), we obtain that, for all t,

$$\dot{\check{a}}(t \mid \omega_L) = -\alpha^L \dot{\check{s}}(t \mid \omega_L) + \frac{\gamma^L}{\beta^L}\frac{\dot{\check{s}}(t \mid \omega_L)}{\check{s}(t \mid \omega_L)}.$$

Integrating the latter expression between 0 and t, we obtain:

$$\check{a}(t \mid \omega_L) = a(0 \mid \omega_L) - \alpha^L(\check{s}(t \mid \omega_L) - s(0 \mid \omega_L)) + \frac{\gamma^L}{\beta^L}\ln\left(\frac{\check{s}(t \mid \omega_L)}{s(0 \mid \omega_L)}\right).$$

Plugging that into the dynamics of $\check{s}$ yields

$$\dot{\check{s}}(t \mid \omega_L) = -\beta^L \check{s}(t \mid \omega_L)\left(a(0 \mid \omega_L) - \alpha^L(\check{s}(t \mid \omega_L) - s(0 \mid \omega_L)) + \frac{\gamma^L}{\beta^L}\ln\left(\frac{\check{s}(t \mid \omega_L)}{s(0 \mid \omega_L)}\right)\right),$$

thus

$$-\frac{\dot{\check{s}}(t \mid \omega_L)}{\beta^L \check{s}(t \mid \omega_L)\left(a(0 \mid \omega_L) - \alpha^L(\check{s}(t \mid \omega_L) - s(0 \mid \omega_L)) + \frac{\gamma^L}{\beta^L}\ln\left(\frac{\check{s}(t|\omega_L)}{s(0|\omega_L)}\right)\right)} = 1$$

Integrating both hands between 0 and τ_{ω_L}, we obtain:

$$\tau_{\omega_L} = -\int_0^{\tau_{\omega_L}} \frac{\dot{\check{s}}(t \mid \omega_L)}{\beta^L \check{s}(t \mid \omega_L)\left(a(0 \mid \omega_L) - \alpha^L(\check{s}(t \mid \omega_L) - s(0 \mid \omega_L)) + \frac{\gamma^L}{\beta^L}\ln\left(\frac{\check{s}(t|\omega_L)}{s(0|\omega_L)}\right)\right)} dt$$

Proceeding to the variable change $u := \check{s}(t \mid \omega_L)$ and $du = \dot{\check{s}}(t \mid \omega_L)dt$, we further obtain:

$$\tau_{\omega_L} = -\int_{s(0|\omega_L)}^{\gamma^L/(\alpha^L\beta^L)} \frac{1}{\beta^L u\left(a(0 \mid \omega_L) - \alpha^L(u - s(0 \mid \omega_L)) + \frac{\gamma^L}{\beta^L}\ln\left(\frac{u}{s(0|\omega_L)}\right)\right)} du$$

$$= \int_{\gamma^L/(\alpha^L\beta^L)}^{s(0|\omega_L)} \frac{1}{\beta^L u\left(a(0 \mid \omega_L) - \alpha^L(u - s(0 \mid \omega_L)) + \frac{\gamma^L}{\beta^L}\ln\left(\frac{u}{s(0|\omega_L)}\right)\right)} du$$

Using the identity $a(0 \mid \omega_L) = \alpha^L(1 - s(0 \mid \omega_L))$, the above expression is rewritten

$$\tau_{\omega_L} = \int_{\frac{\gamma^L}{\alpha^L\beta^L}}^{s(0|\omega_L)} \frac{1}{u\left(\alpha^L(1-u) + \gamma^L \ln\left(\frac{u}{s(0|\omega_L)}\right)\right)} du.$$

C.2 Proof of Claim 2

Let $\hat{\Phi}(t)$ be defined by

$$\hat{\Phi}(t) = \frac{c_S}{E_{\hat{\mu}(t)}\left[\hat{p}(t)\hat{a}(t)\beta v_I\right]},$$

so that $\hat{k}(t) = \min\{1, \hat{\Phi}(t)\}$. Also, let us recall the notation $\hat{k}_H(t)$, $\hat{a}_H(t \mid \omega)$ and $\hat{p}_H(t \mid \omega)$ for equilibrium values when $\mu^0(\omega_H) = 1$, and

$$\Phi_H(t) = \frac{c_S}{\hat{p}_H(t \mid \omega_H)\hat{a}_H(t \mid \omega_H)\beta^H v_I^H},$$

so that $\hat{k}_H(t) = \min\{1, \Phi_H(t)\}$. By definition, $\Phi_H(\hat{\tau}_H) = 1$.

It is straightforward to show that

$$\lim_{\mu^0(\omega_H)\to 1} \hat{\Phi}(t) = \Phi_H(t) \;\forall\; t. \tag{29}$$

Taking $t = \tau_{\omega_L}$, (29) implies that, for all $\varepsilon > 0$, there is μ_ε such that for all $\mu^0(\omega_H) \geq \mu_\varepsilon$, $\hat{\Phi}(\tau_{\omega_L}) > \Phi_H(\tau_{\omega_L}) - \varepsilon$. As $\tau_H < \tau_{\omega_L}$ and $\Phi_H(t)$ is strictly increasing in t, $\Phi_H(\tau_{\omega_L}) > \Phi_H(\tau_H) = 1$, thus we can take $\varepsilon = \Phi_H(\tau_{\omega_L}) - 1 > 0$ and state that there is $\mu_\varepsilon < 1$ such that, for all $\mu^0(\omega_H) \geq \mu_\varepsilon$, $\hat{\Phi}(\tau_{\omega_L}) > 1$.

Taking $t = 0$, (29) implies that, for all $\varepsilon' > 0$ there is $\mu_{\varepsilon'} < 1$ such that for all $\mu^0(\omega_H) \geq \mu_{\varepsilon'}$, $\hat{\Phi}(0) < \Phi_H(0) + \varepsilon'$. As $\tau^H > 0$, $\Phi_H(0) < 1$. Thus we can take $\varepsilon' := 1 - \Phi_H(0)$, and state that there is $\mu_{\varepsilon'} < 1$ such that $\forall\, \mu^0(\omega_H) \geq \mu_{\varepsilon'}$, $\hat{\Phi}(0) < 1$.

Taking $\mu_1^* = \max\{\mu_\varepsilon, \mu_{\varepsilon'}\}$, we can state that for all $\mu^0(\omega_H) > \mu_1^*$, $\hat{\Phi}(0) < 1 < \hat{\Phi}(\tau_{\omega_L})$. As $\hat{\Phi}$ is continuous in t, if $\mu^0(\omega_H) > \mu_1^*$ there exists $\tau \in (0, \tau_{\omega_L})$ such that $\hat{\Phi}(\tau) = 1$, i.e., $\hat{k}(\tau) = 1$.

C.3 Proof of Claim 3

The aim of this claim is to prove that if individuals stop social distancing in epidemic ω_L before the cutoff time τ_{ω_L}, then there must be a time interval on which the infection rate

increases. We work towards a contradiction. As 1) $\dot{\check{a}}(t \mid \omega_L) > 0$ on $[0, \tau_{\omega_L})$, 2) $\tau < \tau_{\omega_L}$ and 3) $\hat{a}(0 \mid \omega_L) = \check{a}(0 \mid \omega_L)$, it is clear that if $\dot{\hat{a}}(t \mid \omega_L) \leq 0$ for all $t \leq \tau$, then $\hat{a}(t \mid \omega_L) < \check{a}(t \mid \omega_L)$ for all $0 < t \leq \tau$ As, in addition, $\hat{k}(t) \leq 1$ for all t, it follows that,

$$\hat{a}(t \mid \omega_L)\hat{k}(t)^2 < \check{a}(t \mid \omega_L) \; \forall \; 0 < t \leq \tau,$$

which implies

$$\hat{s}(t \mid \omega_L) > \check{s}(t \mid \omega_L) \; \forall \; t \leq \tau.$$

By definition of τ_{ω_L}, $\check{s}(t \mid \omega_L) > \gamma^L/(\alpha^L\beta^L)$ for all $t \leq \tau_{\omega_L}$. This implies in particular:

$$\hat{s}(\tau \mid \omega_L) > \gamma^L/(\alpha^L\beta^L). \tag{30}$$

Because $\hat{k}(\tau) = 1$,

$$\dot{\hat{a}}(\tau \mid \omega_L) = \hat{a}(\tau \mid \omega_L)(\alpha^L\beta^L\hat{s}(\tau \mid \omega_L) - \gamma^L).$$

By (30), $\dot{\hat{a}}(\tau \mid \omega_L) > 0$, which contradicts the assumption $\dot{\hat{a}}(t \mid \omega_L) \leq 0$ for all $t \leq \tau$. Therefore, there is $0 \leq t_1 < t_2 \leq \tau$ such that $\dot{\hat{a}}(t \mid \omega_L) > 0$ on (t_1, t_2).

D Calibration

To illustrate our findings, we simulate the equilibrium dynamics in the case where there are two possible epidemic states ω_L and ω_H, which differ only by the initial penetration value $a(0 \mid \omega)$. We set $\boldsymbol{a(0 \mid \omega_L) = 0.01}$ and $\boldsymbol{a(0 \mid \omega_H) = 0.05}$. Otherwise, the medical parameters α^ω, β^ω, γ^ω and ν^ω are the same in both states and are calibrated to the COVID-19 pandemic, thus ω_H is more severe than ω_L.

- Pollan et al. (2020) find that the proportion of asymptomatic individuals in the Spanish population who developed antibodies to the SARS CoV-2 ranges from 21.9% to 35.8%, thus we arbitrary set $\boldsymbol{\alpha^\omega = 0.3}$.
- Recovery from SARSCOV2 was usually estimated to take two weeks (see, e.g., Remuzzi and Remuzzi (2020)), which implies $\boldsymbol{\gamma^\omega = 1/15}$.
- To calibrate β^ω, we use the value of the *basic reproduction number* $\mathcal{R}_0$, i.e., the average number of secondary infections produced by a typical infected individual in a population where everyone is susceptible, whose estimation ranges between 2.5 and 3.5. In our

model, $\mathcal{R}_0 = \beta^\omega \alpha^\omega / \gamma^\omega$. We set $\mathcal{R}_0 = 3.2$, and therefore set $\beta^\omega = 3.2\gamma^\omega/\alpha^\omega$, which yields $\boldsymbol{\beta^\omega \simeq 71\%}$.

- Finally, the Fatality-Infected ratio has been estimated between 0.3% (Gudbjartsson et al. (2020)) and 0.7% (Verity et al. (2020)). We arbitrarily set a fatality rate $\nu^\omega/(\nu^\omega + \gamma_s^\omega) = 0.5\%$, which yields $\boldsymbol{\nu^\omega \simeq 0.033\%}$.

We adapt the calibration of Fenichel et al. (2011), who study a discrete-time model in which they set the discount rate to $\delta = 0.99986$, which corresponds to a 5% annual discount rate. The analog of δ in a continuous-time model is $r = -\ln(\delta)$, thus we set the discount rate to $\boldsymbol{r = -\ln(0.99986) \simeq 0.014\%}$.

Finally, we arbitrarily set the cost parameters to $\boldsymbol{c_S = 1}$ and $\boldsymbol{v_I^\omega = -37}$.