

BENNETT INSTITUTE WORKING PAPER

On the Management of Population Immunity¹

AUTHORS

Flavio Toxvaerd²
Robert Rowthorn³

ABSTRACT

This paper considers a susceptible-infected-recovered type model of infectious diseases, such as COVID-19 or swine flu, in which costly treatment or vaccination confers immunity on recovered individuals. Once immune, individuals indirectly protect the remaining susceptibles, who benefit from a measure of herd immunity. Treatment and vaccination directly induce such herd immunity, which builds up over time. Optimal treatment is shown to involve intervention at early stages of the epidemic, while optimal vaccination defers intervention to later stages. Thus, while treatment and vaccination have superficial similarities, their effects and desirability at different stages of the epidemic are radically different. Equilibrium vaccination is qualitatively similar to socially optimal vaccination, while equilibrium treatment differs in nature from socially optimal treatment. The optimal policies are compared to traditional non-economic public health interventions which rely on herd immunity thresholds.

Keywords: Economic epidemiology, herd immunity, treatment, vaccination, externalities

JEL Classification: C73, I18

Published August 2020

Publication from the Bennett Institute for Public Policy, Cambridge University
bennettinstitute.cam.ac.uk

¹We gratefully acknowledge helpful feedback from Joshua Ross, Chryssi Giannitsarou and Stephen Kissler. We also benefited from feedback at talks at the University of Cambridge (at the Faculty of Economics, the Department of Plant Sciences and the Bennett Institute for Public Policy, respectively), at the IJCF Lisbon Meetings in Game Theory and Applications (2015), at the FFA-FSFM Meetings in Manchester (2019) and at the Department of Physics at the Universidad de Buenos Aires. Some of the work for this paper was completed while Flavio Toxvaerd was a Visiting Fellow at the Bennett Institute for Public Policy (2018-19). Their kind support is gratefully acknowledged.

²Faculty of Economics, University of Cambridge. Address for correspondence: Faculty of Economics, University of Cambridge, Austin Robinson Building, Sidgwick Avenue, Cambridge CB3 9DD, United Kingdom. Email: fmot2@cam.ac.uk

³Faculty of Economics, University of Cambridge.

1. INTRODUCTION

Across the world, policy makers, businesses, civil society and individual citizens have grappled with the very serious health and economic consequences of the ongoing COVID-19 pandemic. A highly contagious disease, COVID-19 is still not vaccine preventable and without exception, policy makers have instead relied on so-called non-pharmaceutical interventions (NPIs) such as lockdowns, social distancing measures, workplace rotation schemes and so forth (see e.g. Ely et al., 2020). Common to all such measures is that they work by creating physical distance between people, thereby avoiding the spread of the disease from infected to susceptible individuals.

It is no exaggeration to say that these measures have come at an enormous economic and social cost. During the deepest recession since the Great Depression, businesses are struggling to keep afloat, workers are cut off from their livelihoods, the elderly and ill have had access to health care curtailed and many children have had their schooling and education put on hold or greatly diminished in quality. From any number of perspectives, the current pandemic has been a calamity unlike any in recent memory.

The central reason for this fallout from the epidemic is the exclusive reliance on non-pharmaceutical interventions, which by their very nature cut deep into many of the activities that create welfare and wellbeing in normal times. For this same reason, there is a general consensus that there are only three possible end games to the current epidemic. The worst possible outcome is that the population does not build up any natural immunity and that the infection becomes endemic (see Giannitsarou et al., 2020). The second is that some immunity builds up in the population (including cross-immunity from other infectious diseases), thereby reducing the potency of the epidemic to the point that it can be effectively managed with a combination of testing, track and trace procedures and NPIs such as isolation of infected cases. The last, and by far the best outcome, would be a scenario in which safe and effective pharmaceutical interventions such as vaccines and antivirals become widely and commercially available.

The focus of the present paper is this third scenario. We model the effects of mass vaccination and treatment in an economic-epidemiological framework based on the SIR compartmental model. Such pharmaceutical interventions directly influence the rates of transition between health states of individuals and therefore also have important aggregate effects. In many recent studies, the advent of pharmaceutical interventions has been treated as the end of active disease control, assuming that the post-vaccine world is a return to normality. But vaccines and antivirals are themselves imperfect health policy tools that can and should be wielded judiciously. In this paper, we explicitly consider the socially optimal way to do so. We make two main contributions: First, we analyze vaccination and treatment as policy instruments, which can therefore be optimally chosen by decision makers. Second and at the same time, we study the direct effects of vaccination and treatment on population immunity, in order to understand its role in optimal policies. This contrasts with the effects of NPIs, which only influence the accumulation of population immunity indirectly.

In our analysis, we derive the equilibrium and socially optimal level of treatment with antivirals (that speeds up the rate of recovery) and vaccination (that speeds up the rate of immunity). We consider the incentives of both individuals under decentralized decision making and of a utilitarian social planner under centralized decision making.

We show that while treatment and vaccination have superficial similarities in that they both contribute to accumulating population immunity, the desirability of and external effects associated with these two measures are very different and change across the stages of the epidemic.

For treatment, we find that the external effects of intervention are proportional to the fraction of remaining susceptible individuals, which is decreasing over time. This means that socially optimal treatment may involve a switch from an initial phase with full treatment to a second phase with no treatment. In contrast, for individuals under decentralized decision making, treatment choices are qualitatively different in nature and wholly disconnected from the stage of the epidemic.

For vaccination, we find that the external effects of intervention are proportional to disease prevalence, which is hump-shaped. This means that optimal vaccination may involve as many as two shifts in policy, with no vaccination at early and late stages of the epidemic (when prevalence is low) and full vaccination in the intervening period (when prevalence is sufficiently high). For individuals under decentralized decision making, vaccination choices are qualitatively similar to the socially optimal ones. But as individuals do not take account of externalities, their vaccination decisions will tend to be socially suboptimal.

We next compare our results on socially optimal policies to the traditional non-economic approach to treatment and vaccination policies. We show that not only do the non-economic policies involve lower social welfare than the socially optimal ones, but also that they may be outperformed by completely non-responsive policies that do not maximize any objective at all. These results highlight that in order to formulate effective and welfare-enhancing pharmaceutical interventions, they must be based on an analysis of how to best achieve well-defined social objectives. When such policies explicitly take into account the costs and benefits of inducing population immunity, they can help strike the right balance and lead to improved social wellbeing.

In addition to understanding the nature of vaccination and treatment in shaping the course of the epidemic, our paper also aims to clarify the role of population (or herd) immunity in formulating optimal treatment and vaccination policy. Population immunity is perhaps one of the most central concepts in public health and yet it is widely misunderstood. In part, the confusion stems from the way that epidemiologists have often operationalized the concept. But clarity is important, because the conceptual confusion has led to ambiguous messaging, with public health experts asserting that herd immunity is a desirable outcome, while at the same time proclaiming that there is no explicit strategy in place that relies on herd immunity.¹

Since herd immunity is often the explicit goal guiding vaccine policy, it is therefore useful to first unpack exactly what the concept means and to dispel some unhelpful, but widely held views. What then, is population immunity? According to Fine (1993), population immunity refers to "[...] *the indirect protection afforded to nonimmune individuals by the presence and proximity of others who are immune*". In the same way that a protective measure such as an imperfect vaccine can offer the individual some protection against infection, the presence of immune individuals in the population offers at-risk individuals some indirect protection against infection, because

¹See e.g. statements by policy makers in the UK in a recent news story from the BBC at <https://www.bbc.com/news/uk-53433824>

immune individuals cannot pass on the disease to others. It should be emphasized that whenever individuals get infected and recover with immunity, some amount of population immunity indeed builds up in the population, regardless of what specific policy measures we put in place. In other words, if the biology of the disease is such that recovered people cannot be re-infected, then some population immunity is not only unavoidable, but also desirable as it works to protect those who are still at risk of infection. It is precisely in this sense that some public health practitioners refer to herd immunity as necessary to end the epidemic.

Much of the public health and epidemiology literature holds that population immunity is achieved when some critical level of immune individuals in the population is reached. For example, Brauer and Castillo-Chavez (2012) state that "*A population is said to have herd immunity if a large enough fraction has been immunized to ensure that the disease cannot become endemic*". However, this definition is problematic for two reasons. First, it does not recognize that even if falling short of this critical threshold, the presence of immune individuals still confers some indirect protection to the non-immune. Second and more importantly, it may cause people to think that achieving a threshold level of immunity is a worthwhile policy goal *per se*. As we show in this paper, this is not the case. A socially optimal vaccine or treatment policy that seeks to maximize overall social welfare may well fall short of achieving the critical threshold of immune individuals. Similarly, it may prove optimal to induce immunity in a higher fraction of the population than the so-called herd immunity threshold. This is because the optimal policy is guided by overall welfare considerations, rather than being dictated by a desired level of population immunity.

In summary, inducing population immunity has both costs and benefits and so any optimal policy will involve balancing these across the stages of the epidemic, carefully taking into account both the contemporaneous and intertemporal effects on disease dynamics and welfare.

Our paper sits in between the literatures on treatment and vaccination. The literature on treatment includes contributions by Sanders (1971), Sethi (1974), Sethi and Staats (1978), Goldman and Lightwood (1995, 2002), Gersovitz and Hammer (2004), Rowthorn (2006), Toxvaerd (2009) and Rowthorn and Toxvaerd (2020). In all these analyses, treatment increases the rate of recovery, but individuals do not acquire immunity and thus make a transition back to susceptibility as in the classical susceptible-infected-susceptible (SIS) model. In these models, treatment therefore works by increasing the measure of susceptibles, whereas in the present analysis, treatment works by increasing the measure of recovered individuals.

There is a very large literature on different aspects of the economic control of infectious diseases through vaccination. Chen and Toxvaerd (2014) provide a detailed review and synthesis of this literature. Of direct relevance to the work here is Francis (1997, 2007), who considers fully discounted economic models of vaccination and characterizes optimal and equilibrium outcomes.² Yusuf and Benyah (2012), Bakare et al. (2014), Bakare (2015), Ledzewicz and Schattler (2011) and Joshi et al. (2015) consider treatment and vaccination in compound models, but either use non-standard objective functions or assume no discounting and finite horizons. Arino et al. (2008) also consider the dynamics of a model with both vaccination and treatment, but disregard

²Morton and Wickwire (1974) and Francis (2004) consider infinite horizon undiscounted problems.

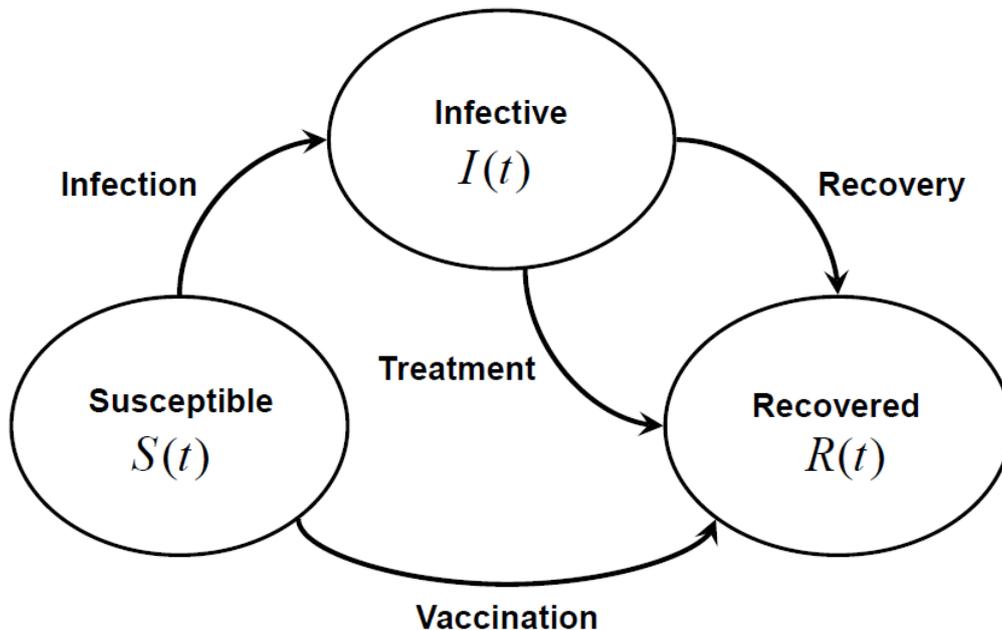


Figure 1: SIR Model with Treatment and Vaccination.

issues of optimal control altogether. In the public health and epidemiology literatures, the research frontier is exemplified by contributions by Milne et al. (2013), Ferguson et al. (2006), Germann et al. (2006) and Longini et al. (2004). These papers deal with issues of feasibility and efficiency, but not with optimality or equilibrium.

Other relevant contributions include Hethcote and Waltman (1973), Barrett (2003), Barrett and Hoel (2007), Auld (2003), Bauch (2005), Bauch and Earn (2004), Boulier et al. (2007), Brito et al. (1991) and Gersovitz and Hammer (2004).³ Last, Gersovitz (2003) considers vaccination in a model with a growing population.

The remainder of the paper is structured as follows: In Section 2, we set out the classical epidemiological and economic versions of the model. In Section 3, we analyze the model with treatment, under decentralized and centralized decision making, respectively. In Section 4, we analyze the model with vaccination, under decentralized and centralized decision making, respectively. In Section 5, we compare our results to the standard non-economic analysis found in the epidemiology literature. In Section 6, we consider the effects of infection-induced mortality. Section 7 contains the Conclusion.

2. THE MODEL

To model the build-up of population immunity, we make use of a classical compartmental model to describe the underlying disease dynamics. We then superimpose an economic model of decision making to understand the interaction between vaccination and treatment decisions and the course of the epidemic.

2.1. The Epidemic SIR Model. The classical *susceptible-infected-recovered* (or SIR) model, used as a building block in this paper, is simple to describe and is il-

³Early non-economic contributions include Anderson and May (1992) and Smith (1964).

illustrated in Figure 1. Time is continuous and runs indefinitely. A closed population $\mathcal{N} = [0, 1]$ consists of a continuum of infinitely lived individuals who can at each instant $t \geq 0$ each be in one of three states, namely *susceptible*, *infected* or *recovered*.⁴ The set of susceptible individuals is denoted by $\mathcal{S}(t)$ and has measure $S(t)$, the set of infected individuals is denoted by $\mathcal{I}(t)$ and has measure $I(t)$ and the set of recovered individuals is denoted by $\mathcal{R}(t)$ and has measure $R(t)$. Because the population size is normalized to one, these measures can be interpreted as fractions.

At each instant, the population mixes homogeneously. The rate at which infection is transferred in such a match is denoted by $\beta > 0$. Coupled with the assumption of homogeneous mixing, this means that the rate at which susceptible individuals become infected is given by $\beta I(t)S(t)$.

Last, individuals spontaneously recover at rate $\gamma \geq 0$. This means that on aggregate, the rate at which recovery occurs is $\gamma I(t)$. Recovered individuals are immune to further infection and also cannot carry the disease.

The dynamic SIR system is described by the following differential equations and initial conditions:

$$\dot{S}(t) = -\beta I(t)S(t) \quad (1)$$

$$\dot{I}(t) = I(t) [\beta S(t) - \gamma] \quad (2)$$

$$\dot{R}(t) = \gamma I(t) \quad (3)$$

$$S(t) = 1 - I(t) - R(t) \quad (4)$$

$$S(0) = S_0 > \gamma/\beta, \quad I(0) = I_0 \approx 0, \quad S_0 + I_0 = 1 \quad (5)$$

The restriction that $S_0 > \gamma/\beta$ ensures that the epidemic can take hold in the population. With this assumption in place, the overall behavior of the system can be described as follows. The measure of susceptible individuals $S(t)$ decreases over time, while the measure of recovered individuals increases over time. In contrast, the measure of infected individuals initially increases, peaks at $S(t) = \gamma/\beta$ and then tends to zero. Note that at the peak of the epidemic, disease prevalence takes the value⁵

$$\bar{I} \equiv I_0 + S_0 - \frac{\gamma}{\beta} \left[\log S_0 - \log \left(\frac{\gamma}{\beta} \right) + 1 \right] \quad (6)$$

It is easily verified that this threshold is decreasing in the recovery rate γ and increasing in the infectiousness β . Both findings are intuitive.

Well-known steps lead to the central result that the final epidemic size is characterized by the equations

$$S(\infty) = 1 - R(\infty) = S_0 \exp(-R(\infty)\mathbb{R}_0) \geq 0 \quad (7)$$

where $\mathbb{R}_0 \equiv \beta/\gamma$ is the basic rate of reproduction (see e.g. Brauer and Castillo-Chavez, 2012). Since $R(0) = 0$, the cumulative incidence, i.e. the total case count across the

⁴The typical duration of an epidemic outbreak is sufficiently short to justify the assumption that there are no births in the model. In a later section, the effects of infection-induced mortality are discussed.

⁵See Brauer and Castillo-Chavez (2012).

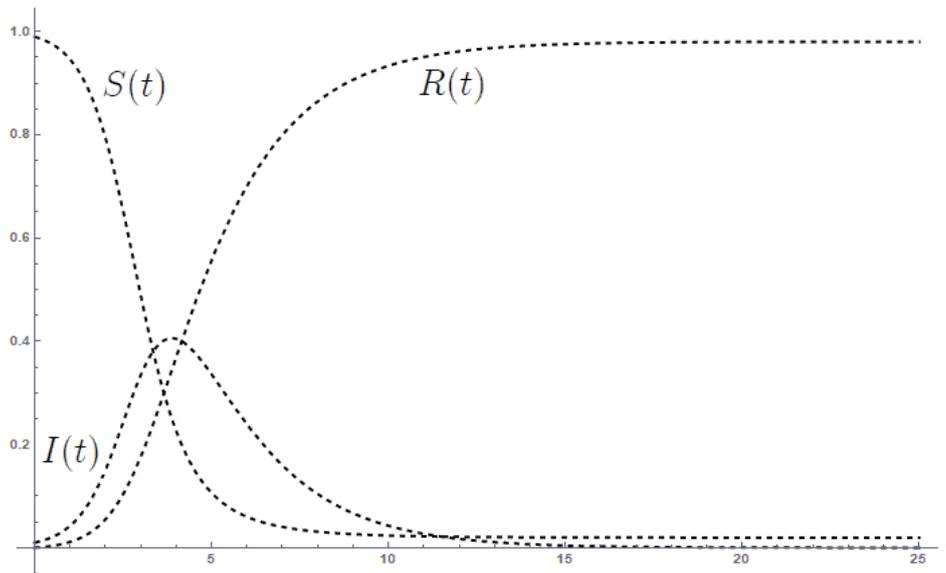


Figure 2: Dynamics in the SIR Model.

epidemic, is given by $R(\infty) = 1 - S(\infty)$. As is to be expected, cumulative incidence is an increasing function of the infectivity parameter β and a decreasing function of the rate of spontaneous recovery γ .

The basic rate of reproduction represents how many secondary infections are caused by the insertion of a single infected individual into a fully susceptible population. The second equation in (7) defines $R(\infty)$ implicitly in terms of parameters and initial conditions and the first equation in (7) defines $S(\infty)$ as the residual, which is possible since $I(\infty) = 0$. The limiting proportions $S(\infty)$ and $R(\infty)$ are easily found for any particular parameterization of the model.

In what follows, the notion of herd immunity will feature prominently. A measure of herd immunity, although imperfect, is the mass of susceptible individuals remaining at the end of the epidemic, $S(\infty)$, since these individuals are not immune yet at no risk of infection.⁶

Note the central role played by the basic rate of reproduction. If $\mathbb{R}_0 < 1$, then infection cannot take hold while if $\mathbb{R}_0 > 1$, then infection first flares up and then tapers off. As will become clear in what follows, the optimal (centralized) control of the epidemic through vaccination or treatment will work by modifying the magnitude of the rate of reproduction \mathbb{R}_0 .⁷

Throughout, we will maintain the following assumption:

Assumption 0: $\beta > \gamma \geq 0$.

Typical dynamics for the SIR model are illustrated in Figure 2.

⁶The reason that it is only a rough measure of herd immunity is that it ignores the time profile of infection. In particular, this measure ignores early protection enjoyed by an individual who eventually becomes infected and recovers.

⁷Note that $R_0 = R(0)$ is the initial condition for the measure of recovered individuals and should not be confused with the basic rate of reproduction \mathbb{R}_0 .

2.2. The Economic Model. Having outlined the classical version of the SIR dynamics, we now make a number of additions in order to turn it into an economically meaningful model. We will consider the economic control of the SIR model via two instruments, namely *treatment* and *vaccination*. Treatment, which is costly, increases the rate at which agents recover (and become immune to further infection). In particular, for some treatment intensity $\tau(t) \in [0, 1]$, the rate at which the individual transitions from $\mathcal{I}(t)$ to $\mathcal{R}(t)$ is given by $\tau(t)\alpha_T + \gamma$, where $\alpha_T > 0$ is interpreted as the efficiency of the treatment. This means that treatment increases the rate of recovery over and above the background rate γ .⁸ Treatment is assumed to cost $c_T > 0$ per instant per individual.

Turning to vaccination, denote by $v(t) \in [0, 1]$ the rate at which individuals are vaccinated. The vaccine offers imperfect protection, being subject to potential failure or delay in achieving effectiveness. Thus vaccination at rate $v(t)$ induces transition from $\mathcal{S}(t)$ to $\mathcal{R}(t)$ at rate $\alpha_V v(t)$, entirely bypassing the class of infected individuals $\mathcal{I}(t)$. The parameter $\alpha_V > 0$ can be interpreted as the speed at which immunity becomes effective.⁹ Vaccination costs $c_V > 0$ per instant per individual.

The transitions between health states caused by vaccination and treatment are illustrated in Figure 1.

Having described the available tools for managing the epidemic, we will describe the payoffs of the decision makers. It will be assumed that the individuals in the sets $\mathcal{S}(t)$, $\mathcal{I}(t)$ and $\mathcal{R}(t)$ earn flow payoffs π_S , π_I and π_R respectively and discount the future at rate $\rho > 0$. We assume that the flow payoffs are ranked as follows:

Assumption 1: $\pi_S \geq \pi_R \geq \pi_I$.

The simplifying assumption $\pi_R = \pi_S$ is common in the vaccination literature, as it allows the modeler to focus on infected versus non-infected individuals, rather than on the three state variables in the present formulation. Instead, we will allow for the possibility of after-effects of infection as these can be serious.¹⁰

For simplicity, we have modeled treatment and vaccination as continuous variables, but they can alternatively be interpreted as randomizations over discrete choices.

In what follows, control variables with a subscript i refer to those of individuals under decentralized decision-making, while control variables without such a subscript refer to the those of a social planner under centralized decision making.

3. DISEASE DYNAMICS UNDER TREATMENT

In this section, we first characterize equilibrium treatment behavior under decentralized decision-making and then characterize the optimal policy under centralized decision-making. As will become clear, the outcomes under centralized decision making will not only differ from the optimal outcome quantitatively, but also qualitatively. This difference will be shown to be intimately related to how the planner and the individuals view the benefits of herd immunity.

⁸The expected time to recovery under continuous treatment is $1/(\gamma + \alpha_T)$.

⁹The expected time to immunity under continuous vaccination is $1/\alpha_V$.

¹⁰One can think of $\pi_R = \pi_S - \varepsilon\pi_A$, where $\varepsilon \in [0, 1]$ is the probability of after-effects and $\pi_A > 0$ is the disutility cost of those effects.

3.1. Treatment Under Decentralized Decision Making. Consider an individual's problem. For any fixed treating intensity $\tau(t)$, the health state of the individual follows a three-state continuous-time Markov process. Fortunately, the actual problem to be solved by an individual can be considerably simplified by noting that in two of these states, the privately optimal choice is trivial. Since treatment is costly, it is optimal for a susceptible or recovered individual to seek no treatment at all. The problem is therefore reduced to determining the privately optimal policy for an infected individual. Without loss of generality, consider an individual who is infected at $t = 0$. Since susceptibility is not feasible for this individual, all he or she is concerned with is the possible transition from the infected to the recovered state. The individual $i \in \mathcal{I}(t)$ then solves the following problem:

$$\max_{\tau_i(t) \in [0,1]} \int_0^{\infty} e^{-\rho t} [Q_{\mathcal{I}}(t) (\pi_{\mathcal{I}} - \tau_i(t)c_T) + (1 - Q_{\mathcal{I}}(t))\pi_{\mathcal{R}}] dt \quad (8)$$

$$s.t. \quad \dot{Q}_{\mathcal{I}}(t) = -Q_{\mathcal{I}}(t) [\alpha_T \tau_i(t) + \gamma], \quad Q_{\mathcal{I}}(0) = 1 \quad (9)$$

where $Q_{\mathcal{I}}(t)$ is the probability of residence in the infected state at time t . This formulation of the individual's problem is the population game formulation developed by Reluga and Galvani (2011). The integrand is simply the expected flow payoff of an individual, while the differential equation governs the evolution of the transition rate between the infected state and the recovered state. Because an infected individual is essentially trading off the costs and benefits of a transition from one state to another, the problem can be reduced to one with a single state variable $Q_{\mathcal{I}}(t)$.

The individual's problem is equivalent to the following simplified problem, which differs only by the constant $\pi_{\mathcal{R}}$:

$$\max_{\tau_i(t) \in [0,1]} \int_0^{\infty} e^{-\rho t} Q_{\mathcal{I}}(t) [\pi_{\mathcal{I}} - \pi_{\mathcal{R}} - \tau_i(t)c_T] dt \quad (10)$$

$$s.t. \quad \dot{Q}_{\mathcal{I}}(t) = -Q_{\mathcal{I}}(t) [\alpha_T \tau_i(t) + \gamma], \quad Q_{\mathcal{I}}(0) = 1 \quad (11)$$

This objective is simply the expected, discounted utility for an individual pursuing treatment strategy $\tau_i(t)$. Note that in steady state, $\dot{Q}_{\mathcal{I}}(t) = 0$ and so it must be that $Q_{\mathcal{I}} = 0$ eventually if $\gamma > 0$, even if no treatment is sought.

The associated current-value Hamiltonian for this problem is given by¹¹

$$H_T^D \equiv Q_{\mathcal{I}}(t) [\pi_{\mathcal{I}} - \pi_{\mathcal{R}} - \tau_i(t)c_T] - \lambda_T^D(t) Q_{\mathcal{I}}(t) [\alpha_T \tau_i(t) + \gamma] \quad (12)$$

where $\lambda_T^D(t)$ is the costate variable. Since the state variable in the individual's problem is the probability of being in the infected state, the costate variable can be interpreted as the shadow value of recovery.

Differentiating the current value Hamiltonian with respect to the treatment rate

¹¹In the individual's problem, an admissible pair of functions $(Q_{\mathcal{I}}(t), \tau_i(t))$ is such that for all $t \geq 0$, $Q_{\mathcal{I}}(t)$ satisfies the differential equation for the state variable $Q_{\mathcal{I}}(t)$ and $\tau_i(t) \in [0, 1]$ is piecewise continuous.

$\tau_i(t)$ yields the following necessary condition for optimality:

$$\frac{\partial H_T^D}{\partial \tau_i(t)} = Q_{\mathcal{I}}(t) [-c_T - \alpha_T \lambda_T^D(t)] = 0 \quad (13)$$

This Hamiltonian condition simply states that the marginal cost of treatment equals the marginal benefit of treatment. To see this, recall that α_T is the rate at which treatment induces recovery and that the costate variable $\lambda_T^D(t)$ is the marginal benefit of recovery.

The evolution of the multiplier is given by the following differential equation:

$$\dot{\lambda}_T^D(t) = \rho \lambda_T^D(t) - \frac{\partial H_T^D}{\partial Q_{\mathcal{I}}(t)} \quad (14)$$

$$= \lambda_T^D(t) [\rho + \alpha_T \tau_i(t) + \gamma] + [\pi_{\mathcal{R}} - \pi_{\mathcal{I}} + \tau_i(t) c_T] \quad (15)$$

It is immediately clear that the net benefit of recovery is independent of the aggregate state of the system and of time. That is, conditional on being infected, the problem is stationary. But then it must be that $\dot{\lambda}_T^D(t) = 0$, which implies that

$$\lambda_T^D(t) = \frac{\pi_{\mathcal{I}} - \pi_{\mathcal{R}} - \tau_i(t) c_T}{\rho + \alpha_T \tau_i(t) + \gamma} \quad (16)$$

Substituting this in the Hamiltonian condition (13) yields the following result:

Proposition 1. *The equilibrium treatment policies under decentralized decision making are given by*

$$\tau_i(t) = 0 \quad \text{for} \quad c_T(\rho + \gamma) > \alpha_T(\pi_{\mathcal{R}} - \pi_{\mathcal{I}}) \quad (17)$$

$$\tau_i(t) \in [0, 1] \quad \text{for} \quad c_T(\rho + \gamma) = \alpha_T(\pi_{\mathcal{R}} - \pi_{\mathcal{I}}) \quad (18)$$

$$\tau_i(t) = 1 \quad \text{for} \quad c_T(\rho + \gamma) < \alpha_T(\pi_{\mathcal{R}} - \pi_{\mathcal{I}}) \quad (19)$$

The privately optimal policy for an individual (i.e. his or her best response function) simply states that treatment should be sought if and only if the expected discounted benefit (to the individual) is larger than the cost of treatment. If the benefit is large enough, then all infected individuals will always seek full treatment and the model reduces to the classical SIR model (but with an increased recovery rate $\alpha_T + \gamma$). If the benefit is not large enough, then no infected individual will ever seek any treatment and the model then reduces to the classical SIR model with recovery rate $\gamma > 0$.

These findings are summarized as follows:

Corollary 2. *Under decentralized decision making: (i) if $c_T(\rho + \gamma) > \alpha_T(\pi_{\mathcal{R}} - \pi_{\mathcal{I}})$, then the equilibrium outcome coincides with that of the SIR model with recovery rate $\gamma > 0$; (ii) if $c_T(\rho + \gamma) < \alpha_T(\pi_{\mathcal{R}} - \pi_{\mathcal{I}})$, then the equilibrium outcome coincides with that of the SIR model with recovery rate $\alpha_T + \gamma$.*

The comparative statics of the privately optimal decentralized policy are straightforward. The higher the discount rate ρ , the recovery rate γ or the treatment cost c_T , the less attractive does treatment become. Conversely, treatment becomes more

attractive the higher the efficiency of the treatment α_T or the higher the recovery premium $(\pi_{\mathcal{R}} - \pi_{\mathcal{I}})$.

It is interesting to note that the treatment decision is not strategic, in the sense that an individual's privately optimal action depends on those of other individuals as would be the case with treatment in a *susceptible-infected-susceptible* model (see e.g. Toxvaerd, 2010). This is because while infected individuals' treatment decisions do influence the prospects of the susceptibles, this influence is ignored by the infected individual since there is no feedback from these decisions to the individual's future welfare. Therefore infected individuals seek treatment if and only if doing so is privately worthwhile, a decision that is not influenced by other infected individuals' treatment decisions. Formally, the lack of strategic interaction follows from the absence of disease prevalence $I(t)$ in the individual's maximization problem.

It is notable that under decentralized decision making, each individual's problem is wholly independent of the aggregate evolution of the epidemic. Disease prevalence $I(t)$ only influences susceptible individuals and not infected or recovered individuals. But the only ones that can actually influence disease prevalence, through the evolution of disease incidence, are the infected individuals (collectively); and they have no direct incentive to do so. This observation is the key difference between the outcomes under centralized and decentralized decision making, which will be explored in detail below.

It is worth emphasizing that herd immunity is a good enjoyed exclusively by susceptible individuals, but provided by recovered individuals. An added twist is that under treatment, it is determined by the infected individuals how much of this good is provided. In other words, the benefits flow from the decisions of individuals in one class to individuals in another class to which the former can never return (and hence from which they will themselves never benefit).

To appreciate how inefficient equilibrium treatment can be, suppose that treatment is too costly to be privately optimal but sufficiently inexpensive for the social planner to want to treat. Furthermore, suppose $S(0) = 1 - \varepsilon$ and $I(0) = \varepsilon$, with $\varepsilon > 0$ arbitrarily small. It is clear that a desirable outcome would involve immediate treatment of the small group of infected individuals and asymptotic eradication with treatment. Yet however small ε is, this would not happen in equilibrium.

3.2. Treatment Under Centralized Decision Making. The problem of the utilitarian central planner is as follows:

$$\max_{\tau(t) \in [0,1]} \int_0^{\infty} e^{-\rho t} [S(t)\pi_{\mathcal{S}} + I(t)(\pi_{\mathcal{I}} - \tau(t)c_T) + R(t)\pi_{\mathcal{R}}] dt \quad (20)$$

In the planner's problem, the policy $\tau(t) \in [0, 1]$ can interchangeably be thought of as the proportion of infected individuals who undergo treatment, or as the intensity of treatment that all individuals are subjected to.

The problem is solved subject to the following laws of motion for the measures of susceptible, infected and recovered individuals, respectively:

$$\dot{S}(t) = -\beta I(t)S(t) \quad (21)$$

$$\dot{I}(t) = I(t) [\beta S(t) - \alpha_T \tau(t) - \gamma] \quad (22)$$

$$\dot{R}(t) = I(t) [\alpha_T \tau(t) + \gamma] \quad (23)$$

$$S(t) = 1 - I(t) - R(t) \quad (24)$$

$$S(0) = S_0 > \gamma/\beta, \quad I(0) = I_0 \approx 0, \quad S_0 + I_0 = 1 \quad (25)$$

The problem solved by the central planner is similar to that solved by individuals under decentralized decision making, but there are some notable differences. First, the planner aggregates the welfare of all individuals into its objective function. Second, the constraints take into account the fact that the planner directly controls the evolution of the aggregate variables through its choice of aggregate treatment. Therefore the fractions $S(t)$, $I(t)$ and $R(t)$ are endogenous for the planner, whereas they are exogenous for any one individual.

In considering the overall effects of treatment, it is useful to make the following analogy. Since recovery confers immunity on the (previously infected) individual, treatment may be interpreted as a kind of immunization at the aggregate level. Immunization transfers susceptible individuals directly into the recovered class; therefore it dilutes the effects of infection, since the rate of contact between infected and susceptible individuals is reduced.¹² Treatment has a similar effect by transferring individuals from $\mathcal{I}(t)$ to $\mathcal{R}(t)$, rather than from $\mathcal{S}(t)$ to $\mathcal{R}(t)$ as is the case with vaccination. Note, however, that from the perspective of the particular individual, treatment and vaccination are quite different in that the former presupposes that the individual has a spell of infection, while the latter does not. These differences make optimal treatment and vaccination policies qualitatively different, as will be discussed further below.

Returning to the characterization of the optimal treatment policy, using that $S(t) = 1 - I(t) - R(t)$, the planner's current-value Hamiltonian can be written as

$$\begin{aligned} H_T^C \equiv & [1 - I(t) - R(t)] \pi_S + I(t) (\pi_I - \tau(t)c_T) + R(t)\pi_R \\ & + \lambda_T^C(t) I(t) [\beta (1 - I(t) - R(t)) - \alpha_T \tau(t) - \gamma] \\ & + \mu_T^C(t) I(t) [\alpha_T \tau(t) + \gamma] \end{aligned} \quad (26)$$

Note that $\lambda_T^C(t)$ and $\mu_T^C(t)$ are the costate variables associated with the laws of motion for infected and recovered individuals, respectively. Because of the normalization of the population size, we can treat this as a maximization problem with only two state variables, $I(t)$ and $R(t)$.

Differentiating with respect to the treatment rate $\tau(t)$ yields the following necessary Hamiltonian condition for optimality:

$$\frac{\partial H_T^C}{\partial \tau(t)} = -I(t) [c_T + \alpha_T (\lambda_T^C(t) - \mu_T^C(t))] = 0 \quad (27)$$

This condition equates the marginal social cost of treatment with its marginal social

¹²This is simply because a number of infected and/or susceptible individuals are matched with recovered individuals instead of with each other.

benefit.

The evolution of the multipliers is governed by the following system of differential equations:

$$\dot{\lambda}_T^C(t) = \rho \lambda_T^C(t) - \frac{\partial H_T^C}{\partial I(t)} \quad (28)$$

$$= \lambda_T^C(t) [\rho + \gamma + \alpha_T \tau(t) + \beta (2I(t) + R(t) - 1)] - \mu_T^C(t) [\gamma + \alpha_T \tau(t)] - (\pi_{\mathcal{I}} - \tau(t)c_T - \pi_{\mathcal{S}}) \quad (29)$$

$$\dot{\mu}_T^C(t) = \rho \mu_T^C(t) - \frac{\partial H_T^C}{\partial R(t)} \quad (30)$$

$$= \rho \mu_T^C(t) + \lambda_T^C(t) \beta I(t) + (\pi_{\mathcal{S}} - \pi_{\mathcal{R}}) \quad (31)$$

From the Hamiltonian condition (27), it follows that the optimal policy is of the bang-bang type and can be characterized as follows:

Proposition 3. *The socially optimal treatment policy under centralized decision making is given by*

$$\tau(t) = 0 \quad \text{for } c_T > \alpha_T (\mu_T^C(t) - \lambda_T^C(t)) \quad (32)$$

$$\tau(t) \in [0, 1] \quad \text{for } c_T = \alpha_T (\mu_T^C(t) - \lambda_T^C(t)) \quad (33)$$

$$\tau(t) = 1 \quad \text{for } c_T < \alpha_T (\mu_T^C(t) - \lambda_T^C(t)) \quad (34)$$

This policy has a nice interpretation. Increasing the treatment rate has two effects, namely to increase the measure of recovered individuals and to reduce the measure of infected individuals. The marginal benefit of reducing the measure of infectives is $-\lambda_T^C(t) \geq 0$, while the marginal benefit of increasing the measure of recovered individuals is $\mu_T^C(t) \geq 0$. Thus the expression $\alpha_T (\mu_T^C(t) - \lambda_T^C(t))$ is simply the rate at which the combined benefits of treatment accrue. The optimal policy is thus dictated by a comparison of the marginal cost and the marginal benefit of treatment, the latter of which is a function of the state of the system and thus not constant, as was the case under decentralized decision making.

3.3. Optimal Treatment Across the Stages of the Epidemic. To understand how socially optimal treatment changes across the stages of the epidemic, it is instructive to consider the source of externalities and how the strength of these depend on the state of the system.

The source of the externalities in this model is the effect that infected individuals have on the susceptible individuals. Specifically, the externality is that individuals only care about their own transition from $\mathcal{I}(t)$ to $\mathcal{R}(t)$, without taking into account the welfare of individuals in $\mathcal{S}(t)$. But note that since the measure of susceptibles $S(t)$ is monotonically decreasing over time, so are the externalities. This means that the external effects are strongest at the beginning of the epidemic and then decrease over time, causing the wedge between private and social values to narrow. In terms of policy, this means that unless treatment is prohibitively expensive, it will typically be optimal to treat to the maximal extent possible from the outset. The only remaining question is then whether the direct private and the external effects remain sufficiently high to continue to treat in perpetuity.

There are in principle two regimes to consider, both starting with a policy of full treatment. In the first regime, even though treatment is at its maximum level, the susceptibles decrease over time and converge to some positive limit. In approaching this limit, the social value of treatment remains sufficiently high throughout to justify the cost of treatment. Thus there is in this case no switch in the treatment policy. In the second regime, the planner also starts out by fully treating anyone who becomes infected, but now the social value of treatment decreases sufficiently over time to become lower than the cost of treatment. Once this happens, it is no longer justifiable to continue treatment and thus it is discontinued. After this point, the dynamics mirror those of a standard SIR epidemic (with a suitably modified initial condition).

To summarize, unless treatment is prohibitively expensive, the optimal treatment policy will start with full treatment and contain *at most one switch to no treatment*. Whether there is a switch depends on the biomedical and preference parameters and on the limiting distribution of the measure of susceptibles $S(\infty)$.

3.4. A Simulated Example of Optimal Treatment. To appreciate the character of optimal treatment, it is useful to consider a simulated example of the evolution of the system. Figure 3 illustrates the evolution of $S(t)$, $I(t)$ and $R(t)$ over time. The solid lines show the paths under optimal treatment, while the dotted lines show the paths of the uncontrolled system. This allows us to highlight the differences due to policy intervention. Under socially optimal treatment, the recovery rate at the initial stages of the epidemic is effectively increased from γ to $(\gamma + \alpha_T)$. This results in a less steep increase in prevalence $I(t)$ and as a consequence, a less marked decline in the susceptibles $S(t)$. The effect on the recovered individuals $R(t)$ is in general ambiguous, because of two competing effects. On the one hand, there is an increase that stems from the contribution of treated individuals. On the other hand, there is a decrease in disease prevalence, which reduces the contribution from natural recovery.

In the figure, it is seen that for the chosen parameterization, the initial treatment essentially increases the recovery rate, relative to the uncontrolled model. This causes a decrease in peak prevalence and to an overall delay in the peak of the epidemic. Across the entire epidemic, the initial phase with treatment reduces the overall number of infections relative to the no-treatment outcome.

Some sample simulation results are summarized in the lower panel of Table 1. We find in our simulations that higher treatment costs or lower treatment efficiency leads the optimal treatment to stop earlier, relative to the base case. Higher infectiousness or lower recovery rate also lead to earlier cessation of treatment. In contrast, the more severe the symptoms of infections become, the longer is the optimal treatment sustained. Last, if treatment is not started at the outset but delayed till infection has taken hold, say because treatment was not available at the start of the epidemic, then the treatment is implemented for a shorter time.

4. DISEASE DYNAMICS UNDER VACCINATION

In this section, we turn our attention to equilibrium and to socially optimal outcomes under vaccination. This will enable us to make a clear comparison with the outcomes under treatment. As emphasized in the introduction, both treatment and vaccination work by boosting the mass of recovered individuals, thereby inducing herd immunity to protect the remaining susceptible individuals. But the similarities turn out to be

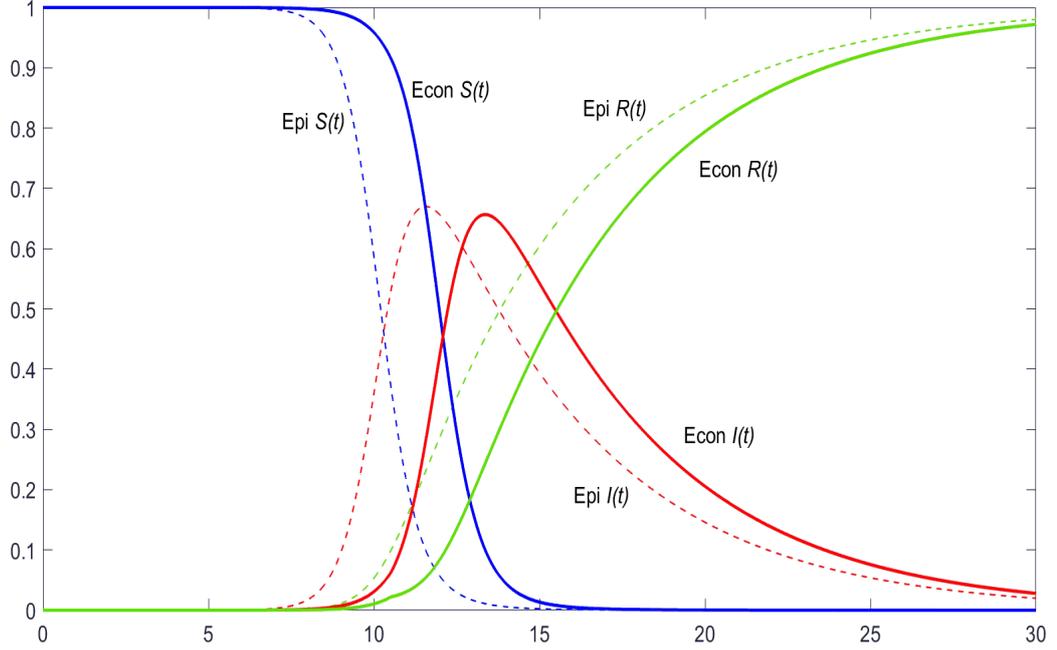


Figure 3: Dynamics Under Socially Optimal Treatment.

superficial only. In this section, we show that seemingly subtle differences between treatment and vaccination have radical effects on equilibrium outcomes and that even the optimal policies differ qualitatively.

4.1. Vaccination Under Decentralized Decision Making. First, consider the vaccination problem of individual $i \in \mathcal{S}(t)$, which is as follows:

$$\max_{v_i(t) \in [0,1]} \int_0^\infty e^{-\rho t} [Q_{\mathcal{S}}(t)(\pi_{\mathcal{S}} - v_i(t)c_V) + Q_{\mathcal{I}}(t)\pi_{\mathcal{I}} + Q_{\mathcal{R}}(t)\pi_{\mathcal{R}}] dt \quad (35)$$

Here, $Q_j(t)$, $j = \mathcal{S}, \mathcal{I}, \mathcal{R}$ is the probability of individual i inhabiting health state j at time $t \geq 0$. The relevant constraints are given by

$$\dot{Q}_{\mathcal{S}}(t) = -Q_{\mathcal{S}}(t) [\beta I(t) + \alpha_V v_i(t)], \quad Q_{\mathcal{S}}(0) = 1 \quad (36)$$

$$\dot{Q}_{\mathcal{I}}(t) = Q_{\mathcal{S}}(t)\beta I(t) - Q_{\mathcal{I}}(t)\gamma \quad (37)$$

$$\dot{Q}_{\mathcal{R}}(t) = Q_{\mathcal{S}}(t)\alpha_V v_i(t) + Q_{\mathcal{I}}(t)\gamma \quad (38)$$

The current-value Hamiltonian for the individual's problem can be written as

$$H_V^D \equiv (1 - Q_{\mathcal{I}}(t) - Q_{\mathcal{R}}(t))(\pi_{\mathcal{S}} - v_i(t)c_V) + Q_{\mathcal{I}}(t)\pi_{\mathcal{I}} + Q_{\mathcal{R}}(t)\pi_{\mathcal{R}} \quad (39)$$

$$+ \mu_V^D(t) [(1 - Q_{\mathcal{I}}(t) - Q_{\mathcal{R}}(t))\alpha_V v_i(t) + Q_{\mathcal{I}}(t)\gamma] \quad (40)$$

$$+ \lambda_V^D(t) [(1 - Q_{\mathcal{I}}(t) - Q_{\mathcal{R}}(t))\beta I(t) - Q_{\mathcal{I}}(t)\gamma] \quad (41)$$

where $\mu_V^D(t)$ and $\lambda_V^D(t)$ are the costate variables for the laws of motion for $Q_{\mathcal{R}}(t)$ and $Q_{\mathcal{I}}(t)$, respectively. Note that because the three state variables in the individual's

maximization problem are probabilities and thus have to sum to one, the problem can be reduced to one with just two state variables $Q_{\mathcal{I}}(t)$ and $Q_{\mathcal{R}}(t)$. Differentiating with respect to the vaccination rate $v_i(t)$ yields the following necessary Hamiltonian condition for optimality:

$$\frac{\partial H_V^D}{\partial v_i(t)} = Q_{\mathcal{R}}(t) [-c_V + \alpha_V \mu_V^D(t)] = 0 \quad (42)$$

This expression simply equates the marginal cost and marginal benefit of vaccination.

The evolution of the multipliers is given by the following differential equations:

$$\dot{\mu}_V^D(t) = \rho \mu_V^D(t) - \frac{\partial H_V^D}{\partial Q_{\mathcal{R}}(t)} \quad (43)$$

$$= \mu_V^D(t) [\rho + \alpha_V v_i(t)] + \lambda_V^D(t) \beta I(t) + [\pi_{\mathcal{S}} - \pi_{\mathcal{R}} - v_i(t) c_V] \quad (44)$$

$$\dot{\lambda}_V^D(t) = \rho \lambda_V^D(t) - \frac{\partial H_V^D}{\partial Q_{\mathcal{I}}(t)} \quad (45)$$

$$= \lambda_V^D(t) [\rho + \beta I(t) + \gamma] + \mu_V^D(t) [\alpha_V v_i(t) - \gamma] + [\pi_{\mathcal{S}} - \pi_{\mathcal{I}} - v_i(t) c_V] \quad (46)$$

We can therefore state the following result:

Proposition 4. *The equilibrium vaccination policies under decentralized decision making are given by*

$$v_i(t) = 0 \quad \text{for } c_V > \alpha_V \mu_V^D(t) \quad (47)$$

$$v_i(t) \in [0, 1] \quad \text{for } c_V = \alpha_V \mu_V^D(t) \quad (48)$$

$$v_i(t) = 1 \quad \text{for } c_V < \alpha_V \mu_V^D(t) \quad (49)$$

To understand how the individual's incentives to vaccinate change across the stages of the epidemic, it is useful to first consider the benchmark studied in Francis (1997) and Chen and Toxvaerd (2014). First, the incentive to vaccinate is proportional to the infection hazard, which is itself proportional to disease prevalence. Thus higher infection levels in the population increases the value of vaccination for the individual. Suppose now that the vaccine is perfect and that immunity is instantaneous. In this case, there is no strategic interaction between individuals' vaccination decisions. The reason is that if people are homogeneous, they all face the same risk and therefore all vaccinate at the same time. But since vaccination offers perfect immunity, a vaccinated individual is unaffected by the vaccination decisions of others. In other words, there is in this case no population immunity in equilibrium. But this is not a robust finding. As shown in Chen and Toxvaerd (2014), with heterogeneous individuals, some will decide to vaccinate earlier than others (i.e. have lower critical thresholds of disease prevalence), and thus the latter benefit from the formers' vaccination.¹³ Similarly, if vaccination only confers immunity with a delay, as is the case in the present setup, then even vaccinating individuals benefit from the vaccination of others till the moment they become immune themselves. In both cases, equilibrium will involve some measure of population immunity, although not as much as that induced by a socially

¹³This point is also made in Fine et al. (2011).

optimal vaccination policy. As disease prevalence is non-monotone in this model, so are the incentives to vaccinate. At the start and the end of the epidemic, when prevalence is low, so is the private value of vaccination. Thus individuals only vaccinate when prevalence is sufficiently high and cease doing so once prevalence has become sufficiently low (in case the individual's vaccination has not yet induced immunity).

4.2. Vaccination Under Centralized Decision Making. Turning to centralized decision-making, the planner's problem can be written as follows:

$$\max_{v(t) \in [0,1]} \int_0^\infty e^{-\rho t} [S(t)(\pi_S - v(t)c_V) + I(t)\pi_I + R(t)\pi_R] dt \quad (50)$$

The relevant constraints are then

$$\dot{S}(t) = -S(t) [\beta I(t) + \alpha_V v(t)] \quad (51)$$

$$\dot{I}(t) = I(t) [\beta S(t) - \gamma] \quad (52)$$

$$\dot{R}(t) = \gamma I(t) + S(t)\alpha_V v(t) \quad (53)$$

$$S(t) = 1 - I(t) - R(t) \quad (54)$$

$$S(0) = S_0 > \gamma/\beta, \quad I(0) = I_0 \approx 0, \quad S_0 + I_0 = 1 \quad (55)$$

Strictly speaking, it is not necessary for $R(0) = 0$ exactly; we only require that there are sufficiently few immune individuals at the outset to make vaccine-induced immunity socially desirable.

The current-value Hamiltonian for the planner's problem is given by

$$H_V^C \equiv [1 - I(t) - R(t)](\pi_S - v(t)c_V) + I(t)\pi_I + R(t)\pi_R \quad (56)$$

$$+ \mu_V^C(t) [\gamma I(t) + (1 - I(t) - R(t))\alpha_V v(t)] \quad (57)$$

$$+ \lambda_V^C(t) I(t) [\beta(1 - I(t) - R(t)) - \gamma] \quad (58)$$

Note that $\mu_V^C(t)$ and $\lambda_V^C(t)$ are the costate variables associated with the laws of motion for recovered and infected individuals, respectively. Because of the normalization, this can be treated as a maximization problem with only two state variables $I(t)$ and $R(t)$. Differentiating with respect to the vaccination rate $v(t)$ yields the following necessary condition for optimality:

$$\frac{\partial H_V^C}{\partial v(t)} = S(t) [-c_V + \alpha_V \mu_V^C(t)] = 0 \quad (59)$$

This expression simply equates the marginal social cost of increasing the vaccination coverage with its marginal social benefit. This can be seen from the fact that $\mu_V^C(t)$ measures the social value of a marginal increase in recovered individuals, while α_V is the rate at which vaccination induces immunity.

The laws of motion for the two costate variables are given by

$$\dot{\mu}_V^C(t) = \rho\mu_V^C(t) - \frac{\partial H_V^C}{\partial R(t)} \quad (60)$$

$$= \mu_V^C(t)[\rho + \alpha_V v(t)] + \lambda_V^C(t)\beta I(t) + [\pi_S - \pi_R - v(t)c_V] \quad (61)$$

$$\dot{\lambda}_V^C(t) = \rho\lambda_V^C(t) - \frac{\partial H_V^C}{\partial I(t)} \quad (62)$$

$$= \lambda_V^C(t)[\rho - \beta(1 - I(t) - R(t)) + \gamma + \beta I(t)] \\ + \mu_V^C(t)[\alpha_V v(t) - \gamma] + [\pi_S - \pi_I - v(t)c_V] \quad (63)$$

We can now state the following result:

Proposition 5. *The socially optimal vaccination policy under centralized decision making is given by*

$$v(t) = 0 \quad \text{for } c_V > \alpha_V \mu_V^C(t) \quad (64)$$

$$v(t) \in [0, 1] \quad \text{for } c_V = \alpha_V \mu_V^C(t) \quad (65)$$

$$v(t) = 1 \quad \text{for } c_V < \alpha_V \mu_V^C(t) \quad (66)$$

We briefly compare the dynamics under centralized and decentralized decision making. Aggregating the laws of motion (36), (37) and (38) across the population, we recover the aggregate laws of motion (51), (52) and (53), respectively. This means that controlling for aggregate vaccination, the dynamics are exactly the same under both decision-making assumption. But it is easily verified that the costate variables in the individual's problem are different from those of the social planner, because the latter maximizes aggregate social welfare while the former only maximizes its own well-being. This means that from a social perspective, the individual does not value vaccination as much as the planner does and will therefore tend to under-vaccinate. For a detailed decomposition of such external effects, see Rowthorn and Toxvaerd (2020).

Optimal Vaccination Across the Stages of the Epidemic. To understand how socially optimal vaccination changes across the epidemic, we again consider the source of externalities. Under vaccination, the source of externalities is the effect that vaccinating susceptibles have on other susceptibles. This is because an immune individual cannot become infected and so cannot pass on infection to third parties. But whether the immunity of the vaccinated individual actually comes into play at all depends on disease prevalence (because this is the quantity that determines infection risk) and so we find that the external effects of vaccination are proportional to disease prevalence. But as we have already seen, in the SIR model prevalence is non-monotone, first increasing and then decreasing. This means that the external effects of vaccination are also non-monotone, first increasing and then decreasing, approaching zero as the disease dies out. In combination with the bang-bang nature of the optimal policy, this implies that along an optimal path, there can be as many as two switches in vaccination policy. Again, there are two regimes to consider. In the first regime, the value of vaccination is never high enough to merit costly vaccination, even at the peak of the epidemic. In the second regime, things are more complicated. At initial stages of the epidemic when infection is negligible, the optimal policy will be to not

vaccinate anyone (unless it is costless or there is no discounting). As infection picks up, the social value of vaccinations increases and reaches a point at which it outweighs the cost of vaccinations. Thus there is a switch in the optimal policy and the planner then vaccinates as many individuals as possible. Even with full vaccination, infection may keep rising for a while, if vaccine effect is delayed or confers imperfect protection. But eventually, disease prevalence will peak and infections start to decrease. At this point, the social value of vaccination also starts decreasing, slowly approaching zero as infection dies out. In terms of optimal policy, this implies that there is a point after which the social value of vaccination is so small that even if there are still remaining at-risk individuals in the population (who did not achieve immunity during the vaccination drive), it is no longer worthwhile vaccinating them. At this point, there is an additional switch in the optimal vaccine policy, back to no vaccination.

Note that while the socially optimal treatment policy was different in nature to the privately optimal treatment decisions under decentralized decision making, this is not so under vaccination. The difference here is one of degree, not of kind. For an individual contemplating getting vaccinated, the external effects play no role. Instead, what will determine the vaccination decision will be the disease risk facing the individual. But even disregarding the external effects which are central to the socially optimal policy, the individual disease risks are also proportional to disease prevalence and are thus also non-monotone. In other words, the main difference between equilibrium and socially optimal vaccine uptake is that individuals attach a lower value to vaccination than does the social planner. This is a fairly general insight.

In the special case $\gamma = 0$, $\alpha_V = 1$ and $\pi_S = \pi_R$, it is known that both the individuals' privately optimal vaccination decisions and the planner's optimal policy are of the bang-bang variety and that their critical thresholds exactly coincide. Equilibrium in this special case is thus socially optimal and there are neither external effects in equilibrium nor any herd immunity (see Francis, 1997 and Chen and Toxvaerd, 2014 for a detailed exposition of this property). In the limit, $I(\infty) + R(\infty) = 1$, so no-one benefits from herd immunity.

The threshold property of socially optimal and equilibrium vaccination completely describe the outcomes in this special case, involving at most one switch from no vaccination to full vaccination. This is because if $\gamma = 0$, the setup is a modified SI type model in which disease prevalence $I(t)$ is indeed monotone.¹⁴

In the case with $\gamma > 0$, $I(t)$ is non-monotone and the optimal policy has only been partially characterized (see Sethi and Staats, 1978 and Francis, 2007).¹⁵ The optimal policy is still of the bang-bang variety and is never singular (i.e. interior) for any positive interval of time.¹⁶ Furthermore, it is known that there can be at

¹⁴Even though the presence of vaccine imperfections and after-effects of infection does modify the relevant critical thresholds, the monotonicity property of disease prevalence vis-à-vis the recovery rate γ remains unchanged.

¹⁵In Francis (2007), even though $\gamma > 0$ and disease prevalence is hump-shaped, equilibrium behaviour still exhibits only one switch because vaccination is assumed to yield instantaneous immunity. This means that as soon as the hazard of infection is high enough, all remaining susceptible individuals immediately and effectively vaccinate. From their perspective, it is at this point irrelevant whether aggregate disease prevalence will eventually decrease. When vaccination is imperfect or only effective with a delay as in the present model, this is no longer the case.

¹⁶Sethi and Staats (1978) consider a finite time horizon and focus mostly on the undiscounted case.

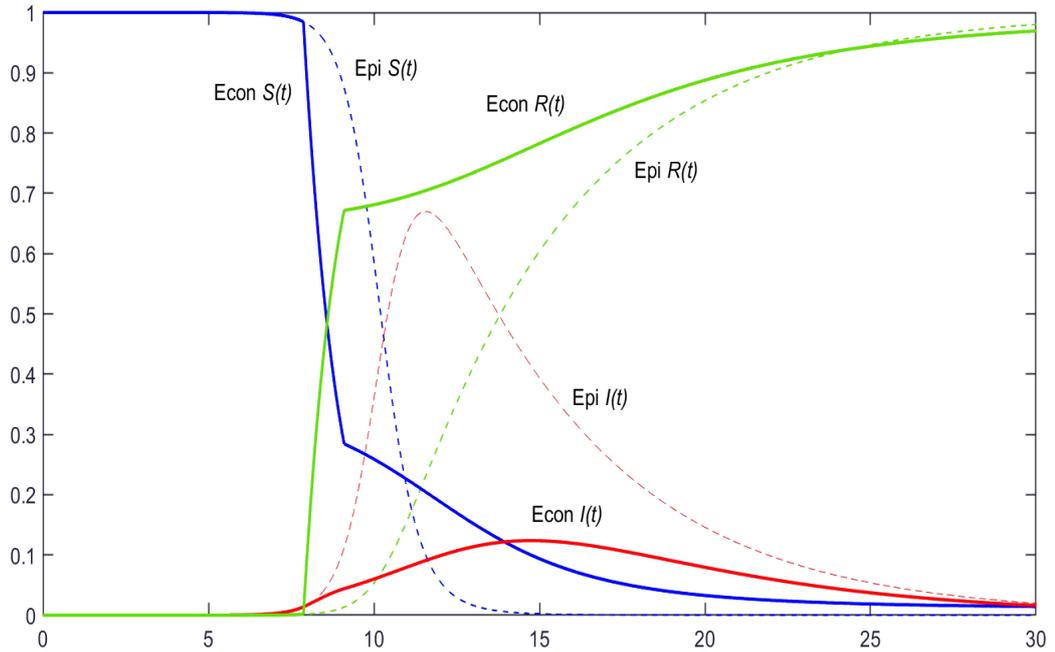


Figure 4: Dynamics Under Socially Optimal Vaccination.

most two switches in policy and that the last switch is always from full vaccination to no vaccination. So the planner, if it finds it optimal to vaccinate at all, either vaccinates at the outset and then ceases doing so, or vaccination is initially delayed, then implemented at full force and then eventually ceased. Despite the difficulty in formally characterizing the optimal policy in this case, the main features are intuitively clear. At early stages of the epidemic, the hazard is moderate and thus it is optimal not to vaccinate, as is the case when $\gamma = 0$. At some point, the optimal policy switches and full vaccination is implemented. This is done in part with a view to benefit from herd immunity (i.e. the planner values vaccination higher than do the individuals, who only maximize their own welfare and disregard the social benefits flowing from their vaccination). As vaccination is pursued, in conjunction with natural recovery, the mass of recovered and immune becomes sufficiently large that the remaining non-vaccinated susceptibles are effectively protected by herd immunity. At this point, the optimal policy switches back and no further vaccination is pursued. The fully discounted problem has received surprisingly little attention. Exceptions include Sethi and Staats (1978) and Francis (1997, 2007). Of these, none consider imperfect vaccines and only the latter contribution features spontaneous recovery. It is precisely the combination of imperfect vaccination and the possibility of spontaneous recovery in our model that complicates the analysis and prevents us from obtaining further analytical results.

4.3. A Simulated Example of Optimal Vaccination. We now consider a simulated example of socially optimal vaccination. Figure 4 shows the evolution of the optimally controlled system in solid lines and of the uncontrolled system in dotted lines.

Under socially optimal vaccination, the initial dynamics coincide with those of the classical uncontrolled model until prevalence $I(t)$ reaches a critical threshold. At this point, full vaccination is implemented, causing an increase in the measure of immune individuals $R(t)$. At the same time, the measure of susceptibles $S(t)$ decreases more sharply from this point onwards (because successfully vaccinated individuals are drawn from that set of individuals). This means that there are fewer individuals that can become infected, which explains why the increase in infected individuals $I(t)$ is attenuated.

It is worth noting that while there is full vaccination of all susceptible individuals, the number of infections keeps increasing regardless. This is an artefact of the possibility that the vaccine only confers immunity with a delay. The faster the vaccine works, the fewer new infections will there be once widespread vaccination is implemented. Last, note that when infection again falls below some critical threshold, it is optimal to cease vaccination of any remaining susceptibles. The switches in the optimal vaccination policy cause there to be kinks in the curves. These are most visible in the path of susceptible individuals but can also be found in the paths of the other state variables.

Overall, vaccination causes peak prevalence to be significantly reduced, and it is also somewhat delayed. It is also seen that the susceptibles converge to a higher level under vaccination than under the uncontrolled epidemic, a fact that stems from the population immunity that the vaccination confers on susceptible individuals.

Next, we will briefly outline the results of our simulations, which are summarized in the upper panel of Table 1. The table reports results on a number of key outcomes from a base case and for a number of simulations in which we have varied each of the parameters relative to the base case.

We see in the simulation results that when the cost of vaccination increases or the efficiency of the vaccination decreases, the optimal vaccination policy involves a delay in vaccination relative to the base case. In contrast, higher infectiousness or a lower recovery rate, biological measures of a worse infection, both lead to an earlier start of vaccination. Similarly, the worse the symptoms from infection become, the earlier should vaccination commence. Last, if vaccination is not started before the infection has taken hold, say because a vaccine was not yet available at the start of the epidemic, then the optimal vaccination policy will prescribe that vaccination should be sustained for a longer period of time.

5. COMPARISONS TO NON-ECONOMIC CONTROL

To put this analysis into perspective, we will next consider the typical non-economic approach to infection control common in the public health and epidemiology literature and to compare it to the optimal policies characterized in the previous sections. As noted in the introduction, the mathematical epidemiology literature has featured a multitude of analyses focused almost entirely around the basic rate of reproduction in different settings and on how the public health authority may bring about eradication by influencing this rate. For a discussion of this literature and the central role accorded to herd immunity and the basic rate of reproduction, see e.g. Fine (1993) and Fine et al. (2011).

First, recall the special role played by the basic rate of reproduction $\mathbb{R}_0 = \beta/\gamma$. For $\mathbb{R}_0 < 1$, the disease starts decreasing immediately and dies out over time while for

$\mathbb{R}_0 \geq 1$, the epidemic takes hold in the population. Perhaps not surprisingly, a large part of the traditional analysis of disease control takes \mathbb{R}_0 as the implicit objective function. We will show how this approach is inconsistent with standard cost-benefit analysis.

5.1. Vaccinating to Achieve Population Immunity. The textbook analysis of imperfect (leaky) vaccination argues that if a proportion $v \in [0, 1]$ of the population is immunized, then the basic rate of reproduction changes to

$$\mathbb{R}_0^V \equiv \frac{[1 - \alpha_V v]\beta}{\gamma} = [1 - v\alpha_V]\mathbb{R}_0 \quad (67)$$

In this equation, one can think of the effectiveness of the vaccine $\alpha_V \equiv (1 - \phi)$, where $\phi \in [0, 1]$ is the probability that the vaccine fails or as the fraction of the population for which the vaccine does not work. In order to force the basic reproductive ratio below one, the health authority must vaccinate at least a fraction

$$v \geq \bar{v} \equiv \frac{\mathbb{R}_0 - 1}{\mathbb{R}_0 \alpha_V} \quad (68)$$

of the population. Indeed, this is the typical policy prescription in most epidemiology textbooks (see e.g. Keeling and Rohani, 2008). Whether this is feasible depends on parameter values.

Note that the critical threshold is increasing in the failure probability ϕ . The reason is simply that as vaccines become increasingly ineffective in securing immunity, one must vaccinate a larger proportion of the population to ensure that the needed proportion of immune individuals in the population is reached.

The weakness of this analysis is that it completely disregards the desirability of using vaccination at different stages of the epidemic (for the individual or for the planner) and therefore ignores the important tradeoffs that determine privately and socially optimal decisions. To make meaningful policy recommendations and sensible comparisons between instruments, one must explicitly consider the problems faced by decision makers under centralized and decentralized decision making. Furthermore, one must fully take into account that any decision, whether on treatment or vaccination, has both costs and benefits and that these must be appropriately balanced. The purely mechanistic approach inherent in the \mathbb{R}_0 -focused analysis may lead to socially undesirable outcomes and simplistic policy recommendations.¹⁷

There is a stark contrast between the economic and non-economic control of the disease, best exemplified by the case of an imperfect vaccine. Under the non-economic approach, a higher failure rate ϕ would call for a higher critical vaccination threshold to be achieved (to ensure eradication). In contrast, the economic analysis shows that a higher failure rate can be understood as an effective increase in the cost of vaccination (or as a decrease in the benefits to vaccination), thereby reducing the desirability of

¹⁷The treatment of vaccination in Keeling and Rohani (2008) is a good example of the ambiguity on policy issues in the existing literature. While noting that it may be infeasible to eradicate the disease through vaccination, they note that some vaccination may still be desirable. Yet, they stop short of letting the desirability of vaccination be the guiding principle in the formulation of a vaccine policy.

vaccination. Thus, *ceteris paribus*, it is optimal to vaccinate at a lower rate. Intuitively, similar points would apply for a decrease in treatment efficiency α_T . In a sense, the non-economic approach confounds feasibility with desirability. Some of these issues will become apparent in our simulation results that we report in what follows.

While the classical approach identifies the critical threshold of susceptible individuals that needs to be immunized in order for infection to start decreasing, it is not clear about how best to achieve this outcome. There are literally infinitely many ways in which the outcome can be achieved, but there are two natural candidates that deserve special attention. These are so-called *pulse vaccination* and *continual vaccination*. Under pulse vaccination, a fraction of the susceptible population all receive vaccination at the same time—say, at the outset of the epidemic—to ensure that the desired level of immunized individuals comes about. This can also be followed up by additional pulses as needed to boost the effects of the vaccination. Under pulse vaccination, only a highly effective vaccine can ensure that the critical level of immunity is achieved in a single pulse. In contrast, under continual vaccination, a fraction of the population is immunized over time till the cumulative measure of immunized individuals reaches the desired level. While neither of these vaccination policies maximize social welfare, it is interesting to explicitly consider these as they are used extensively in practice. Our simulations show that a policy of continuous vaccination (described in detail below) always yields higher social welfare than pulse vaccination. To not unduly stack the cards against the non-economic approach to vaccination, we will confine our attention to comparing the economic vaccination policy to the continuous non-economic policy.

In our model, vaccination $v(t)$ is a flow variable indicating the instantaneous rate of vaccination. We can model the alternative biological approach in flow terms as follows. Consider the vaccination policy

$$v(t) = \begin{cases} 1 & \text{for } t \leq t^* \\ 0 & \text{for } t > t^* \end{cases} \quad (69)$$

where t^* is implicitly defined by

$$S(t^*) = \frac{\gamma}{\beta} \quad (70)$$

We know that

$$\dot{S}(t) < 0 \text{ for all } t \quad (71)$$

whereas

$$\dot{I}(t) > 0 \text{ for } t < t^* \quad (72)$$

$$\dot{I}(t) = 0 \text{ for } t = t^* \quad (73)$$

$$\dot{I}(t) < 0 \text{ for } t > t^* \quad (74)$$

This policy is the most effective at curbing the spread of the disease, but it takes no account of cost and is therefore not optimal in the economic sense. In the Appendix, we show that this policy achieves the same measure of total vaccinations as the successful pulse vaccination of a fraction \bar{v} of the population.

Turning to our simulation results, we report the discounted social welfare under four different policies, namely (i) the socially optimal one (*), the continuous biological

one (B) as just described, a policy of never vaccinating (N) and a policy of always vaccinating (A). In addition, we indicate the time interval during which the optimal and the biological vaccination policy is active. The simulation results are reported in Table 1.

Comparing discounted social welfare levels, we find that there are often substantial differences. While it is no surprise that welfare is higher under the optimal policy than under the biological one, it turns out that the latter is sometimes inferior even to the policies of never vaccinating or always vaccinating. This reinforces the point that the biological path does not just maximize an ad-hoc objective. In some cases, it is worse than policies that maximise no objective at all.

5.2. Treating to Achieve Population Immunity. To see what the non-economic approach would imply for a policy on treatment, suppose that treatment $\tau \in [0, 1]$ increases the recovery rate to $(\gamma + \alpha_T \tau)$, where $\alpha_T > 0$ is the (finite) efficiency of treatment. In other words, treatment boosts the patient's own immune defence and speeds up recovery. The basic rate of reproduction now changes to

$$\mathbb{R}_0^T \equiv \frac{\beta}{\tau \alpha_T + \gamma} \quad (75)$$

Again, in order to ensure that the disease declines over time, the policy must induce a reproductive ratio $\mathbb{R}_0^T < 1$, which is achieved by permanently treating infected individuals at some level

$$\tau \geq \bar{\tau} \equiv \frac{\beta - \gamma}{\alpha_T} = \frac{\mathbb{R}_0 - 1}{(\alpha_T/\gamma)} \quad (76)$$

Since treatment is not infinitely efficient in inducing recovery, i.e. recovery is not instant, such eradication may not be feasible as treatment may fail to force \mathbb{R}_0^T below one.

For treatment, there is no sensible equivalent to pulse vaccination because there may not be enough infected individuals initially to achieve the desired threshold of treatment-induced recovered people. But it is straightforward to replicate the continuous vaccination policy described above in terms of treatment. The policy takes the form

$$v(t) = \begin{cases} 1 & \text{for } t \leq t^* \\ 0 & \text{for } t > t^* \end{cases} \quad (77)$$

where

$$S(t^*) = \frac{\gamma}{\beta} \quad (78)$$

As for the vaccination simulations, we compare the four different policies of optimal treatment (*), the biologically determined continuous treatment (B) as just described, the policy of never treating (N) and that of always treating (A). The simulation results are reported in Table 1.

Again, we find that there are sometimes very substantial welfare losses associated with the biological approach. In addition, the optimal treatment policy always ceases treatment before the biological one, emphasizing that the objective is to maximize social welfare rather than reaching a particular threshold of immune individuals. Again, we also see that there are cases in which the biological policy is outperformed by the

policies of never treating or always treating, respectively.

6. DISEASE-INDUCED MORTALITY

In the previous sections, we have assumed that while infection decreases the wellbeing of an infected individual, it does not lead to death. In this section, we consider the alternative case in which infected individuals may die from the disease.

First, introduce a new class of deceased individuals, $\mathcal{D}(t)$, of measure $D(t)$. Next, assume that while infected, individuals transition to the deceased class at rate $\delta \in [0, \infty)$. The uncontrolled system is now given by

$$\dot{S}(t) = -\beta I(t)S(t) \quad (79)$$

$$\dot{I}(t) = I(t) [\beta S(t) - \gamma - \delta] \quad (80)$$

$$\dot{R}(t) = \gamma I(t) \quad (81)$$

$$\dot{D}(t) = \delta I(t) \quad (82)$$

$$\dot{N}(t) = -\delta I(t) \quad (83)$$

There are several properties of this extended model that merit comment. First, note that the population decreases in size over time (as the deceased individuals are for simplicity not replaced by new births). Second, the qualitative properties of the dynamics are unchanged. To wit, the basic rate of reproduction is now reduced to $\mathbb{R}_0 = \beta/(\gamma + \delta)$, with the limiting distribution over classes suitably modified. It is still the case that $\lim_{t \rightarrow \infty} S(t) > 0$, although this quantity must now be interpreted as a measure rather than as a fraction of the population. The higher the mortality rate δ is, the fewer susceptible and recovered individuals remain in the limiting distribution.¹⁸

We next consider the economic control of this model and assume that the payoff to a deceased individual is $\pi_{\mathcal{D}} = 0$. As argued by Gersovitz and Hammer (2004), this assumption means that the shortfall in welfare for an individual in the deceased class is the opportunity cost of not earning $\pi_{\mathcal{I}}$ until recovery from the disease and not earning $\pi_{\mathcal{R}}$ thereafter.

Consider first the problem of an individual under decentralized decision making. For such an individual, mortality unambiguously increases the cost of infection, thereby increasing the incentive to vaccinate and treat. Qualitatively, the decentralized problem remains unchanged. Turning to the centralized problem, mortality has somewhat more subtle effects. On the one hand, society's welfare is decreased because the specific individual passes away, incurring an opportunity cost like that of the individual. On the other hand, there are additional and countervailing effects of mortality, due to external effects. When an individual passes away, society loses a current infected individual who no longer infects others. In determining the overall impact of mortality on the socially optimal policy on vaccination, one must weigh these different considerations. To the extent that the overall effect on welfare is negative, the qualitative effects on the optimal vaccination policy are unchanged. In turn, the possibility of

¹⁸Note that this is an 'early death' model, in which mortality shortens the time spent in the infectious state. This modeling choice is also made in Gersovitz and Hammer (2004), Getz and Lloyd-Smith (2005) and Keeling and Rohani (2008). Alternative 'late death' models, in which recovery time and time till death are equal, are treated by Keeling and Rohani (2008), Naevdal (2012) and Gallos and Fefferman (2015). These models are less suitable for the present purposes.

disease-induced mortality (weakly) increases the value of treatment.

7. CONCLUSION

In this paper, we have considered the economic control of the classical susceptible-infected-recovered model of infectious disease. Two costly measures are considered, namely treatment and vaccination. Treatment increases the rate of recovery and confers immunity from future infection on the recovered individual. We find that in equilibrium, if the treatment cost is sufficiently large, individuals adopt socially sub-optimal treatment policies, leading to too little treatment and recovery. This is because decentralized and non-cooperative individuals disregard the socially beneficial external effects that treatment and recovery have on susceptible individuals (through the effects of herd immunity). Optimal treatment will (weakly) decrease over time with at most one switch from full treatment to no treatment.

Vaccination decisions turn out to be different to those of treatment, whether made by individuals or by a social planner. Individuals will delay vaccination until the hazard of infection becomes sufficiently high. While this is also true of the socially optimal vaccination policy, the vaccination threshold of the planner is typically lower than the equilibrium threshold. Furthermore, depending on parameter assumptions, the equilibrium and the socially optimal vaccination policy may involve more than once switch between no vaccination and full vaccination.

The optimal policy will depend delicately on a combination of biomedical and economic (preference) parameters. But the overall nature of the external effects is this. External effects from treatment are proportional to the measure of remaining susceptibles, while those from vaccination are proportional to disease prevalence. While the former is monotone, the latter is non-monotone and therefore outcomes under vaccination, in equilibrium as well as under social planning, may involve as many as two switches.

Looking at the comparisons from a different angle, with vaccination there are cases in which both the planner and individuals have qualitatively similar behavior, but where different valuations lead to quantitative differences in outcomes. In contrast, with treatment, equilibrium behavior differs qualitatively from the socially optimal policy. The individual's decision problem is a simple and state independent comparison of costs and benefits, while that of the planner is a delicate function of the measures of susceptible, infected and recovered individuals.

Under decentralized decision making no self-interested individual would consciously contribute towards herd immunity. Having said that, the presence of herd immunity may well influence individual decision making, as is the case with vaccination when the population is heterogeneous or when the vaccine confers imperfect protection against infection.

A topic of both theoretical and practical importance, but which this paper has not treated, is the equilibrium and the socially optimal combination of treatment and vaccination. It is immediately clear that treatment and vaccination at different stages of the epidemic changes the value of such efforts. The simplest case to consider is that of decentralized decision making. For an individual, treatment is an imperfect substitute for vaccination, while vaccination is no substitute for treatment (once infected, it is simply too late to vaccinate). Thus for an infected individual, the presence or otherwise of vaccination is immaterial and cannot influence the treatment decision.

Next, consider the effects that the presence of treatment has on the privately optimal use of vaccination. Whether treatment is available or not, the individual will choose to vaccinate once the cost of doing so is justified by the expected future cost of infection. Now, the option to treat infection in the future must decrease the expected future cost of becoming infected - at least weakly so since the option to treat can always be rejected. This means that if the individual would choose not to vaccinate without the option of treatment, then it would continue not to do so with that option present. On the other hand, an individual who would choose to vaccinate without the possibility of treatment may or may not choose to do so once treatment becomes an option. It depends on how much the treatment option lowers the infection cost, relative to the cost of vaccination. This completes the characterization of the hybrid model under decentralized decision making.

Next, turn to centralized decision making. Perhaps not surprisingly, this case poses a significant challenge that will not be taken on here. When studied in isolation, optimal policy suggests using treatment at early stages, while delaying vaccination to later stages. Once treatment and vaccination can be combined, it is by no means clear that their timing would reflect that pattern. Depending on costs, parameters and initial conditions, it is conceivable that an optimal policy would involve so much initial treatment that no subsequent vaccination would be desirable.

In this connection it is worth noting that while an individual is always restricted to vaccination first and then treatment later (if infected), the central planner would have no such restrictions at the population level. Of course, the same restriction would apply in vaccinating or treating any given individual, but at the aggregate level, the planner can make any number of switches between combinations of vaccination and treatment across the stages of the epidemic.

There are some overall lessons to be learned from the present analysis. First, rather than focusing on the mechanics of population immunity and on bringing down the basic rate of reproduction, optimal treatment and vaccination policy should be formulated with a view to optimally trade off costs and benefits of each intervention. Once this has been achieved, it emerges endogenously whether and the extent to which herd immunity plays a role in the optimal policy. Second, it is not always optimal to suppress infection by lowering the basic rate of reproduction. This follows from the simple observation that under certain circumstances, the costs of suppressing infection (by lowering the basic rate of reproduction) outweigh the social benefits of doing so. Third, the intensity of optimal interventions may well vary across different stages of the epidemic in non-obvious ways. In particular, depending on the available tools, optimal intervention may be either front-loaded or back-loaded. This finding contrasts with the common assertion in the public health literature that interventions such as vaccination should be made earlier rather than later in order to maximize impact.

It is worth emphasizing the importance of properly formulated policy in the management of epidemics through pharmaceutical interventions (once and if they become available). This is relevant not only for the COVID-19 epidemic but also for other diseases that fit the present framework such as H1N1, seasonal influenza, hepatitis B, norovirus, pertussis (whooping cough) and a host of other viral and bacterial infections.¹⁹ The tools available for a health authority depends on the specific infectious

¹⁹While for some of these diseases acquired immunity is not permanent, the model is still appro-

disease at hand. Unfortunately, there are many vaccine-preventable diseases for which there are no effective treatments available. Similarly, there are many treatable diseases for which there are no viable vaccines. This paper has taken a first step towards characterizing equilibrium behavior and optimal policy to control population immunity when either of these instruments is available.

A. SIMULATION RESULTS

For the simulations, we used a fourth order Runge-Kutta method. The notation for the parameters $(\gamma, \rho, \pi_S, \pi_I, \pi_R, I_0, S_0, R_0, \alpha_V, \alpha_T, c_V, c_T)$ are set out in the text. \bar{T} denotes the time horizon used in the simulations. The values W_k^* and W_k^B , $k = V, T$ denote discounted social welfare under the optimal policy (*) and the biological policy (B), respectively. The values W_k^N and W_k^A , $k = V, T$ denote the discounted social welfare under the policies of never (N) or always (A) using the instrument k . Last, Δ_k^* and Δ_k^B , $k = V, T$ denote the time interval during which the relevant intervention is being applied. The values for ρ, β, γ are chosen to be in line with those for COVID-19 for the vaccination simulations (see e.g. Giannitsarou et al., 2020), but the remaining parameters were chosen for purposes of illustration of specific points. Since we are interested in exploring the structural properties of socially optimal treatment and vaccination across a range of infectious diseases, we have chosen parameters to showcase the most interesting cases, namely those where there are most shifts in the optimal policies. For the treatment simulations, we have chosen parameters β, γ that yield a somewhat higher rate of reproduction than for COVID-19, but quantitatively similar results can be obtained by setting a lower rate of reproduction but changing other parameters, e.g. by increasing the cost of treatment. In order to do a careful calibration to COVID-19, we would need additional information not yet available, namely the biomedical properties of treatment and vaccine efficiency and their associated costs.

B. CONTINUOUS NON-ECONOMIC VACCINATION

Since $v(t) = 1$ for $t \leq t^*$, we can write the evolution of susceptibles as

$$\dot{S}(t) = -S(t) [\beta I(t) + \alpha_V] \text{ for } t \leq t^* \quad (84)$$

For small values of $I(t)$, this equation can be approximated as

$$\dot{S}(t) = -\alpha_V S(t) \quad (85)$$

which has the solution

$$S(t) = S_0 e^{-\alpha_V t} \quad (86)$$

Suppose that $R_0 = 0$ and that I_0 is vanishingly small. Then $S_0 \approx 1$ and hence

$$S(t) \approx e^{-\alpha_V t} \quad (87)$$

Since $S(t^*) = \gamma/\beta$, it follows that

$$\frac{\gamma}{\beta} = e^{-\alpha_V t^*} \quad (88)$$

Thus we find the switching time as

$$t^* = -\frac{1}{\alpha_V} \ln \left(\frac{\gamma}{\beta} \right) \quad (89)$$

This is an accurate approximation of the dynamics for small values of I_0 . Next,

Simulated Vaccination Scenarios (boldfaced numbers indicate deviations from benchmark case)																			
Case	\bar{T}	ρ	I_0	S_0	R_0	β	γ	α_T	c_T	π_S	π_T	π_R	W_T^*	W_T^B	W_T^N	W_T^A	Δ_T^*	Δ_T^B	
Simulated Treatment Scenarios (boldfaced numbers indicate deviations from benchmark case)																			
Case	\bar{T}	ρ	I_0	S_0	R_0	β	γ	α_T	c_T	π_S	π_T	π_R	W_T^*	W_T^B	W_T^N	W_T^A	Δ_T^*	Δ_T^B	
Base	40	0.05	10^{-8}	$1-10^{-8}$	0	2	0.2	0.3	10	1	-5	1	4.03	2.65	2.90	1.84	0-10.51	0-14.34	
Low α_T	40	0.05	10^{-8}	$1-10^{-8}$	0	2	0.2	0.15	10	1	-5	1	3.37	-0.59	2.92	-5.52	0-8.58	0-12.77	
High I_0	40	0.05	0.05	0.944	0.006	2	0.2	0.3	10	1	-5	1	-4.67	-7.02	-4.67	-8.41	0-0.16	0-3.94	
High β	40	0.05	10^{-8}	$1-10^{-8}$	0	8	0.2	0.3	10	1	-5	1	-3.96	-5.08	-4.03	-8.44	0-1.91	0-3.00	
High γ	40	0.05	10^{-8}	$1-10^{-8}$	0	2	0.4	0.3	10	1	-5	1	10.53	8.51	9.85	7.44	0-11.14	0-15.75	
High c_T	40	0.05	10^{-8}	$1-10^{-8}$	0	2	0.2	0.3	100	1	-5	1	3.76	-54.46	2.90	-85.00	0-8.07	0-14.34	
High \bar{T}	40	0.05	10^{-8}	$1-10^{-8}$	0	2	0.2	0.3	10	1	-5	1	6.59	5.21	5.47	4.41	0-10.51	0-14.34	
High ρ	40	0.10	10^{-8}	$1-10^{-8}$	0	2	0.2	0.3	10	1	-5	1	3.68	2.65	2.60	2.09	0-10.56	0-14.34	
Low π_T	40	0.05	10^{-8}	$1-10^{-8}$	0	2	0.2	0.3	10	1	-8	1	-1.06	-1.61	-4.29	-1.06	0-32.82	0-14.34	
Low π_R	40	0.05	10^{-8}	$1-10^{-8}$	0	2	0.2	0.3	10	1	-5	0	-2.10	-4.01	-3.99	-5.17	0-10.63	0-14.34	
High π_S	40	0.05	10^{-8}	$1-10^{-8}$	0	2	0.2	0.3	10	2	-5	1	13.06	11.87	10.90	11.16	0-10.97	0-14.34	
Simulated Vaccination Scenarios (boldfaced numbers indicate deviations from benchmark case)																			
Case	\bar{T}	ρ	I_0	S_0	R_0	β	γ	α_V	c_V	π_S	π_T	π_R	W_V^*	W_V^B	W_V^N	W_V^A	Δ_V^*	Δ_V^B	
Base	40	0.05	10^{-8}	$1-10^{-8}$	0	1	0.2	0.95	0.1	1	-5	1	17.24	17.21	9.85	17.19	7.87-9.10	0-1.69	
Low α_V	40	0.05	10^{-8}	$1-10^{-8}$	0	1	0.2	0.2	0.1	1	-5	1	17.08	16.95	9.85	16.89	8.10-13.94	0-8.04	
High I_0	40	0.05	0.05	0.937	0.013	1	0.2	0.95	0.1	1	-5	1	14.61	12.88	-2.27	14.61	0-23.20	0-1.51	
High β	40	0.05	10^{-8}	$1-10^{-8}$	0	4	0.2	0.95	0.1	1	-5	1	17.20	17.20	-1.54	17.19	0.80-3.70	0-3.15	
High γ	40	0.05	10^{-8}	$1-10^{-8}$	0	1	0.4	0.95	0.1	1	-5	1	17.26	17.23	14.73	17.19	11.01-11.69	0-0.96	
High c_V	40	0.05	10^{-8}	$1-10^{-8}$	0	1	0.2	0.95	10	1	-5	1	12.70	9.12	9.85	7.29	13.31-14.31	0-1.69	
High \bar{T}	40	0.05	10^{-8}	$1-10^{-8}$	0	1	0.2	0.95	0.1	1	-5	1	19.81	19.78	12.29	19.77	9.82-11.50	0-1.69	
High ρ	40	0.10	10^{-8}	$1-10^{-8}$	0	1	0.2	0.95	0.1	1	-5	1	12.94	9.74	7.79	9.71	11.40-12.94	0-1.69	
Low π_T	40	0.05	10^{-8}	$1-10^{-8}$	0	1	0.2	0.95	0.1	1	-8	1	17.24	17.21	9.85	17.19	6.34-7.48	0-1.69	
Low π_R	40	0.05	10^{-8}	$1-10^{-8}$	0	1	0.2	0.95	0.1	1	-5	0	12.45	3.85	7.46	0.90	15.65-16.83	0-1.69	
High π_S	40	0.05	10^{-8}	$1-10^{-8}$	0	1	0.2	0.95	0.1	2	-5	1	29.65	21.14	23.52	18.19	14.84-15.88	0-1.69	

the ratio of vaccinations to the initial population of susceptibles is in general equal to

$$\tilde{v} = \frac{\int_0^\infty v(t)S(t)dt}{S_0} \quad (90)$$

In the case we are considering, we get

$$\tilde{v} = \frac{\int_0^{t^*} S_0 e^{-\alpha_V t} dt}{S_0} \quad (91)$$

$$= \int_0^{t^*} e^{-\alpha_V t} dt \quad (92)$$

$$= \frac{1}{\alpha_V} (1 - e^{-\alpha_V t^*}) \quad (93)$$

$$= \frac{1}{a_V} \left(1 - \frac{\gamma}{\beta}\right) \quad (94)$$

$$= \frac{\mathbb{R}_0 - 1}{\mathbb{R}_0 \alpha_V} = \bar{v} \quad (95)$$

Thus with the suggested version of continuous vaccination, the non-economic approach induces the same total number of vaccinations as under the pulse vaccination of the critical threshold, but it achieves this by spreading vaccinations out over time rather than doing it instantaneously at the outset.

Incidentally, the continuous vaccination method allows for vaccinating the same person more than once (if missed or failed in initial rounds) and is therefore compatible with $\tilde{v} > 1$. The same could be true in the pulse vaccination case, if we were to allow for instantaneous repeat vaccinations.

REFERENCES

- [1] ANDERSON, R. M. AND R. M. MAY (1991): Infectious Diseases of Humans: Dynamics and Control, *Oxford University Press*.
- [2] ARINO, J., F. BRAUER, P. VAN DEN DRIESSCHE, J. WATMOUGH AND J. WUE (2008): A Model for Influenza with Vaccination and Antiviral Treatment, *Journal of Theoretical Biology*, 253(1), 118-130.
- [3] AULD, C. (2003): Choices, Beliefs, and Infectious Disease Dynamics, *Journal of Health Economics*, 22(3), 361-77.
- [4] BAKARE, E. A. (2015): On the Optimal Control of Vaccination and Treatments for an SIR Epidemic Model with Infected Immigrants, *Journal of Applied & Computational Mathematics*, 4(4), doi:10.4172/2168-9679.1000230.
- [5] BAKARE, A. E., A. NWAGWO AND E. DANSO-ADDO (2014): Optimal Control Analysis of an SIR Epidemic Model with Constant Recruitment, *International Journal of Applied Mathematical Research*, 3(3), 273-285.
- [6] BARRETT, S. (2003): Global Disease Eradication, *Journal of the European Economic Association*, 1(2-3), 591-600.
- [7] BARRETT, S. AND M. HOEL (2007): Optimal Disease Eradication, *Environment and Development Economics*, 12(5), 627-652.
- [8] BAUCH, C. T. (2005): Imitation Dynamics Predict Vaccinating Behaviour, *Proceedings of the Royal Society B*, 272, 1669-1675.
- [9] BAUCH, C. T. AND D. J. D. EARN (2004): Vaccination and the Theory of Games, *Proceedings of the National Academy of Sciences*, 101(36), 13391-13394.
- [10] BRAUER, F. AND C. CASTILLO-CHAVEZ (2012): Mathematical Models in Population Biology and Epidemiology, 2nd edition, *Springer*.
- [11] BOULIER, B. L., T. S. DATTA AND R. S. GOLDFARB (2007): Vaccination Externalities, *B.E. Journal of Economic Analysis & Policy*, 7(1), (Contributions), Article 23.
- [12] BRITO, D. L., E. SHESHINSKI AND M. D. INTRILIGATOR (1991): Externalities and Compulsory Vaccinations, *Journal of Public Economics*, 45(1), 69-90.
- [13] CHEN, F. AND TOXVAERD, F. (2014): The Economics of Vaccination, *Journal of Theoretical Biology*, 363, 105-117.
- [14] ELY, J., A. GALEOTTI AND J. STEINER (2020): Rotation as Contagion Mitigation, *CEPR Discussion Paper No. DP14953*.
- [15] FERGUSON, N. M., D. A. T. CUMMINGS, C. FRASER, J. C. CAJKA, P. C. COOLEY AND D. S. BURKE (2006): Strategies for Mitigating an Influenza Pandemic, *Nature*, 442, 448-452.

- [16] FINE, P. E. M. (1993): Herd Immunity: History, Theory, Practice, *Epidemiologic Reviews*, 15(2), 265-302.
- [17] FINE, P., K. EAMES AND D. L. HEYMANN (2011): “Herd Immunity”: A Rough Primer, *Clinical Infectious Diseases*, 52(7), 911-916.
- [18] FRANCIS, P. J. (1997): Dynamic Epidemiology and the Market for Vaccinations, *Journal of Public Economics*, 63(3), 383-406.
- [19] FRANCIS, P. J. (2004): Optimal Tax/Subsidy Combinations for the Flu Season, *Journal of Economic Dynamics and Control*, 28(10), 2037-2054.
- [20] FRANCIS, P. J. (2007): Dynamic Vaccination Externalities with Heterogeneity, Recovery and Discounting, *mimeo*.
- [21] GALLOS, L. K. AND N. H. FEFFERMAN (2015): The Effect of Disease-Induced Mortality on Structural Network Properties, *PLoS ONE*, 10(8): e0136704. doi:10.1371/journal.pone.0136704.
- [22] GERSOVITZ, M. (2003): Births, Recoveries, Vaccinations and Externalities, in *Economics for an Imperfect World: Essays in Honor of Joseph E. Stiglitz*, R. J. Arnott, et al. (eds.), MIT Press, 469-483.
- [23] GERSOVITZ, M. AND J. S. HAMMER (2004): The Economical Control of Infectious Diseases, *Economic Journal*, 114(492), 1-27.
- [24] GERMANN, T. C., K. KADAU, I. M. LONGINI, JR. AND C. A. MACKEN (2006): Mitigation Strategies for Pandemic Influenza in the United States, *Proceedings of the National Academy of Sciences*, 103(15), 5935-5940.
- [25] GETZ, W. M. AND J. O. LLOYD-SMITH (2005): Basic Methods for Modeling the Invasion and Spread of Contagious Diseases, *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*.
- [26] GIANNITSAROU, C., S. KISSLER AND F. TOXVAERD (2020): Waning Immunity and the Second Wave: Some Projections for SARS-CoV-2, *CEPR Discussion Paper DP14852*.
- [27] GOLDMAN, S.M. AND J. LIGHTWOOD (1995): The SIS Model of Infectious Disease with Treatment, *mimeo*.
- [28] GOLDMAN, S.M. AND J. LIGHTWOOD (2002): Cost Optimization in the SIS Model of Infectious Disease with Treatment, *Topics in Economic Analysis and Policy*, 2(1), 1-22.
- [29] HETHCOTE, H. W. AND P. WALTMAN (1973): Optimal Vaccination Schedules in a Deterministic Epidemic Model, *Mathematical Biosciences*, 18(3-4), 365-381.
- [30] JOSHI, H. R., S. LENHART, S. HOTA AND F. AGUSTO (2015): Optimal Control of an SIR Model with Changing Behavior Through an Education Campaign, *Electronic Journal of Differential Equations*, 2015(50), 1-14.

- [31] KEELING, M. J. AND P. ROHANI (2008): Modeling Infectious Diseases in Humans and Animals, *Princeton University Press*.
- [32] LEDZEWICZ, U. AND H. SCHATTLER (2011): On Optimal Singular Controls for a General SIR-Model with Vaccination and Treatment, *Discrete and Continuous Dynamical Systems*, 2011, Special Issue, 981-990.
- [33] LONGINI, JR., I. M., M. E. HALLORAN, A. NIZAM AND Y. YANG (2004): Containing Pandemic Influenza with Antiviral Agents, *American Journal of Epidemiology*, 159(7), 623-33.
- [34] MORTON, R. AND K. H. WICKWIRE (1974): On the Optimal Control of a Deterministic Epidemic, *Advances in Applied Probability*, 6(4), 622-635.
- [35] MILNE, G. J., N. HALDER AND J. K. KELSO (2013): The Cost Effectiveness of Pandemic Influenza Interventions: A Pandemic Severity Based Analysis, *PLoS One*, 8(4):e61504.
- [36] NAEVDAL, E. (2012): Fighting Transient Epidemics-Optimal Vaccination Schedules Before and After an Outbreak, *Health Economics*, 21(12), 1456-1476.
- [37] RELUGA, T. C. AND A. P. GALVANI (2011): A General Approach for Population Games with Application to Vaccination, *Mathematical Biosciences*, 230(2), 67-78.
- [38] ROWTHORN, R. (2006): The Optimal Treatment of Disease Under a Budget Constraint, in R. Halvorsen and D. Layton (eds), *Explorations in Environmental and Natural Resource Economics: Essays in Honor of Gardner M. Brown, Jr*, Edward Elgar.
- [39] ROWTHORN, R. AND F. TOXVAERD (2020): The Optimal Control of Infectious Diseases via Prevention and Treatment, *Cambridge-INET Working Paper Series* No: 2020/13.
- [40] SANDERS, J. L. (1971): Quantitative Guidelines for Communicable Disease Control Programs, *Biometrics*, 27(4), 883-893.
- [41] SETHI, S. P. (1974): Quantitative Guidelines for Communicable Disease Control Program: A Complete Synthesis, *Biometrics*, 30(4), 681-691.
- [42] SETHI, S. P., P. W. STAATS (1978): Optimal Control of Some Simple Deterministic Epidemic Models, *Journal of the Operational Research Society*, 29(2), 129-136.
- [43] SMITH, C. E. G. (1964): Factors in the Transmission of Virus Infections from Animals to Man, *The Scientific Basis of Medicine Annuals Reviews 1964*, 125-150.
- [44] WICKWIRE, K. H. (1975): A Note on the Optimal Control of Carrier-Borne Epidemics, *Journal of Applied Probability*, 12(3), 565-568.
- [45] TOXVAERD, F. (2009): Recurrent Infection and Externalities in Treatment, *mimeo*.

- [46] YUSUF, T. T. AND F. BENYAH (2012): Optimal Control of Vaccination for an SIR Epidemiological Model, *World Journal of Modelling and Simulation*, 8(3), 194-204.