The Economics of Infectious Diseases[1]

M. Christopher Auld[2] Eli P. Fenichel[3] Flavio Toxvaerd[4]

This version: November 6, 2024.

[1]We thank Steven Durlauf, David Romer, and six anonymous referees for many very helpful comments.

[2]Department of Economics, University of Victoria. Email: auld@uvic.ca.

[3]Yale School of the Environment. Email: eli.fenichel@yale.edu.

[4]Faculty of Economics, University of Cambridge, Bennett Institute for Public Policy and CEPR. Email: fmot2@cam.ac.uk.

Abstract

We synthesize the literature on economic epidemiology, the interdisciplinary field that draws on the ideas and methods of economics to analyze individual behavior, aggregate disease dynamics, and public policy during infectious disease epidemics. We cover the main models of individual behavior during epidemics, related econometric evidence, and models of disease dynamics appropriate for the analysis of a range of infectious diseases. We outline modeling approaches to a range of control measures including non-pharmaceutical interventions such as stay–at–home mandates, quarantines, and sheltering, and pharmaceutical interventions such as vaccines and treatment. Last, we characterize different types of externalities and heterogeneities and discuss the targeting and implementation of policies through restrictions and incentives.

JEL Classification: H12, C73, I12, I18.

Keywords: Economic epidemiology, disease dynamics, social distancing, non-pharmaceutical interventions, vaccination, treatment, econometrics.

1 Introduction

Epidemics, from the Antonine Plague in ancient Rome to the Bubonic Plague to the COVID-19 pandemic, have shaped the course of human civilization. Despite the fact that many decisions and tradeoffs that drive epidemics take place outside of the market, epidemics are an economic problem: The threat of communicable disease generates incentives to avoid becoming infected, and the resulting changes in behavior determine the course of the epidemic. During the COVID-19 epidemic, for example, choices such as whether to mask, socially distance, or receive a vaccination were driven by perceived risk of infection and, in turn, governed the course of the epidemic, including health outcomes, microeconomic, and macroeconomic outcomes. This simultaneity of epidemic outcomes and behavior is the central issue of economic epidemiology, the interdisciplinary field that draws on the ideas and methods of economics to analyze individual behavior, aggregate disease dynamics, and public policy during infectious disease epidemics.[1]

Emerging in the 1980s and 1990s and motivated largely by the human toll taken by the HIV+/AIDS epidemic, economic epidemiology has subsequently focused on malaria, various strains of influenza, ebola, Lyme disease, and a variety of livestock and wildlife diseases. The COVID-19 pandemic attracted economists to the field in large numbers, greatly expanding theoretical and empirical investigation of the interactions among health-affecting behaviors, disease dynamics, and microeconomic and macroeconomic dynamics. In this paper, we synthesize this literature by tracing the long arc of its development, highlighting key ideas from pioneering work in the 1970s, through rapid expansion in response to the HIV+/AIDS epidemic, and culminating with the greatly expanded interest among economists generated by the COVID-19 epidemic.

Economic epidemiology starts with an appreciation of, and willingness to incorporate, the rich background dynamics provided by biological epidemiology, highlighting biological and medical processes and constraints. For example, some pathogens transmit from person to person and some through the environment. Some pathogens convey immunity, while others do not, or the immunity quickly wanes. Some pathogens spread through sexual contacts, whereas other pathogens pass through the air. Understanding and including these kinds of detail in analysis is important for two reasons. First, these features result in different intrinsic dynamics that can pose challenges for economic analysis. As we show in this review, many of these features influence individual incentives and the tradeoffs faced by policy makers. Second, economists interested in having their research influence epidemiological policy must be able to communicate with medical doctors and epidemiologists, who often lead technical public health advice.

Economic epidemiology builds on these foundations. At the microeconomic level, individuals and households make health–related decisions such as whether to voluntarily socially distance, whether to get vaccinated, or whether to engage in risky sexual encounters. Individuals also face trade-offs, similar to other trade-offs familiar to economists, between health and other goals such as consumption or sexual variety. Since such responses in

[1] We continue the convention of using the terminology *economic epidemiology* to refer specifically to economic analysis of communicable disease and not economic analysis of other topics that overlap with epidemiology, such as the economics of obesity or substance use. See Philipson (2000), Funk, Salathé and Jansen (2010), Fenichel et al. (2011), Perrings et al. (2014), Avery et al. (2020), McAdams (2021), Dangerfield et al. (2022), and Bloom, Kuhn and Prettner (2022) for other reviews of aspects of the economic epidemiology literature.

turn affect the spread of disease, we show how biological models of disease spread have been modified to incorporate rational behavioral responses. Understanding how health and infection risk enter into decision making is foundational to forecasting epidemic dynamics, evaluating policy interventions, and designing welfare–improving public health policy. Indeed, in Ferguson (2007), a leading epidemiologist explicitly called for the inclusion of behavior in epidemic models to guide policy. In the wake of the 2020 COVID-19 pandemic, it has become strikingly clear that epidemiological modeling suffers from failure to include endogenous behavioral responses to the threat of disease and to policy interventions.

The importance of microeconomic behavior has elevated the importance of econometrics in the study of epidemics. Even biological epidemiological models depend on behavioral parameters that must be estimated from observational data; economic epidemiological variants emphasize the importance of shifts in behavior that must be measured absent randomized trials, and similarly randomized experiments to assess the effects of policy interventions are typically not available. For example, to our knowledge, during the COVID-19 epidemic only one randomized controlled trial was conducted, globally, on purely non-pharmaceutical interventions (Abaluck et al., 2022), and one randomized trial was conducted on improving vaccine uptake (Meriggi et al., 2024). Since such interventions often have general equilibrium effects, it is particularly important to combine theoretical modeling and econometric evidence to measure their effects. This area of research grew rapidly during the COVID-19 pandemic.

How individual behavior aggregates to generate social outcomes is central to economic theory. Economic epidemiology provides the microfoundations for aggregate responses to disease risk and disease dynamics and helps characterize equilibria in disease settings. Most traditional epidemiological modeling starts from specifying population-level dynamics rather than individual behavioral foundations. This state of affairs is not dissimilar to macroeconomics prior to the neoclassical synthesis, and much of the economic epidemiology literature can be interpreted as an attempt to provide microfoundations for models of aggregate disease dynamics. In canonical microeconomics, agents take prices as parametric and make choices that collectively determine, and are determined by, prices. In economic epidemiology, agents take risk of disease as parametric and make choices that collectively determine, and are determined by, risk of disease. This analogy to conventional economic theory is not exact, but highlights why economic theory developed to study behavior in markets, including frameworks such as rational choice and supply and demand, and modeling conventions such as positing a social planner, are so easily adaptable to epidemiological problems.

Normative economic theory adds important insights to policy analysis during epidemics. Once models are microfounded in ways that enable reasonable forecasts of epidemics, it is possible to consider policy interventions that improve welfare. The focus of economic epidemiology on changes in welfare is a departure from biological epidemiology and the medical literature more broadly, which typically implicitly assumes that public health, as measured for example by disease-induced deaths, is the only measure of well-being of interest. Economic epidemiology emphasizes that the total welfare loss from infectious disease includes loss of outcomes such as social contacts, sexual partners, and income in addition to diminished health. In so doing, it highlights the excess burden of protective measures to reduce infection in addition to direct health impacts (Philipson, 1995). Further, many decisions that generate infection risk are not market transactions, creating the potential for a range of externalities, which we characterize within this review.

Policy analysis proceeds from studying these external effects. Economics may improve understanding of whether and when equilibrium dynamics are welfare maximizing and whether government intervention is called for. For example, drawing on welfare economics, economists can design incentive schemes such as Pigouvian taxes or subsidies to address problems arising from contagion and other externalities. Economics also provides the tools for analyzing possible second-best policies (Lipsey and Lancaster, 1956), which do not necessarily achieve socially optimal outcomes, but can substantially improve welfare in the presence of pre-existing distortions.

In this survey, we present an overview of the development of the key concepts in the rapidly expanding field of economic epidemiology. Most of the review is divided into two broad sections: positive and normative analysis. Within the section on positive analysis, we review compartmental models used in classical epidemiology, such as the canonical susceptible–infected–recovered (SIR) model. We present the economic theory of rational behavior during an epidemic, outlining the standard model of response to risk of disease, and show how it integrates with those canonical epidemiological models. Then, we turn to the empirical evidence regarding this theory. In the normative section of the paper, we clarify the role of economics in providing a formal welfare-based approach to assessing policy interventions within epidemics. Next, we develop a taxonomy of epidemiological externalities, which we use to explain the variety of results in the normative economic epidemiology literature. We connect these results to different epidemiological structures from the susceptible-exposed-infectious-recovered-susceptible (SEIRS) family of models, and discuss other policy considerations and topics which have so far received scant attention. Having summarized the literature to date, we conclude with a brief discussion of promising future directions for economic epidemiology.

2 Positive Economic Epidemiology

In this section we describe the positive science of modeling epidemics. First, we introduce the canonical models of disease dynamics that form the basis for most epidemiological and public health research and on which epidemic containment policies are often based. These are *compartmental models*, described by systems of differential equations. As will become clear, these models are wholly abehavioral in the sense that the parameters of the models jointly capture and confound biomedical features of the diseases and features that are functions of individual decisions and policy interventions. Next, we develop the microeconomic foundations for disease dynamics, describe how decision making in a disease context by a single individual can be modeled, and discuss a range of heterogeneities that can influence individual choices. We then connect this structure back to the macroscale dynamics of the epidemic. Last, we review the empirical literature on the economic approach to behavior and disease dynamics.

2.1 Compartmental Models of Disease Dynamics

Compartmental models describe flows of people between various disease states. A "compartment" is a health or other classification within which all individuals are identical. We outline conventional compartmental models by describing a moderately complex variant,

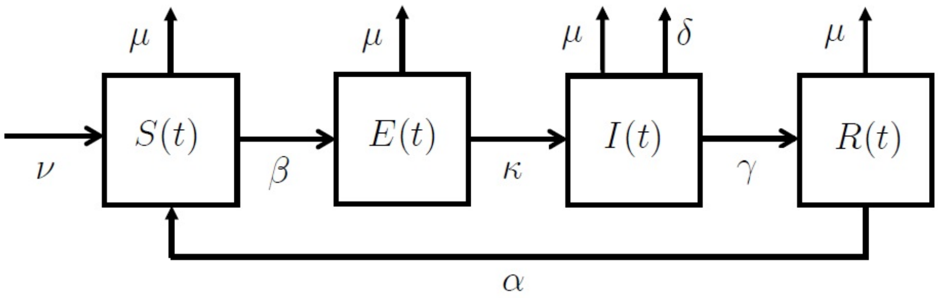

Figure 1: Dynamics in a SEIRS compartment model.

the susceptible-exposed-infectious-recovered-susceptible (SEIRS) model with vital statistics, which nests many important cases.[2] In this model, vital statistics account for new individuals who flow into the population through either births or immigration, and people who flow out of the population through either death or emigration. In the compartmental modeling framework, notation is commonly abused, such that S describes the health class and measure or the population in that class, with the same abuse of notation extended to the other health classes (Figure 1). In addition to health classes, it is common to track the overall population N and cumulative deaths D.

The foundations of the model are as follows. There exists a continuum of individuals, so there is no aggregate uncertainty. Time is continuous. Each individual at any time t is within a given compartment, and by definition there is no heterogeneity within compartments, so, for example, two people within the infected compartment are interchangeable regardless of how long they have been infected. Transitions from one compartment into another follow a Poisson-like process at rates determined by the parameters and functional form of the model.

We now describe these compartments and transition rates for this SEIRS model. All individuals are initially in the susceptible compartment, S. In the economic models discussed below, the behavior of susceptible individuals is typically the focus of analysis, as this behavior is crucial to understanding the rate of new infections. In the SEIRS model, all (living) individuals, regardless of their infection status, contact each other randomly at an exogenously given rate. Interaction between susceptible individuals and infectious individuals generates infections, and those susceptible individuals transition into an exposed class, E. Class E introduces a delay between contracting a pathogen and becoming infectious, i.e. transitioning to class I.[3] In the canonical model the average delay is $1/\kappa$, that is, κ measures the conversion rate from exposed to infectious. Only infectious people experience adverse effects of infection, usually modeled as disease-induced mortality at rate δ. In these models, R represents a class that is recovered with immunity, where recovery occurs at rate γ. In the SEIRS framework, immunity wanes at a rate of α, and individuals return to the S class. SEIRS models may include births (or immigration) at rate v and background mortality (or emigration) at rate μ. However, it is also common to assume that births and immigration perfectly balance all deaths and emigration to hold

[2] See e.g. Hethcote (2000) for a review of biological compartmental models.

[3] In models with compartments in which infected people cannot transmit the disease (e.g., an exposed class with zero transmission), there is a distinction between being infected and being infectious. For the remainder of this paper, we refer to "infected" individuals as those who are both infected and infectious.

the total population, N, constant. We discuss the implications of this assumption further below. When N is assumed constant, it is common for the compartments to be modeled in shares of the total population as opposed to absolute numbers. Finally, the equations for D and N track cumulative deaths and the size of the population, respectively.

Epidemics are typically initiated by introducing a small number of infected people into a population composed only of susceptible individuals, although immunity through prior infection, immunization, or natural causes can easily be modeled by initiating some fraction of the population in the recovered compartment. Steady-state equilibria can often be characterized analytically, but dynamics are difficult to characterize analytically even in simple models and are typically studied by numerical simulation.[4]

The dynamics of this model are captured by the following system of differential equations:

$$\dot{S}(t) = \nu N(t) - [\beta I(t)/N(t)]S(t) + \alpha R(t) - \mu S(t) \tag{1}$$

$$\dot{E}(t) = [\beta I(t)/N(t)]S(t) - (\kappa + \mu)\, E(t) \tag{2}$$

$$\dot{I}(t) = \kappa E(t) - (\gamma + \delta + \mu)\, I(t) \tag{3}$$

$$\dot{R}(t) = \gamma I(t) - (\alpha + \mu)\, R(t) \tag{4}$$

$$\dot{N}(t) = \nu N(t) - \mu N(t) - \delta I(t) \tag{5}$$

$$\dot{D}(t) = \delta I(t) + \mu N(t) \tag{6}$$

$$N(t) = S(t) + E(t) + I(t) + R(t). \tag{7}$$

Here, the rate at which susceptible people become infected, $\beta I(t)/N(t)$, is the core of the epidemiological model, and is called the *force of infection*. It is the only transition in the canonical SEIRS model that is not described by a rate parameter. It takes the general form,

$$\frac{\beta I(t)}{N(t)} = \frac{\tilde{\beta} A(\cdot) I(t)}{N(t)}, \tag{8}$$

where $\tilde{\beta}$ represents biological factors related to transmission conditional on an interaction, or contact, between infectious and susceptible individuals. The function $A(\cdot)$ is a contact function. In conventional models, behavior and the contact function are assumed to be constant throughout the epidemic and $\tilde{\beta}A(\cdot)$ is written as the single parameter, β. As is well known in epidemiology, doing so conflates the biological and behavioral aspects of the epidemic (Fenichel et al., 2011).

Our general specification accommodates different conventional transmission processes through the specification of $A(\cdot)$. When transmission depends on how many infectious people are in a given area, it is said to be *density-dependent*. One specification assumes that contacts increase proportionately with population size, $A(\cdot) = aN(t)$, where a is the density-dependent contact rate, leading to a force of infection that is proportional to the number of infected people,

$$\tilde{\beta} a I(t). \tag{9}$$

When transmission instead depends on the contact frequencies, which vary independently of population density, it is said to be *frequency-dependent*. When all people behave identically and contact each other randomly at rate a, the probability that a given contact is with an infected person equals the proportion of the population who are infected (the

[4]See Prodanov (2022) for a discussion of analytical solutions to the biological SIR model.

prevalence of disease), hence $A(\cdot) = a$ and the force of infection is proportional to the share rather than to the number of infected people,

$$\tilde{\beta} a \left[\frac{I(t)}{N(t)}\right]. \tag{10}$$

Density-dependent transmission is often preferred for airborne diseases, while frequency-dependent transmission is usually used for sexually transmitted diseases.[5] These different biological features can be incorporated into the SEIRS model simply by changing the form of the force of infection. Furthermore, if $N(t)$ is constant, then the two forms are identical after controlling for units. Hence, in the standard formulation whether matching is frequency or density dependent has no qualitative effect on the model's predictions as it amounts to scaling a parameter. As we will subsequently see, in economic models matching patterns have profound microeconomic implications, and the force of infection depends on behavioral responses to the risk of infection. In what follows, we will omit the tilde, and use β to denote the biological transmission coefficient, for clarity.

It is useful to introduce another important epidemiological concept here, the *basic reproduction number* of the disease, denoted by $\mathcal{R}_0$. The basic reproduction number is conceptually defined as the expected number of secondary cases resulting directly from the initial case in an otherwise susceptible population. In the SEIRS model above with, for simplicity, no births or deaths ($\nu = \mu = \delta = 0$, $N(t) = 1.0$), the average duration of an infection is $1/\gamma$, and in an otherwise susceptible population a single infection causes infections at rate β, so $\mathcal{R}_0 = \frac{\beta}{\gamma}$. For example, if the only infectious person remains so for an average of 4 days and on average infects 0.3 people per day, that person on average infects $\mathcal{R}_0 = 1.2$ others. It is common to generalize $\mathcal{R}_0$ to $\mathcal{R}_t$, the effective reproduction number, to represent the number of secondary cases generated by a new infection at time t. Note that if $\mathcal{R}_0 < 1.0$ at the start of an outbreak, then a small number of infected individuals does not start an epidemic, while if $\mathcal{R}_t < 1.0$ at any subsequent time during the epidemic, then it is waning.

The SEIRS model nests many compartmental model formulations used in mathematical epidemiology, and a large family of models has been developed by choosing parameters to remove one of the health classes from the system. In Table 1, we provide a taxonomy of the characteristics of the most common of these models, and illustrate their dynamics in Figure 2. These different forms are often compared in the absence of entrants into the population through birth or immigration. Many models omit the exposed class because in the canonical models this is just used to introduce a delay. Some forms of the model do not allow exit from the $R(t)$ compartment, providing permanent immunity once infected. Others impose permanent infection, which is useful for pathogens like HIV. Either of these assumptions make the transition graph acyclical, leading to a single long-run steady-state. The ability to return to the susceptible health class results in the possibility of multiple qualitative types of steady-state. In other forms of the model, there are only susceptible and infectious classes, which enables the models to be reduced to a single state variable, easing analysis.

Compartmental models can capture a wide variety of health or other states by introducing or removing compartments. For example, some mathematical epidemiological models

[5]See McCallum, Barlow and Hone (2001) and Begon et al. (2002) for more discussion and generalizations of transmission functions.

Table 1: Special cases of the SEIRS model and their key properties.

	Monotone dynamics	**Possible types of steady states**	**Permanent infection**	**Reducible to a single state variable** [6]	**Common applications**
SEIRS	No, potential for long running oscillations to a steady state	Two, eradication and endemic disease	No	No	Influenza and COVID-19
SI	Yes, and entire population becomes infected in a constant population.	Full infection	Yes	Yes	HIV/AIDS, livestock and wildlife diseases
SIS	No	Two, eradication and endemic disease	No	Yes	Syphilis, fungal infections
SIR	Yes for S and R classes, no for I class	One, with some fraction of the population remaining susceptible	No	No	Influenza and COVID-19
SEIR	Yes for S and R classes, no for E and I classes	One, with some fraction of the population remaining susceptible	No	No	Influenza and COVID-19
SIRS	No, potential for long running oscillations to a steady state	Two, eradication and endemic disease	No	No	Influenza and COVID-19

Note: monotone dynamics assumes no new entrants, and possible steady states hold population constant.

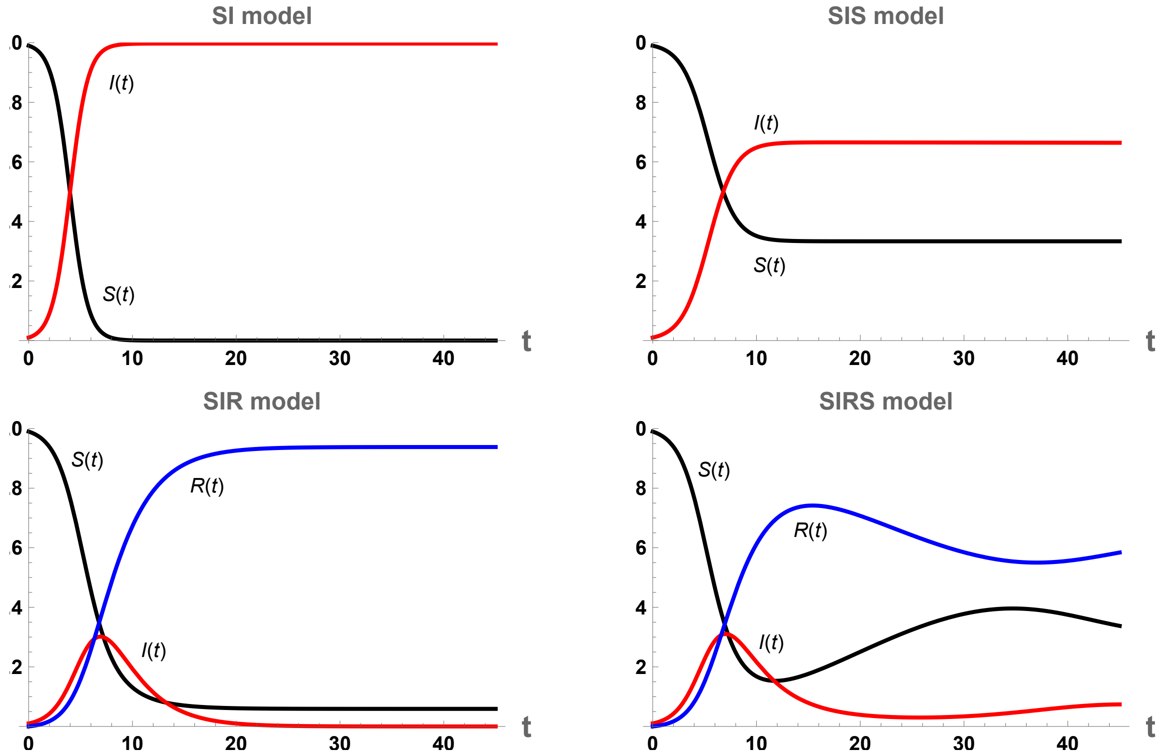

Figure 2: Various canonical compartmental model simulations.

introduce additional health classes to address symptomatic states, such as models with an asymptomatic but infectious class, SEAIRS. Asymptomatic individuals mix like susceptibles, transmit pathogens, but do not experience disease-induced mortality or other costs (e.g. Baril-Tremblay, Marlats and Ménager 2021; Engle et al. 2021 and Toxvaerd 2022). Some models let a share of asymptomatics recover directly with immunity. Other models introduce more complex epidemiological-behavior compartments, or compartments defined by demographic characteristics. However, behavioral responses to changes in the state of the epidemic are difficult to capture using compartments, and such behavioral responses are typically ignored or addressed in an *ad hoc* fashion in biological epidemiology.

Vital statistics are often omitted from epidemiological models based on the assumption that epidemics play out over a relatively short horizon, such as e.g., influenza, or involve a relatively small share of the population, such as e.g., HIV. However, vital statistics can qualitatively affect epidemic dynamics. Take for example the simple SIR model. Without any inflow of new susceptibles, the disease dies out asymptotically. In contrast, in the augmented SIR model with births and deaths, disease can become endemic (persist in steady state). This issue is then not merely a technical curiosity. For example, measles are completely vaccine preventable and recovery confers perfect and lifelong immunity. Yet measles persist to this day, due to a combination of births and incomplete vaccination coverage (Grenfell, Bjørnstad and Kappey 2001).

2.2 Behavior and Economic Epidemiology

Economic epidemiology starts from the premise that human behavior is intentional: When incentives change because the risk of contracting a disease has changed, behavior changes. These changes in behavior in turn affect how diseases spread through human populations.

Since interventions such as vaccines, quarantines, and treatments interact with behavioral changes to generate changes in the spread of disease, it is critical to consider endogenous behavioral responses when evaluating interventions. In this section, we discuss the development of the literature that derived from these ideas.

2.2.1 Modelling Individual Risk Exposure Decisions

Much of the early economic epidemiology literature was motivated by the HIV+/AIDS epidemic. The suffering and lost lives were devastating, projections were alarming, and compelling empirical evidence emerged documenting reductions in high-risk behaviors in response to the increasing threat of HIV infection. Echoing results in economic theory on risk–offsetting behavior such as Peltzman (1975), these papers emphasized that interventions that reduce disease prevalence will be at least partially offset by behavioral responses. These insights were formalized in models in which risky contacts are selected through individual–level optimization problems in which people face trade-offs between risky contacts and their health status or other goals, bridging the epidemiological literature and the health economics literature drawing on Grossman (1972).

The first contributions to bring economic theory to bear on the spread of disease appeared in the early 1990s. Brito, Sheshinski and Intriligator (1991) consider a model of endogenous vaccination, demonstrating that the conventional wisdom that compulsory vaccination is sound policy is generally false when vaccination is perfect. However, in this model disease dynamics are treated in an *ad hoc* fashion. Philipson and Posner (1993) studied simple rational choice models of whether to exhibit a binary protective behavior. Kremer (1994, 1996), and Geoffard and Philipson (1996) formally modeled optimal individual behavior and aggregate epidemic dynamics. Kremer (1994, 1996) considers continuous risky behaviors in a static behavioral framework, while Geoffard and Philipson (1996) present a susceptible-infected model in which agents solve an optimal control problem to decide whether to engage in a costly protective behavior.

Kremer's (1996) seminal contribution provides a simple model that highlights the key concepts that underlie the economic approach to modeling individual behavior and characterizing the resulting aggregate equilibrium, which we now sketch. Kremer begins with the SI model with a continuum of identical agents of unit mass in which the rate of change of the infected population is,

$$\dot{I}(t) = \beta a(t) I(t) S(t) - \gamma I(t), \tag{11}$$

where γ denotes the exogenous birth and death rate, and $A(\cdot) = a(t)$ denotes the action taken by the agent, that is, the contact rate. Recall that in the standard epidemiological model $a(t)$ is parametric and can therefore be combined with the biological transmissibility parameter, β. However, when economic behavior is introduced, these two factors must be separated. To understand the dynamics of the abehavioral model, observe that a single infection in an otherwise susceptible population causes an epidemic if $\dot{I}(t) > 0$ evaluated in the limit as $I(t) \to 0$, and therefore, in this model, if $\beta a > \gamma$. This reasoning is closely related to $\mathcal{R}_0$, which is in this case $\frac{\beta a}{\gamma}$. Setting $\dot{I}(t) = 0$ also reveals that steady-state equilibrium prevalence is $I^* = 1 - \frac{\gamma}{\beta a}$. Thus, the contact rate must be sufficiently high for this population to be at risk of epidemic disease, and if an epidemic takes hold, then the steady-state proportion of infected people is increasing in the exogenous contact rate.

To endogenize the contact rate, suppose that the agent is narrowly selfish and has preferences over the number of contacts and health status, h, represented by a utility function $U(a,h)$. If the agent has a contacts, then she believes her probability of infection is $p(a)$, which is increasing in its argument, and is either assumed to be linear as a simplifying approximation for low-risk agents, or is otherwise concave. The agent's problem is,

$$\max_{a} U(a, \text{healthy}) - p(a)\Delta, \tag{12}$$

where $\Delta = [U(a, \text{healthy}) - U(a, \text{infected})]$ is the loss, in units of utility, from becoming infected. The first-order condition implies that the agent continues to have risky contacts until the marginal benefit obtained from an additional contact equals the expected loss from that contact. In dynamic, forward-looking variants of this model, the loss due to infection can be interpreted as the present value of the expected difference in the value functions of being in the susceptible and infected states, respectively. That is, in an obvious notation and in discrete time, at time t the loss is $\Delta_t = \psi E[V^S(t+1) - V^I(t+1)]$, where ψ is the discount factor. It follows that any shock that affects the net expected benefit of remaining susceptible affects current behavior, for example, the announcement of a future intervention such as a vaccine or treatment will immediately affect behavior, and expectations that the epidemic will worsen may be self-fulfilling (Auld, 2003; Makris and Toxvaerd, 2020).

This problem can exhibit nonconvexities when the probability of infection $p(a)$ is concave. In this case, the marginal expected cost of an additional contact is declining in contacts, as the probability that the agent has already been infected by an infra-marginal contact is increasing in contacts. For example, if each contact causes an infection with probability β, then the probability that infection is not transmitted by any of a contacts is $p(a) = 1 - (1-\beta)^a$, such that $p'(a) > 0$ and $p''(a) < 0$. Let $a^*(I)$ denote optimal contacts at prevalence I. Due to this non-convexity, there may be multiple extrema, the corner solution $a^*(I) = 0$ may be preferred to the interior solution with $a^*(I) > 0$, and "demand" for risky contacts may not slope down everywhere: At sufficiently high levels of risk, increases in risk–per–contact may *reduce* the marginal probability of infection and induce a fatalistic increase in risky behavior.[7]

At low levels of activity, the probability of infection is approximately linear in the number of contacts, which rules out such fatalistic responses. Suppose such an agent believes that

$$p(a) = \beta I a. \tag{13}$$

Optimal contacts, $a^*(I)$, are then unambiguously declining in prevalence, virulence, and the cost of infection. Even if the agent does not become infected, then their well-being is still reduced because they obtain fewer contacts, which in this model represents the excess burden of disease.

An increase in prevalence reduces the number of risky contacts but, importantly, has an ambiguous effect on the probability of infection. Equation (13) implies that

$$\frac{dp(a^*)}{dI} = \beta a^* \left(1 + \eta \right), \tag{14}$$

[7]Philipson and Posner (1993), Kremer (1996), Auld (2003), and Kerwin (2012) develop the concept of fatalistic responses.

where $\eta = \frac{I}{a^*}\frac{da^*}{dI}$ is the prevalence elasticity of risky behavior. If an increase in prevalence causes a sufficiently large reduction in risky behavior, a response that is *prevalence-elastic*, i.e. $\eta \leq -1$, then the probability of infection is *decreasing* in prevalence (Philipson and Posner, 1993). For example, if a 10% increase in prevalence from 20% to 22% causes a 30% reduction in the contact rate, then the probability of infection falls. For this reason, the prevalence elasticity of behavioral response is a key parameter to be investigated in empirical work.

This model predicts that, other things being equal, a decrease in the biological transmission coefficient, β, leads to more risky behavior, whereas an increase in the cost of illness, Δ, reduces risky behavior. These behavioral responses suggest, then, that epidemics may be self–limiting: As prevalence increases, so do incentives to avoid infection, which reduces peak prevalence. In contrast, any intervention, technological advance, or other shock that reduces disease prevalence, will be at least partially offset by a resulting substitution effect that increases risky behavior. For example, quarantining some but not all infectious individuals reduces prevalence in the remaining population, thereby increasing the preferred contact level of susceptible individuals, spurring the epidemic. Moreover, effective but imperfect treatments may spur the epidemic, blunting or reversing the gains in health they would otherwise produce. Similarly, interventions that reduce the per-contact probability of transmission, β, such as masking or partially effective prophylactic vaccines, will either be less effective at reducing spread of disease, or may increase spread of disease if behavioral response is sufficiently strong (Philipson and Posner 1993; Toxvaerd 2019). Disease eradication becomes more difficult, or impossible, when incentives to protect against infection disappear as prevalence falls toward zero. Finally, profit-maximizing firms have perverse incentives to exploit this mechanism by pricing vaccines or treatments high enough to prevent eradication (Geoffard and Philipson, 1997). However, formal policy analysis requires determining how these individual responses aggregate to determine epidemic dynamics, a topic we now consider.

2.2.2 Individual Responses and Epidemic Dynamics

To understand how behavioral response aggregates and alters epidemic dynamics, consider the properties of the SI model when the contact rate, a, is determined endogenously through the process of agents optimizing (by solving the problem in equation 12) instead of being taken as an exogenous parameter. Notice that the force of infection in this model is,

$$\beta a^*(I(t))I(t). \tag{15}$$

Since optimal contacts are decreasing in prevalence, the force of infection may be increasing or decreasing in prevalence, as opposed to the biological model in which the force of infection is necessarily increasing in prevalence, as in equation (8).

In a steady-state equilibrium, prevalence satisfies (at an equilibrium with positive prevalence),

$$I^*(a) = 1 - \frac{\gamma}{\beta a^*(I)}, \tag{16}$$

which is increasing in the contact rate, a. So long as $a^*(I)$ is decreasing, the model predicts a unique steady state equilibrium, as illustrated in the left panel of Figure 3, which displays prevalence (I) and risky behavior (a) as determined in equilibrium by the response function $a^*(I)$ and the epidemiological relationship $I^*(a)$. The steady state,

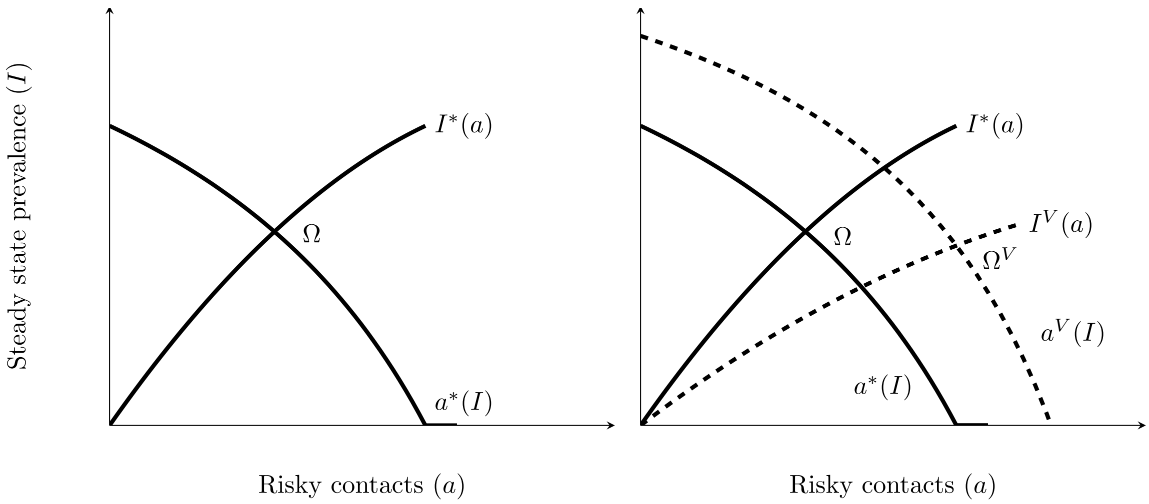

Figure 3: Equilibrium in an economic SI model

endemic equilibrium is point Ω.

An immediate implication of endogenous behavioral responses in this model is that epidemics begin exactly as in the abehavioral model, but generally converge to lower steady–state prevalence than would be obtained if the contact rate were held fixed throughout the epidemic. At the start of an epidemic prevalence is approximately zero, so that as in the abehavioral model, an epidemic starts if $\mathcal{R}_0 = \beta a^*(0)/\gamma > 1$. But as prevalence increases, $a^*(I)$ decreases, and in equilibrium $I^* = \beta a^*(I^*)/\gamma < \beta a^*(0)/\gamma$.

To illustrate how the economic SI model may exhibit different properties and policy implications than a conventional SI model, consider modeling the effects of a partially effective prophylactic vaccine. The vaccine is modeled as reducing, but not to zero, the virulence parameter, β. All else equal, an individual who receives this vaccine generally responds by increasing his or her contact rate and faces a *higher* probability of infection if this response is prevalence-elastic.

The right panel of Figure 3 shows how a partially effective prophylactic vaccine may affect the spread of disease. The vaccine reduces equilibrium prevalence at any level of risky behavior by reducing the probability that a contact between a susceptible and an infected person causes an infection (the $I^*(a)$ schedule shifts out to $I^V(a)$) and increases risky behavior at any given prevalence for the same reason (the $a^*(I)$ shifts out to $a^V(I)$). The new equilibrium is Ω^V. Since both schedules shift out, the model unambiguously predicts that the vaccine is expected to increase the contact rate, but whether the vaccine retards or spurs the epidemic depends on the biological effectiveness of the vaccine and on the magnitude of the behavioral response. In contrast, if behavior were fixed as in standard epidemiological models (the $a^*(I)$ schedule is perfectly inelastic), then the vaccine unambiguously decreases prevalence.

A subtle yet important difference between epidemiological models and most conventional economic models is that heterogeneity in behavior can affect epidemic dynamics in potentially counterintuitive ways. This effect occurs because heterogeneous preferences lead to heterogeneous behaviors that alter the composition of the pool of contacts. To see this, suppose that there are equal masses of two types of people, denoted l and h, both of whom behave according to the model discussed above, who differ in their preferences. One has higher marginal utility for risky contacts and consequently has more contacts and higher prevalence, that is, $a^l < a^h$ and $I^l < I^h$ (dropping time indexing for clarity). Assuming that mixing is at random conditional on contact rates, such that the probability that a randomly selected match is with type j is $a^j/(a^l + a^h)$, then the probability that a given match is with an infected person, P, is the product of the probability of matching with a low-activity type times the probability that a low-activity type is infected plus the probability of matching with a high-activity type times the probability that a high-activity type is infected,

$$P = \frac{a^l I^l + a^h I^h}{a^l + a^h} = wI^l + (1 - w)I^h, \tag{17}$$

where w is the proportion of total contacts from low-activity people.

Then, an increase in risky behavior by one type affects outcomes in part because every match is now more likely to occur with that type of person. Due to this dependence on matching patterns, increases in average risky activity may decrease spread of disease when they decrease the probability of matches with high-activity types. Here, P is decreasing in a^l, that is, if low–activity people *increase* their activity rates all else equal, then the probability that any given match is with an infected person *decreases*. Similarly, notice that a mean-preserving spread in contact rates, equally increasing a^h and decreasing a^l, increases P, and that generally the mean of the distribution of contact rates is not sufficient to determine epidemic dynamics. This positive feedback may generate multiple equilibria, one in which only high-activity people have many partners and risk per contact is high, and another in which all types have a moderate number of partners and risk per contact is low. This is a classic economic selection process. Other types of heterogeneity that affect epidemics are endogenous heterogeneity and situational heterogeneity, either of which can also affect the pool of available contacts. We now discuss these and others sources of heterogeneity in behavior.

2.3 Sources of Heterogeneity in Behavior

Most, but not all, epidemiological modeling relies on the simplifying assumption that the population is well-mixed conditional on modeled class, so that the probability of a random encounter between two individuals can be characterized by using population frequencies of different types across the population. The assumption of homogeneous mixing corresponds to the use of mean-field approximations in network-based models. This assumption simplifies the analysis but does not do justice to the many sources of heterogeneity in behavior observed in practice. The standard dynamic disease model, based on the assumption of homogeneous mixing, can be extended to heterogeneous population settings by specifying equations for each subgroup of the population and by determining the frequency of interaction across these subgroups (see e.g. Ellison 2020). Such metapopulation models are alternately interpreted as capturing heterogeneity in some pertinent characteristic such as age, gender, preferences, or vulnerability, or as geographical (location) heterogeneity. Yet, this approach retains a significant amount of homogeneity, since within groups there is still homogeneous mixing. In addition, it requires a specification of a large number of hard-to-identify parameters, even when behavioral responses to the epidemic are assumed away.

An alternative approach to compartmental models and metapopulation models are "micro-dynamic" models such as cellular automata and agent-based models (see Wang, Zheng and Liu (2022) for a review). Agent-based models (ABMs) build synthetic populations, agent interactions, and disease transmission from the ground up, as described in e.g. Eubank et al. (2004, 2006) and in Venkatramanan et al. (2018). First, a statistically accurate synthetic population is created on a grid. The geographical distribution of the population can take into account a number of pertinent characteristics, such as age, gender, and socioeconomic status. Second, each individual is endowed with a sequence of activities tied to specific locations, such as home, work, school, retail, and public transport. Last, a disease model is postulated, where disease transmission between individuals is tied to co-location on the grid. The two first components induce a dynamic social network that keeps track of how individuals are connected at each point in time. Together with the disease model, this enables the analyst to simulate the movements and contacts of individuals and the spread of the disease. Although agent-based models allow for very sophisticated dynamics, detailed descriptions of individuals and contacts, and a substantial amount of heterogeneity, they often suffer from some of the same drawbacks as traditional homogeneous-mixing compartmental models. Both are essentially mechanistic in nature, with individuals' movements across space during an epidemic unchanged from stable movement patterns during normal times. In other words, behavior is exogenous: Individuals are assumed not to change activity patterns with a view to avoid infection. However, we know that during epidemic outbreaks individuals shun locations perceived to be crowded (and thus risky), such as public transport, entertainment venues, or retail outlets, while spending more time at home or in open spaces such as parks. This means that the movement patterns that researchers record during disease-free times are no longer applicable when infections spread through the population, and in addition, will tend to change over time as the epidemic evolves.

2.3.1 Innate Heterogeneity

Individuals have a number of different innate characteristics that may condition their behavior. Some of these characteristics are purely preference-related, such as preferences over exposure and activity, costs of disease mitigation, rates of time preference, and the extent to which infection is detrimental to their individual well-being. Heterogeneity in biology also affects behavior and transmission. Some people have innate immunity to some diseases, for example, and susceptibility to infection may vary with characteristics such as age, ethnicity, and biological sex. All these characteristics influence best responses and consequently influence transition rates and aggregate dynamics. In Section 3.1.1 we will see that heterogeneity that causes differences in contact rates can have deep implications for both disease dynamics and welfare economics. Makris (2024) explores heterogeneity in an equilibrium model of social distancing, while Acemoglu et al. (2021) do so in an optimal planning model.

2.3.2 Endogenous and Situational Heterogeneity

Even if individuals are *ex ante* homogeneous, and if there is no aggregate uncertainty in disease dynamics, there will be *ex post* heterogeneity because individual transitions are stochastic, even if they make the same decisions *ex ante*. For example, individuals who adopt an imperfect vaccine at the same time may be homogeneous *ex ante*, but *ex post* some may become infected while others remain healthy. As an epidemic progresses, otherwise homogeneous individuals experience different transition histories between health states. This means that at a given point in time, individuals find themselves in different situations, with widely different state payoffs, incentives, and future prospects.

Heterogeneity across health states that interacts with behavior complicates the contact function, $A(\cdot)$. Following Fenichel (2013), one formulation of transmission is that the contact function introduced in equation (8) takes the form,

$$A(\cdot) = N(t)\left[\frac{a_S(t)a_I(t)}{S(t)a_S(t) + I(t)a_I(t) + R(t)a_R(t)}\right]. \tag{18}$$

Here, a_j is the activity level of agents in health state j (which does not vary among the members of a health classes). Notice that this expression collapses to a common activity level, $a(t)$, if there is no heterogeneity. The term in brackets reflects how the structure of social contacts affects the spread of disease at any given population size. We consider implications of this and other contact patterns in Section 3.1.1.

Policies that target different subgroups depending on disease state have a long history. Optimal quarantine of infected people was first studied by Sethi and Staats (1978) under SIS dynamics, while Gersovitz and Hammer (2004) consider a variety of possibilities, including shielding of the immune, quarantine of the infected, and a blanket lockdown that targets susceptible and infected people alike. Fenichel (2013) considers the imposition of a blanket restriction on the activity of all people and observes that such a general non-targeted policy may yield lower welfare than the laissez-faire equilibrium. This is because a uniform policy is a restriction on the planner's set of policies, which means that the outcome under decentralized decision making is not nested within the set of outcomes achievable under central planning. Allowing people to act on their heterogeneous characteristics may yield higher aggregate welfare as better targeting is achieved, relative to a uniform policy that treats everyone equally. This also points towards the importance of situational targeting and the value of information about health status.

2.3.3 Geographical and Positional Heterogeneity

Models that abstract from how the disease spreads geographically may obscure policy-relevant phenomena. Hoover and Toxvaerd (2022) note that welfare under geographically non-targeted policies, even if chosen with the intent of maximizing social welfare, may be dominated by outcomes under laissez-faire, where individuals make choices in a decentralized, uncoordinated manner. This is because the setting with decentralized decision making is nested within that of an unconstrained social planner, while untargeted (or blunt) policy making cannot achieve the outcomes that emerge under decentralized decision making. Geographic targeting is done at the international level as a matter of course, but such policies have also been applied at the subnational level, such as the tier policy in the UK and the policy of hotspot regulation in Germany during the COVID-19 pandemic.

The analysis of epidemic spread based on homogeneous mixing comes at a cost, namely that interaction is with anonymous, random individuals in the population. In some contexts, as is the case with most sexual interactions, contacts are on the whole neither random nor anonymous, and thus a different modeling approach is necessary to capture such endogenous matching. For this kind of analysis, network-based approaches are particularly useful. The theory of epidemics on networks is well developed, yet the application of such models in integrated economic-epidemic analyses with behavior is still in its infancy.[8]

In addition to epidemics on networks, there are three other main approaches to spatio-temporal analysis in epidemiology, namely microparameter or partial differential equation approaches of the Hochman and Zilberman (1978) form applied to epidemics (e.g., Veliov, 2005), diffusion models (see, e.g., Okubo, Levin et al. 2001 and Otto and Day 2007), and metapopulation models (see Hanski 1999). To the best of our knowledge, only the latter has been applied to economic-epidemic analyses, but there is no reason in principle that the other frameworks could not be fruitfully integrated into economic-epidemic analyses of disease dynamics.

2.4 Evidence and the Positive Economic-Epidemiological Framework

2.4.1 Econometric Research on Behavior Prior to COVID-19

Here, we briefly summarize some of the empirical literature that addresses the predictions of positive models regarding behavior. Even this literature is vast, and the broader literature attempting to estimate the effects of policies on health and other outcomes is even more so. We therefore limit attention to selected papers that provide evidence on endogenous responses to the threat of disease, although such papers often nevertheless involve estimating policy effects in disentangling changes in behavior caused by policy and those caused by the epidemic itself.

The common assumption in biological epidemiology that behavior over the course of an epidemic is fixed was first heavily challenged during the rise of the HIV+/AIDS epidemic in the 1980s. Substantial decreases in risky sexual behavior, including reduced rates of partner change and increased use of condoms among gay men, were well documented (e.g. McKusick, Horstman and Coates 1985). Early econometric work extended these findings by estimating the magnitude of various behavioral responses to a given increase in disease prevalence. Other strands of econometric analysis attempt to estimate the effects of interventions on behavior or disease spread. An important contribution of economic epidemiology to the field of epidemiology and public health has been to demonstrate that behavior is sufficiently responsive to disease risk to shape aggregate risk. These studies deploy econometric methods to estimate causal effects from observational data, such as instrumental variables, difference–in–difference type models for panel data, and estimable structural models.

Pioneering empirical work on estimating behavioral responses to infection risk includes Geoffard and Philipson (1996), who used time series data on AIDS incidence and risky sex among gay men in San Francisco to generate a rough estimate of the prevalence elasticity of risky sex. Philipson (1996) showed that the local prevalence of measles increased measles vaccination rates, while Mullahy (1999) similarly found that the prevalence of local flu, along with a host of economic factors such as wages, drives flu vaccination decisions. Ahituv, Hotz and Philipson (1996) used panel microdata to show that condom use among heterosexual Americans was highly responsive to local AIDS prevalence. Together with

[8] See e.g., Jackson and López-Pintado 2013; Newman 2002; Pastor-Satorras et al. 2015 for examples of abehavioral network models. Talamàs and Vohra (2020) present a network model with behavioral choice, and Julliard, Shi and Yuan (2023) provide an early empirical contribution to this strand of literature.

many other contributions, this early work on behavioral response confirmed that changes in the threat of communicable disease lead to changes in behavior.

Subsequent pre-COVID-19 econometric analyses of the effect of changes in incentives on risky behavior include Li, Norton and Dow (2004) and Oster (2018), who use panel data to study how flu vaccination, pertussis vaccination, or both, respond to local disease outbreaks. Oster (2012) studies behavioral responses to the AIDS epidemic in Africa instrumenting prevalence with distance to the origin of the virus, while Greenwood et al. (2019) estimate a structural model of risky sexual behavior in Malawi. Auld (2006) investigated heterogeneity in behavioral response to the HIV epidemic in San Francisco in the 1980s. Kremer's (1996) theoretical prediction of fatalistic responses, that is, that high-risk people may respond to increases in risk per contact by increasing risky contacts, is supported by Kerwin (2012, 2020) and Matthies and Toxvaerd (2023), who document fatalistic responses to changes in risk among high-risk groups. Similarly, Auld (2006) showed that gay men with high activity in San Francisco in the 1980s reduced risky activity relatively less than their lower-activity counterparts in response to increased prevalence of HIV. Delavande and Kohler (2016) estimate the effects of beliefs on risky sexual behavior, addressing the endogeneity of beliefs using a simultaneous equations method.

The 2009 H1N1 "swine" flu pandemic was the first case in which econometrics was used to explore voluntary social distancing behavior in response to pathogens that could be transmitted through casual contact. Fenichel, Kuminoff and Chowell (2013) find that air travelers abandoned purchased tickets to reduce the risk of infection. Bayham et al. (2015) use the American Time Use Survey to estimate behavioral shifts and show voluntary social distancing. They also showed that the behavioral response was of a sufficient scale to influence the macro-level pathogen dynamics. Springborn et al. (2015) used television viewing data in Mexico City to estimate social distancing and pandemic fatigue.

Other econometric work has explored systematic response to incentives other than changes in risk per contact. Boozer and Philipson (2000) present estimates of the effects of learning one's HIV status on behavior, finding that people surprised by information tend to change behavior in a manner that may spur the spread of disease: Low-risk people learning that they are HIV-positive increase risky behavior, while high-risk people learning that they are HIV-negative decrease risky activity. Several related papers exploit data from the Malawi Diffusion and Ideational Change Project (MDICP), which reports beliefs about own–HIV status to study the effects of beliefs and information on behavior[9]. Mechoulan (2007) and Lakdawalla, Sood and Goldman (2006) estimate the effects of new effective treatments for HIV+ disease and find, consistent with theory, that reducing the expected health costs of the disease increases risky behavior.

2.4.2 Econometric Analysis During COVID-19

The economic literature on estimating endogenous behavioral responses to the threat of communicable disease expanded greatly in response to the COVID-19 epidemic. This work extends previous empirical work primarily in that it focuses on addressing the difficulties inherent in disentangling the effects of policy interventions from endogenous behavioral responses, which allows insights into policy efficacy and the magnitude of the behavioral responses predicted by economic theory.

[9]Some examples include Paula, Shapira and Todd 2014; Thornton 2008, and Kerwin 2020.

Fundamentally, disentangling policy and endogenous responses is challenging because individuals and policy makers are responding to the same changes in the state of the epidemic, so policies may spuriously appear to be more or less effective in changing risky behavior than they actually are. To fix ideas, consider attempting to estimate the effect of a stay–at–home order on social contacts, using data aggregated at a regional level. If prevalence or other measures of the state of the epidemic are not included as controls, estimates are subject to potentially severe biases, as—in addition to the usual issue of unobserved confounders—both behavior and policy are caused by the state of the epidemic. Times and places in which the epidemic is most severe will also tend to have the largest endogenous reductions in social contacts and the highest likelihood of policy responses such as stay–at–home orders. A naive analysis may then conclude that the stay-at-home order was much more effective than it actually was in reducing risky activity, as the endogenous behavioral response to the severity of the epidemic is misattributed to the stay-at-home-order.

Controlling for prevalence when attempting to estimate the effects of policy on behavior is also problematic. However, policy and the state of the epidemic are likely to be endogenous due to "reverse" causation: Policy makers may implement policies to limit disease spread when they believe prevailing behaviors will lead to rapid spread of disease (behavior causes policy), and private behaviors to avoid risk affect prevalence (behavior causes prevalence). With panel data, these considerations can be mitigated, but further problems arise due to the nonlinear spread of cases and the possibility that expectations of future disease outcomes affect current policy making and current risk behaviors.

The econometric approach to these issues differs from methods commonly used in other literatures in how counterfactuals are generated. Counterfactuals generated by the structure of abehavioral compartment models tended to over-forecast deaths during the initial phase of COVID-19 and implied that policies would be extremely effective. A notable example is Ferguson et al. (2020), which predicted more than 20 times more deaths than actually occurred in the US by mid-2020 and estimated that a combination of policies could reduce those deaths by about half. Failure to model endogenous behavioral responses is one reason these models tend to make overly pessimistic forecasts. Korolev (2022) shows that this class of model is not identified in that many different compartmental models generate identical epidemic dynamics, even when policies across these models have markedly differing effects.

Despite the difficulties highlighted above, econometric analysis overwhelmingly shows that endogenous behavioral responses to COVID-19 were large. A simple observation demonstrating that behavior changes absent policy interventions is that social distancing began before any policy intervention occurred. Aum, Lee and Shin (2021) leverage this reasoning to study endogenous responses in Korea, which did not implement lockdowns or business closures, finding large changes in employment as COVID-19 progressed across regions even in the absence of such policies. Similarly, Maloney and Taskin (2020) emphasizes that restaurant reservations in the U.S. and movie spending in Sweden fell largely before any policy responses occurred. Further, voluntary social distancing and compliance with mandates have been shown to be related to partisanship and beliefs about disease risk; for example, Republican voters in the U.S. engaged in less voluntary distancing and believed that they were at lower personal risk of COVID-19 than Democrat voters did (e.g., Allcott et al. 2020; Makridis and Rothwell 2020). The effects of social distancing on the spread of disease, interpretable as estimates of the slope of the theoretical relationship (16), are presented by Glaeser, Gorback and Redding (2022), who found that cases fall by 19% in response to a 10 percentage point decrease in social mobility.

Two-way fixed effects models and modern variants using similar identification strategies (e.g. Sun and Abraham 2021) show similar results over longer time periods, exploiting staggered implementation and retraction of various policies to separate policy and endogenous behavioral responses. For example, Goolsbee and Syverson (2021) disentangle the effects of stay–at–home orders from endogenous behavioral responses by comparing behavior within the same commuting zone subject to different policy regimes. They find that only 7 of a 60 percentage point fall in consumer traffic can be attributed to policies, with the remainder attributable to endogenous behavioral responses. Yan et al. (2021*a*) use a mix of fixed effects and synthetic control identification strategies and estimate that only about 4% to 10% of the additional time people spent at home could be attributed to stay–at–home orders, although school closures and other interventions also increased time spent at home. They also developed the concept of equivalent cases to estimate the additional cases required in a US county needed to induce the same number of individuals spending additional time at home, finding that for 50% of US counties this is less than 30 additional cases. Similarly, Gupta et al. (2021) find that about half of the change in behavior they observed in the initial stages of COVID-19 in the United States can be attributed to all interventions taken together, the remainder attributable to private responses. Further results documenting endogenous responses to increases in prevalence are reported in, for example, Berry et al. (2021) and Gupta, Simon and Wing (2020).

Another branch of this literature augments econometric panel models with structural models of disease spread. For example, Atkeson, Kopecky and Zha (2021) develop and estimate an SIR model in which the contact rate is endogenously determined by current disease prevalence. Chernozhukov, Kasahara and Schrimpf (2021) develop and estimate a causal model that incorporates specification choices based on a SIR model with deaths. In many econometric models, endogenous behavioral responses are implicit, that is, they are "in the residual," while the effects of policies are explicitly estimated, but Chernozhukov, Kasahara and Schrimpf (2021) include measures that proxy information about local and national disease spread, interpreting the estimated effects of these information variables as private behavioral response. They find that while policies are important, the bulk of response would have occurred absent policy. Bisin and Moro (2021) develop an estimable spatially structured network model in which individuals and firms change their behaviors to avoid infection risk and demonstrate the empirical importance of 'epidemiological Lucas Critique' considerations in which policy makers systematically make errors if they fail to address endogenous responses to epidemic dynamics.

Relatively little work has confronted the predictions of rational choice models that effective treatments and interventions that reduce risk–per–contact should increase risky behavior. For example, the model developed in Section 2.2.2 above predicts that a partially effective COVID-19 vaccine should reduce social distancing while having an ambiguous effect on the spread of disease. Examples of work estimating the effects of vaccination on behavior include Auld and Toxvaerd (2021) and Kim and Lee (2022). Both find that the effects of vaccination on retarding the spread of disease are blunted by offsetting decreases in social distancing, but neither uses disaggregated data such that the effects of vaccination on the vaccinated cannot be disentangled from general equilibrium effects. Abaluck et al. (2022) is notable as the only RCT on COVID-19 mask mandates. They report that mask mandates had no discernible effect on social distancing. However, Yan et al. (2021*b*), using a dynamic event study model with smartphone location data, find an increase in time outside the home following mandated mask wearing in the United States. To our knowledge, no studies have investigated whether the introduction of effective treatments for COVID-19, such as paxlovid, caused increases in risky behavior.

The contemporary econometric focus on identification was also important during the COVID-19 pandemic as scientists and policy makers tried to understand the causal effects, as opposed to associations, of interventions in near-real time.[10] The availability of granular, daily data on cell phone location provided researchers with rich and timely opportunities to study endogenous behavioral changes and policy impacts (Grantz et al., 2020). However, many interventions were initiated jointly and finding credible sources of variation was challenging, and the early wave of empirical work received substantial criticism for these reasons.[11]

Overall, despite challenges stemming from observational data, econometric evidence on behavioral response to epidemics, consistent with results from other disciplines, overwhelmingly supports the prediction of positive economic theory that epidemics induce endogenous behavioral responses. We close by emphasizing that our review above is necessarily highly selective due to the sheer number of relevant papers: Bürgi and Wohlrabe (2022) report that, between January 2020 and August 2021, almost 1,500 of about 9,500 working papers issued in five leading economic series were related to COVID-19.

3 Normative Economic Epidemiology

Having established the usefulness of economic methods in generating positive insights, predictions, and evidence in epidemiological contexts, we now turn to the normative insights that economic epidemiology provides.

Murray (2020) observed that "epidemiologists as a discipline are singularly focused on saving lives but are generally unprepared to provide recommendations on how individuals can best live those lives." We begin our discussion by highlighting that economists are not singularly focused on mortality, but rather take a more holistic view of well-being in which health is an important, but not the only, goal. For example, in Section 2.2.2 we showed that a partially effective vaccine might spur the spread of disease. But even if introducing the vaccine causes more disease, it may increase welfare because contacts also increase, and those are valuable too. This focus on a more inclusive notion of welfare than health alone differentiates the economic approach.

Generally, normative economic epidemiology addresses how infectious diseases impact welfare, the potential for infectious diseases to generate externalities, and hence the role for policy interventions to improve social welfare. Tradeoffs made at individual level, considered in the section on positive economics, can generate externalities at the societal level. An efficient economy is one that internalizes these externalities such that those bearing the costs of providing public health benefits to others could potentially be compensated by those enjoying the benefits. The traditional approach in the literature is to analyze socially optimal disease paths by considering the policies implemented by a benevolent,

[10] See Acemoglu et al. (2021); Chu et al. (2020); Flaxman et al. (2020); Hsiang et al. (2020); Quaas et al. (2020); Sears et al. (2023).

[11] See for example Allen (2022); Haber et al. (2022).

utilitarian social planner. The optimally controlled dynamics can then be contrasted with those that result from equilibrium under decentralized decision making, where individuals act in their own private interest. A key difference between those settings is that, while individuals take aggregate disease dynamics as given, the social planner takes into account that their decisions alter the future path of the system.

3.1 Externalities and the Rationale for Policy Intervention

There is a commonly held belief that markets do not coordinate behavior sufficiently to provide an efficient level of protection against infectious diseases. This belief is borne out by the fact that there is a whole field of *public* health. Before discussing different approaches that compare and contrast efficient behavior during an epidemic with potentially inefficient decentralized behavior, it is helpful to discuss some of the potential externalities generated by infectious diseases. As noted by Gersovitz and Hammer (2004), the spread of communicable diseases is rife with externalities. A taxonomy of these externalities is presented in Table 2 to help organize thinking. Most of the economic epidemiology research is not explicit regarding which externalities are at play. Rather, research has focused on the interaction between a model of disease dynamics such as those summarized in Table 1 and a specific policy intervention, for example, vaccination or lockdowns. In practice, multiple types of externality may coexist, and it is often difficult to disentangle the effects of each on overall outcomes. Models of disease dynamics play out over time, and decisions made by individuals at a given moment may influence the welfare of others, both contemporaneously through the matching process and intertemporally through the time path of the dynamic system.

How underlying disease dynamics (e.g., SEIRS, SIR, SIS) influence infection, intertemporal, and matching externalities is at the core of normative economic epidemiology. That infection passes from one individual to the next potentially generates the first three types of externalities in Table 2. The first external effect, the infection externality, occurs when self-regarding susceptible individuals underprotect from infection because they do not account for the fact that if they become infectious, they may infect others. That is, they do not account for the effect of their health on others' probability of contacting an infectious person.

Similarly, matching externalities operate through affecting the probability that a contact is with an infected person. However, matching externalities affect this probability through contact rates, rather than through health status. For example, in equation (17), the infection externality operates through the prevalence terms and the matching externality operates through the contact rates. To illustrate the conceptual difference, consider a person who is completely immune. This person's contact rate does not confer an infection externality but does confer a matching externality: If they increase their contact rate, they reduce the probability of matching with an infectious person for every person (assuming mixing is random given contact rates). Matching externalities result from asymmetric information: A susceptible individual who could observe the infection status of potential matches could in principle avoid matches with infected people, or the infected individual could compensate the susceptible individual for the contact, generating incentives to avoid infection.

The intertemporal externality occurs because actions today affect the dynamics of the epidemic, so actions today can also affect the probability that a match in the future

Table 2: Taxonomy of infectious disease externalities

Externality	**Description**	**Receives benefit**	**Bears cost**
Infection	Susceptible individuals undervalue their own health in the sense that they do not take into account the effect of their health on others.	Susceptible individuals	Other susceptible individuals
Intertemporal	Susceptible individuals undervalue their own health when infection perpetuates the epidemic and increases the likelihood of reinfection.	Susceptible individuals	Future susceptible people
Matching	Occurs when a person's contact rate affects the probability of matches between susceptible and infected people.	Susceptible individuals	Susceptible individuals
Immunity	Immunity prevents transmission to others.	Susceptible individuals	All groups
Defensive	Individuals who engage in defensive behavior may not account for the costs of reduced economic and non-economic social interactions	Those engaging in defensive behavior	All others
Congestion	Medical resources and other "essential services" may be congestible possibly due to preexisting market distortions or the inability to adjust pricing.	Those who use medical facilities first	Those who miss out, perhaps healthcare workers

is with an infectious person. Immunity creates a fourth type of externality. Immunity protects the immune individual, thus providing a private benefit. Immunity also depletes the infection pool for some specifications of the transmission function, providing a public good that is often uncompensated. This process is popularly called *herd immunity*, and is said to be present when the share of immune people in the population is large enough to start reducing new infections even without further intervention. The immunity externality is central to public health and forms the basis of most policy recommendations on the use of vaccines to manage epidemics.[12] Therefore, the hypothesis is that people will tend to underinvest in acquiring immunity relative to the social optimum. This is most often discussed in the context of vaccination and potential free-riding on others' vaccination.

There are two additional potential types of externalities. Defensive and congestion externalities do not depend directly on the epidemiological structure. These externalities may be better thought of as extending or complementing pre-existing market distortions because of behaviors that emerge during an epidemic. They may emerge due to fixed costs, relatively slow market adjustments, or other constraints. Defensive externalities involve the externalized costs of protecting one's own health. Contacts occur because they are valuable, although some contacts may be essential and not merely incidental (Toxvaerd, 2024). Philipson and Posner (1993) and Toxvaerd (2017) consider the case of sexual contacts, whereas Fenichel (2013) and McAdams, Song and Zou (2023) consider other social and economic externalities. For example, if workers stay home to avoid infection, this can reduce the quality of services, increase prices, or reduce wages for others (see Section 3.7 for a related discussion). Bayham and Fenichel (2020) show how closing schools could lead healthcare workers to stay home to care for children, which could reduce hospital services and increase pathogen-related mortality. Defensive externalities also occur when protective measures reduce the benefits of non-market social interactions, for example, if person A refrains from attending a party in an effort to avoid becoming infected, party-goers lose the pleasure of person A's company. The congestibility of healthcare resources was highlighted during the COVID-19 pandemic, manifesting in the public sphere as the call to "flatten the curve." There are large fixed costs associated with the establishment of new health resources, and the development of these resources can be slow relative to surges in demand during an epidemic. In addition, price rationing is generally viewed as socially unacceptable.

In addition to these externalities, there is another effect that is potentially relevant. This effect stems from the observation that in a competitive economy, each individual may act as if his or her actions cannot influence the aggregate evolution of the disease. The analysis in Rowthorn and Toxvaerd 2012 suggests that as a result, in settings with a continuum of individuals, each individual receives a different 'return on investment' than a social planner who can influence the aggregate path of infection, and that this discrepancy causes even completely altruistic individuals to engage in too little mitigation. Relatedly, recent work by Kaufmann, Andre and Kőszegi (2024) describes conditions in models with a countably infinite number of individuals where individuals "internalizing the externality" may not lead to socially optimal outcomes. However, these analyses raise subtle issues whose implications are only beginning to be explored.

[12] See e.g. the exposition of vaccination policy in Keeling and Rohani (2008).

3.1.1 Matching Patterns and External Effects

Here, we highlight the welfare-economic implications of assumptions over matching patterns. As discussed in Section 2, in canonical epidemiological models of spread of disease, whether matching is density or frequency dependent amounts to changing the interpretation of the parameter β, so there are no qualitative implications when the population size does not vary over the epidemic. But in economic models, the assumed matching pattern can have deep implications for welfare analysis. To see this point, consider again the case discussed in section 2.3.2 in which people in different health states may exhibit different contact rates, but there is no heterogeneity within health classes. Every individual within health class j exhibits activity level a_j, for $j \in \{S, I, R\}$. We contrast the implications of this heterogeneity across three types of matching patterns.

First, suppose that matching is purely density dependent, that is, contacts are assumed to simply scale with population density, as considered, without heterogeneity, in equation (9). In this case, "all infectious particles rain down equally (whatever their source) on all susceptibles," in the phrasing of Begon et al. (2002), and the parameter a_I can be interpreted as the amount of pathogen emitted per infected agent. The contact function then takes the form

$$A(\cdot) = a_I N(t). \tag{19}$$

Substituting this expression into equation (8), one can see that this assumption implies that the force of infection is proportional to $a_I I(t)$, the total amount of pathogen emitted by infected agents. Here, the risk to susceptible people is analytically equivalent, despite its nature as stemming from communicable disease, to an atmosphere externality familiar from environmental economics. Since the activity levels of infected and recovered individuals have no effect on disease dynamics, this specification also allows arbitrary differences in activity levels across health states. Since, however, transmission does not depend on the frequency of social contacts, ruling out, for example, susceptible people reducing their contact frequency as a protective measure, this form of density dependence does not seem suitable for most economic epidemiological analysis.

If matching is instead frequency dependent, the analyst must be careful when specifying how the number of potentially transmissive contacts between susceptible and infected people varies with the activity levels of each group. One contact function in this case is

$$A(\cdot) = a_S(t)S(t)a_I(t)I(t), \tag{20}$$

where for convenience the population size is now normalized to unity, and $S(t)$ and $I(t)$ are the proportions of susceptible and infected people. The multiplicative nature of this specification, known consequently as quadratic matching in the terminology of Diamond and Maskin (1979), captures the idea that more social activity from one group of people increases the activity of other groups.[13] Notice in this case that the behavior of recovered people does not affect the number of potentially transmissive contacts between infected and susceptible people, which is appropriate when modeling an airborne disease such as COVID-19. Conversely, under linear matching, increases in the activity of one group change the composition of social contacts but do not change its level. In this case, a specification for the contact function is given by equation (18), which we repeat here for

[13]See Farboodi, Jarosch and Shimer (2021) on the relation between matching technologies and externalities in an epidemiological context.

convenience (recalling that we have now normalized $N(t) = 1.0$),

$$A(\cdot) = \left[\frac{a_S(t)a_I(t)}{S(t)a_S(t) + I(t)a_I(t) + R(t)a_R(t)} \right]. \tag{21}$$

Here, increases in activity by recovered people reduce the risk faced by susceptible people, which is most closely associated with modeling a sexually transmitted disease, but is also helpful for allowing homogeneous behavior to be a special case of heterogeneous behavior. See Chen (2012) for an example of how matching patterns may create conditions for multiple equilibria.

When there is heterogeneity in contact rates within health classes, even the sign of the external effect can depend on matching patterns. For example, in the model with heterogeneous preferences leading to the probability of potentially transmissive matches given previously in equation (17), higher activity of low-activity agents confers a positive matching externality, but higher activity of high-activity agents confers a negative externality.[14] Which specification is appropriate depends on the biological nature of the pathogen and the social structure of the society in which it spreads.

Pragmatically, the modeler must carefully consider how matches are formed when specifying the planner's problem. A planner facing pure density dependent matching (equation 19) would want to control only the behavior of infected agents, while they would want to also control the behavior of susceptible agents in the case of quadratic matching (equation 20), or the behavior of all types under linear matching (equation 21). Further, the modeller must be careful to consider whether the planner should be concerned about heterogeneity of contact rates within health classes. For example, if matching is random conditional on contact rates, the force of infection generally depends on the entire distribution of contact rates, not just the mean, so the planner's problem becomes much more complex. It is possible to show that $\mathcal{R}_0$ is approximately proportional to $(\mu_a + \sigma_a^2/\mu_a)$, where μ_a is the mean and σ_a^2 the variance of contact rates (Anderson and May, 1991). Notice that a mean-preserving spread in contact rates spurs the epidemic, because high-activity types disproportionately spread the disease. In this case, the planner must consider the effects of its directives on the higher-order moments of the distribution of activity.

3.2 The Social Planner Faces an Epidemic

The social planner's problem is used to benchmark efficient outcomes, i.e., those in which externalities are internalized in ways that lead to first-best outcomes. It is useful to distinguish two different ways that the literature has modeled the tradeoffs facing a social planner when formulating optimal policies. Both spring from the observation that interventions are costly and disruptive to economic and social activity, and thus suppressing disease incidence or prevalence imposes costs elsewhere in the system. Most of the literature before COVID-19 took an implicit approach and modeled the cost of interventions in an abstract way through direct costs or opportunity costs directly associated with disease mitigation. This implicit approach elucidates the tradeoffs in a clean way, helping to highlight the structural features of the models that lead to different types of results. In much of the post-COVID-19 literature, which has sought to quantify the effects of different interventions such as stay–at–home orders, research relied on an explicit model-

[14] See Kremer and Morcom (1998) for an extensive discussion of this issue.

ing of costs or of macroeconomic impacts. This explicit approach can help quantify the effects of practical policy proposals and can directly guide cost-benefit analysis. Below, we outline a number of models relying on the implicit approach and give an example of a macroeconomic-epidemic model as an exemplar of the explicit approach.

3.3 Preventive Action

Preventing infections is a major goal of public health, and the undersupply of disease mitigation efforts is a driver of many of the disease externalities in Table 2. Preventive actions can be divided into pharmaceutical interventions, such as vaccines, and non-pharmaceutical interventions (NPIs), such as social distancing.

3.3.1 Behavior-Based Interventions

Behavior-based interventions, such as interventions to increase social distancing, are preventive measures that, in their most basic form, do not require any technological intervention. They are the largest class of NPI. NPIs are often important in response to novel pathogens that lack effective vaccines. Behavior-based NPIs are measures that restrict or alter individuals' behavior or movements with a view to reduce the transmission of a pathogen from infectious to susceptible individuals. NPIs include stay–in–place orders that restrict the movement of all people, school closures, targeted quarantines of infectious individuals, or other measures restricting circulation only to some physical spaces or time intervals. NPIs also include other forms of individual shielding such as wearing masks, hand washing, disinfecting, and community–level technical NPIs such as enhanced ventilation. The economic epidemiology literature has focused mainly on NPIs that are imposed at the level of the individual.

The normative question in economic epidemiology is whether people voluntarily adopt or engage in NPIs at socially optimal levels during an epidemic or whether public policy is necessary to maximize welfare.

It has become standard in the literature to model NPIs as costly reductions in potentially transmissive matches between infectious and susceptible individuals. Concretely, by reducing the exposure of susceptible individuals to risky encounters, disease transmission can be reduced, but that reduction comes at a cost because the change in behavior is a deviation from the desired no-infection level that would have been chosen in a world without disease risk. To model such policies, let $\overline{a}_S(t) \in [0,1]$, $\overline{a}_I(t) \in [0,1]$, and $\overline{a}_R(t) \in [0,1]$ denote the activity levels of each susceptible, infected, and recovered individual in the population, where we have normalized the desired activity level in the absence of an epidemic to unity. We assume that all individuals in a health class make the same decisions. This can be generalized by conditioning this assumption on other observable features. In the models that follow, we normalize the population so that $N(t) = 1$ and drop the $E(t)$ compartment to streamline exposition, unless otherwise stated. In addition, we assume that the behavior of recovered people does not influence matches between infected and susceptible individuals. Suppose that matching between people in different health states is such that the force of infection is

$$\beta \overline{a}_S(t) \overline{a}_I(t) I(t). \tag{22}$$

Setting aggregate activity $\overline{a}_S(t) = \overline{a}_I(t) = 1$ yields the standard non-behavioral disease incidence, while setting $\overline{a}_I(t) = 0$ eliminates new infections altogether. It follows from the multiplicative nature of this expression that the rate of new infections can be reduced by scaling back the activity of either infected individuals, $\overline{a}_I(t)$, that of susceptible individuals, $\overline{a}_S(t)$, or both. The first policy is akin to quarantines, the second to shielding (or focused protection) and the last to a generalized lockdown. Notice that this formulation of the transmission function invokes the assumption that the behavior of recovered people has no effect on the rate at which potentially transmissive matches are formed.

To the extent that the social planner can distinguish susceptible individuals from infectious individuals, the planner will differentially control their activity levels over time, at each moment accounting for the costs and benefits of doing so and continually adjusting these levels to changes in the disease classes (Fenichel, 2013). We will discuss these issues in more detail momentarily, but since focus will usually be on whether private protective actions by susceptible people are socially optimal, we begin by considering a social planner that only controls the activity level of susceptible people. Consider the case where only susceptible individuals have any private benefit from disease mitigation and choose their mitigation non-cooperatively. This provides as reasonable benchmark for comparison a setting in which the planner only controls the activity decisions of susceptible individuals.

In the economic treatment of epidemics, individuals have explicit preferences over health states. A simple way to model this is to assume that individuals earn health-state-dependent *flow utilities* π_j, $j = S, I, R$ with $\pi_{\mathcal{S}} \geq \pi_{\mathcal{R}} > \pi_{\mathcal{I}}$. This simply means that individuals prefer to be healthy and that once recovered, they may recover fully or experience permanent after-effects of the disease. In the planner's objective, these flow utilities are simply aggregated across the population in a utilitarian welfare criterion, with each health class weighted by their measure at any given moment. In the models of vaccination and treatment below, these assumptions will be retained without further mention.

To model the tradeoff between reducing transmission and scaling back intrinsically desirable transmissive activity, we introduce a cost of mitigation of the form

$$C(\bar{a}_S(t)) = c(1 - \bar{a}_S(t)) + \frac{k}{2}(1 - \bar{a}_S(t))^2, \tag{23}$$

where $c, k \geq 0$ are parameters. This formulation nests the two most common cost functions considered in the literature, namely linear costs (when $k = 0$) and quadratic costs (when $c = 0$).

With this cost function, we can write the problem of a benevolent utilitarian social planner who wishes to maximize the net present value of welfare across the population as

$$\max_{\bar{a}_S(t)\in[0,1]} \int_0^\infty e^{-\rho t} \left(S(t) \left(\pi_{\mathcal{S}} - c(1 - \bar{a}_S(t)) - \frac{k}{2}(1 - \bar{a}_S(t))^2 \right) + I(t)\pi_{\mathcal{I}} + R(t)\pi_{\mathcal{R}} \right) dt, \tag{24}$$

where $\rho > 0$ is the discount rate. The tradeoff now becomes easy to describe. Decreasing the activity level $\overline{a}_S(t)$ is costly, but it also decreases disease incidence. This changes the composition of the population so there are fewer infected people, with a commensurate decrease in the disease burden. This tradeoff changes across the stages of the epidemic and therefore the optimal activity level changes over time. To see this, assume for simplicity that $\overline{a}_I(t) = 1$ and note that the optimal activity level for susceptible people implemented by the planner satisfies the necessary optimality condition

$$-\lambda(t)\beta I(t) = c + k(1 - \bar{a}_S(t)),$$

where $\lambda(t)$ denotes the shadow cost of additional infected people, i.e., the lost value in discounted, aggregate social welfare from a small increase in disease prevalence. This condition equalizes the marginal benefit of disease mitigation (which is proportional to the force of infection) to its marginal cost.

It turns out that the exact policy that the social planner will wish to implement depends delicately on the nature of the underlying disease dynamics, as they dictate how the shadow cost of infection $\lambda(t)$ evolves over time as a function of the state of the pandemic and the distribution of the population over health classes. Below, we describe how the nature of the disease influences optimal policy.

3.4 Centrally Planned vs Decentralized Behavior

We briefly contrast the problem of the social planner outlined above with the continuous-time formulation of the problem faced by a representative individual who takes aggregate disease dynamics as given. This approach was introduced by Reluga (2010), Reluga and Galvani (2011) and Gersovitz and Hammer (2004) and was subsequently used by Toxvaerd (2019), Farboodi, Jarosch and Shimer (2021) and others.

There are two features of the planner's constrained maximization problem that are worth emphasizing. First, the utilitarian social planner maximizes the aggregate welfare of the population. Second, in optimally setting agents' activity level, $\bar{a}_S(t)$, the planner *explicitly* take into account that its decisions directly influence the path of aggregate disease prevalence. To see the contrast, the problem solved by the individual is given by

$$\max_{a_S(t)\in[0,1]} \int_0^\infty e^{-\rho t} \left(p_{\mathcal{S}}(t) \left(\pi_{\mathcal{S}} - c(1 - a_S(t)) - \frac{k}{2}(1 - a_S(t))^2 \right) + p_{\mathcal{I}}(t)\pi_{\mathcal{I}} + p_{\mathcal{R}}(t)\pi_{\mathcal{R}} \right) dt. \tag{25}$$

In this objective function, the "individual" state variables $p_i(t) \in [0, 1]$ are the probabilities that the individual is in the health state $i = S, I, R$ at instant t. In this formulation, individuals have perfect foresight and know the future path of aggregate variables, but individually maximize their objectives, taking these dynamics as given. Assuming that the underlying disease dynamics are of the SIR type, the evolution of these state probabilities are governed by the laws of motion

$$\dot{p}_{\mathcal{S}}(t) = -a_S(t)\beta I(t) p_{\mathcal{S}}(t), \quad p_{\mathcal{S}}(0) = 1 \tag{26}$$
$$\dot{p}_{\mathcal{I}}(t) = a_S(t)\beta I(t) p_{\mathcal{S}}(t) - \gamma p_{\mathcal{I}}(t) \tag{27}$$
$$\dot{p}_{\mathcal{R}}(t) = \gamma p_{\mathcal{I}}(t). \tag{28}$$

Note that, in these laws of motion, the current state and future path of disease prevalence $I(t)$ is taken as parametrically given by the individual, whose choices do not influence the aggregate. In a perfect foresight equilibrium, the postulated future path of the aggregate system is confirmed and when one aggregates the individual state variables across the population, the laws of motion of the aggregate state variables are recovered (see Rowthorn and Toxvaerd 2012 for details).

We are now in a position to fully account for the difference between the planner's and the representative individual's problem. First, the individual takes the aggregate path of the system as *given* and maximizes their own objective function with respect to what they can control, namely the probability over their own health states. In contrast, the planner directly controls the path of infection, not just that of a given individual. Second, the two objectives are different, accounting as usual for the fact that the individual does not internalize the effects that their decisions have on the welfare of others. The upshot of this is that there are two separate channels through which the equilibrium path under decentralized decision making differs from that preferred by a social planner. The wedge is driven in part by differences in preferences and in part by the extent to which decisions impact the path of aggregate infection. The necessary optimality condition for the representative individual's optimization problem is

$$-\mu(t)\beta I(t) = c + k(1 - a_S(t)),$$

where $\mu(t)$ denotes the shadow cost of a higher probability of infection for this given individual. Direct comparison with the optimality condition of the planner shows that these are similar in nature. But because the planner and the representative individual have different maximization problems, it is generally the case that $\mu(t) \neq \lambda(t)$ and thus the socially optimal policy will differ from what is privately optimal for the individual.

3.5 Targeting and Interactions Between Interventions

When a social planner can tell apart individuals in different health classes, this raises the possibility that individuals in different disease states are instructed to vary their exposure levels to different degrees so that the mitigation efforts of both sides of a match interact to reduce transmission. In settings where the disease status of individuals cannot be readily verified without diagnostic testing, they must all be treated equally and in a manner consistent with beliefs about population averages (see Farboodi, Jarosch and Shimer 2021 and Toxvaerd 2022). In models in which individuals are aware of their own disease status, if infected people are completely self-regarding, then $a_I(t) = 1$, i.e., infected people have no private incentive to restrict activity, short of symptoms, such as feeling lethargic, that lead to changes in behavior, as modeled e.g. in Chen (2012).

To see how the cross-compartment externalities influence optimal policy, consider a simplified model in which a social planner can differentially control the activity level of infected and susceptible individuals and costs are quadratic. The planner's objective is to maximize the following static objective at each point in time, where each contact between infectious and susceptible individuals that leads to infection carries a welfare cost of $(\pi_S - \pi_I)$ to society:

$$-(\pi_S - \pi_I)\beta I(t)S(t)\overline{a}_I(t)\overline{a}_S(t) - S(t)\frac{k}{2}\left(1 - \overline{a}_S(t)\right)^2 - I(t)\frac{k}{2}\left(1 - \overline{a}_I(t)\right)^2. \qquad (29)$$

Clearly, reductions in activity of either group mitigate infection, but also carry costs for each side of the match. The first-order conditions for optimality with respect to $\overline{a}_S(t)$ and

$\overline{a}_I(t)$ are then

$$\overline{a}_I(t)\beta I(t)(\pi_S - \pi_I) = k\left(1 - \overline{a}_S(t)\right) \quad (30)$$
$$\overline{a}_S(t)\beta S(t)(\pi_S - \pi_I) = k\left(1 - \overline{a}_I(t)\right). \quad (31)$$

We see from these first-order conditions, which equalize the marginal benefits and marginal costs of each intervention, that the optimal activity level of infected people depends on the activity level of the susceptible people and vice versa.

Although this objective ignores intertemporal considerations for expositional purposes, we draw a number of insights from this simplified model. Any of several possible policies reduce contacts between susceptible and infected people and thus reduce transmission. The question then becomes, what is the best way of achieving such a reduction in transmissive contacts? Due to the multiplicative nature of this transmission term, the effects are generally asymmetric and depend on the numbers of people who are susceptible or infected. Second, the cost of implementing the differential activity levels itself depends on the number of people in each class. And because these magnitudes change over time, so must the targeting of the restrictions. Of course, with a fully forward-looking policy, additional intertemporal considerations play a role.

3.5.1 Role of Epidemiological Structure

Different epidemiological structures influence the effects of prevention and mitigation. In this section, we discuss the influence of the epidemiological structure on the basic results of preventive action.

We start by considering SI type diseases, which have particularly simple disease dynamics and were amongst the first type of disease to be studied in the economic epidemiology literature (Geoffard and Philipson, 1996). Under SI dynamics in a closed population, the number of infected people is always non-decreasing, because those who become infected neither pass away, recover, or become susceptible to reinfection. This means in particular that infected people remain a perpetual, non-decreasing source of infection risk to remaining susceptible people. From the perspective of at-risk people, the increased infection risk creates an increasing incentive to engage in protective behavior, but the time-path of mitigation will be dictated by how mitigation costs are modeled. In the simple linear cost case, mitigation will jump from zero to full mitigation, while in less stark settings with convex mitigation costs, protection will be ramped up as infection risk increases. In general, equilibrium mitigation will be lower than the level that maximizes social welfare. While externalities from infected to susceptible individuals eventually decrease as few susceptibles remain, the intrinsic value of protecting these individuals remains high.

The ability to recover and return to susceptibility has little qualitative affect on optimal policy. For SIS type diseases, optimal mitigation takes a simple form in the linear cost case (see, e.g. Aadland, Finnoff and Huang 2013; Toxvaerd 2019). In particular, for prevalence above the optimal endemic steady state level, there is full mitigation, forcing disease prevalence down until the steady state is reached. For prevalence levels below the endemic steady state, optimal policy mirrors that for an SI disease, with optimal mitigation set at zero, thereby allowing disease prevalence to increase to the steady state level. Once in steady state, mitigation is set at an interior level to maintain a constant level of infection in the population. In this model, when prevalence is low, the marginal cost of even the first unit of prevention exceeds the marginal benefits, and it is socially optimal to let the infection spread unhindered initially. If, instead, prevention costs are convex, then mitigation is increased or decreased smoothly as prevalence changes. Under decentralized decision making, the equilibrium path mirrors the socially optimal dynamics qualitatively, but there is suboptimally low private protective behavior in steady state, and accordingly too much infection in the population.

Now consider SIR and other epidemics in which individuals can recover with immunity. This necessarily introduces an additional state variable into the problem. If the setup is linear in the control variable, the solution is also of the bang-bang type. The infection initially spreads unhindered until a threshold is reached (see, e.g., Toxvaerd 2020), at which point individuals maintain an increased intensity of mitigation over a certain interval of time. During this interval, disease incidence (new cases per unit time) is suppressed until enough herd immunity accumulates in the population, and once sufficiently many people have recovered, mitigation measures are lifted. After this point, infection does not increase, even if the mitigation measures are discontinued, but instead slowly declines of its own accord. Equilibrium mitigation is in place only until the point that prevalence naturally declines, a point that is eventually reached in any SIR setting without births.

Turning to the case of convex mitigation costs, adjustment is more gradual and the path of disease incidence is more fine-tuned than in the case with linear costs. Mitigation follows disease prevalence and acts to both decrease peak prevalence and push some infections into the future. Overall, mitigation reduces the aggregate number of infections across the epidemic and does not simply postpone them. Optimal mitigation policy in the SIR model yields disease dynamics similar to those under decentralized decision making, but with more mitigation. See e.g. Makris and Toxvaerd (2020) for a direct comparison under convex mitigation costs.

In the SIRS model, in which immunity is not permanent, preventive action also increases and decreases with disease prevalence, as is the case in the SIR model. In contrast to the case with perfect immunity, when immunity protection wanes, the disease can potentially become endemic. In this case, it may be optimal to control infection even in steady state. Giannitsarou, Kissler and Toxvaerd (2021) show that in this setting, mitigation levels can exhibit dampened oscillations, tracking this pattern in the underlying disease dynamics. Similar results can occur when births and deaths are introduced into SI or SIR models, because then the S compartment is continually replenished (Fenichel and Horan, 2007). There is some herd immunity created by temporary individual immunity in the SIRS model, so it may be optimal to eradicate the disease in a steady state. This contrasts with mitigation in the SIS model without immunity (Rowthorn and Toxvaerd, 2012), where eradication is generally not optimal. To eradicate the disease in the SIS model, full mitigation must be sustained in perpetuity, which violates a transversality condition, unless there is a large, discrete lump sum benefit to eradication (Barrett and Hoel, 2007). In the SIRS model, the system can tip over to a path towards asymptotic eradication even with less than full mitigation. Also, mitigation may be more valuable in the SIRS model than in the SIR model. The reason is that if there is disease-induced mortality, under SIRS dynamics individuals face repeated risk of death over time, and there is therefore added value from keeping disease numbers low.

3.5.2 Vaccines

Vaccines are a technology that can prevent infection, reduce infection risk, reduce disease severity, or reduce the ability to become infectious. The economic epidemiology literature has focused almost exclusively on vaccines that either partially or fully prevent disease transmission.[15] When a vaccine completely protects against infection, it is said to yield *sterilizing immunity*. In this case, by vaccinating oneself and thereby reducing one's risk of becoming infected and infectious, there is a strong positive effect on other individuals and a potential for uninternalized positive externalities. These have been the central focus of normative economic analysis and the other roles of vaccination have received far less attention in economic epidemiology. We take up the case of partially effective vaccines, for which behavioral responses may further complicate policy analysis, in the following section.

It is simplest to start by considering vaccines that create sterilizing immunity. To model disease dynamics with vaccine-induced immunity, it is natural to make use of the classical SIR model, in which some individuals are protected from infection. Vaccination moves individuals from the susceptible compartment directly to the recovered (and thus immune) compartment. The approach can be extended by adding more compartments in order to model more realistic vaccination that provides imperfect or waning protection.

When vaccine acquisition and delivery are costly, optimal rollout schedules call for a comparison of these costs with the benefits of vaccination, which the epidemiological literature has mostly couched in terms of the disease burden. The benefits of vaccination typically change across the stages of the epidemic and so vaccination policy is necessarily forward-looking and state-dependent. Because the build-up of population immunity creates the potential for non-monotonicities in the underlying biological dynamics, the economic benefits from vaccination (and hence, from the creation of costly increases in population immunity) depend delicately on the nature of the disease and on whether individuals spontaneously recover with immunity without the aid of vaccines.

Francis (1997) introduced the canonical dynamic model of vaccination. In this model, once individuals become infectious, they remain so in perpetuity, so the model retains the SIR structure except in that the only path to the recovered state is through vaccination. In this exposition, we follow the formalization in Toxvaerd and Rowthorn (2022), but other closely related formalizations are possible and the qualitative results remain the same.

Denote by $v(t) \in [0, 1]$ the rate at which individuals are vaccinated. Vaccination at rate $v(t)$ induces transition from $S(t)$ to $R(t)$ at rate $\alpha_V v(t)$, entirely evading infection. The parameter $\alpha_V > 0$ can be interpreted as the speed at which immunity becomes effective. Vaccination costs $c_V > 0$ per instant per individual. The central tradeoff in this model is that while vaccines are costly to administer, they boost the build-up of population immunity over and above the immunity that would be created naturally through infection and subsequent recovery (if any).

[15]Models of vaccination were amongst the first to be analyzed in the field of economic epidemiology. See e.g. Sethi and Staats (1978) for treatment and vaccination in the SIS setting, Morton and Wickwire (1974) and Bobisud (1977) in the SIR setting. See Chen and Toxvaerd (2014) for a review of early work on the economics of vaccines.

The planner's problem is then

$$\max_{v(t)\in[0,1]} \int_0^\infty e^{-\rho t}[S(t)(\pi_{\mathcal{S}} - v(t)c_V) + I(t)\pi_{\mathcal{I}} + R(t)\pi_{\mathcal{R}}]dt, \tag{32}$$

subject to the constraints,

$$\dot{S}(t) = -S(t)\left[\beta I(t) + \alpha_V v(t)\right] \tag{33}$$

$$\dot{I}(t) = I(t)\left[\beta S(t) - \gamma\right] \tag{34}$$

$$\dot{R}(t) = \gamma I(t) + S(t)\alpha_V v(t). \tag{35}$$

This problem is linear in the control variable, $v(t)$, and therefore the solution is of the bang-bang variety, where vaccination is set to either zero or to the maximum possible level. This means that if marginal vaccination costs are constant (that is, do not change in the rate of vaccination), increasing vaccine benefits induce at most one switch in optimal policy. Francis (1997) determines that there is an optimal point in time when all remaining susceptible individuals are vaccinated as quickly as possible. In his model, people do not recover spontaneously, and so infected people remain a permanent source of infection. Furthermore, the population is homogeneous, which means that the risk of infection for unvaccinated people increases over time and is the same for all. People remain unvaccinated until a risk threshold is reached, at which point any remaining unvaccinated individuals simultaneously get vaccinated. If vaccination confers perfect and instant protection, then there are no immediate externalities and the decentralized vaccination path is socially optimal.

This striking result, derived by Francis (1997), does not generalize (Chen and Toxvaerd, 2014; Fine, Eames and Heymann, 2011). If only some people are vaccinated, because the population is heterogeneous and some people therefore vaccinate before others, then some people will be exposed to infection risk and benefit from the indirect protection of vaccinated individuals. In this case, decentralized vaccination decisions are no longer socially optimal. This is also the case when vaccines offer imperfect protection against infection, because then even vaccinated people are at (reduced) risk of infection and thus benefit from the vaccination of others.

Now consider a disease for which infectious individuals can recover and gain natural immunity. In the absence of vaccination, this model is of the SIR type, for which we know that population immunity eventually protects remaining susceptible individuals as the disease dies out in the population. This has implications for optimal vaccination policy. Economically, the possibility of spontaneous recovery with immunity makes the benefits of vaccination hump-shaped, with an initial increase and a subsequent decrease as the epidemic progresses. This creates the potential for up to two switches in optimal vaccination policy. It may be optimal to initially delay vaccination until a threshold of people have become infected and then start vaccination at the maximum possible rate. If this rate is finite, for example, because of lack of vaccines or because of constraints on the capacity to deliver them, then not all susceptible people are immunized at once and the vaccination rollout must be sustained over time. But if aggregate immunity builds up sufficiently fast, because of the combination of natural and induced immunity, a point may be reached where it is optimal to cease further vaccination. Another possibility is that the optimal policy first vaccinates at the maximum rate and then switches to no vaccination as population immunity becomes significant (Toxvaerd and Rowthorn, 2022).

3.5.3 Interactions Between Vaccination and Behavior

Since vaccines affect the tradeoffs people face between health and other goals, economists expect vaccination to affect behavior, which will in turn affect the spread of disease. For example, the agent in the model presented in Section 2.2.1 chooses to take on higher risk of infection in response to a decrease in risk per contact, such as induced by a partially effective vaccine, if her behavior is prevalence elastic (see equation 14). Then vaccination generally has external effects, but even their sign is ambiguous. This externality is a special case of the matching externality. In almost all past economic epidemic work on vaccines, the vaccination decision itself is the only choice over which individuals or the social planner maximize. However, as we have seen, the effects of vaccination may influence the incentives to engage in other mitigation efforts and so it is important to understand how such decisions interact.

Here, we focus on how vaccination may influence social distancing during the rollout, when only some individuals are vaccinated but vaccines only confer limited protection against infection. There are multiple interacting effects of such vaccination, through their effects on behavior, disease dynamics, and virulence. Suppose that the social planner rolls out a vaccine such that vaccinated individuals (i) become less susceptible to infection (inward mitigation), (ii) become less infectious to third parties through reduced shedding (outward mitigation), (iii) are infectious for shorter time (the infection clears faster), and (iv) the disease becomes less virulent, that is, vaccinated individuals, if they become infected and ill, have lower disease burden and suffer smaller utility loss. These four properties have direct and indirect effects on individuals' incentives to self-protect and, similarly, would influence the design of optimal lockdown and vaccine rollout policies.

This setup affects incentives in multiple ways. All else equal, vaccinated individuals are less at risk than non-vaccinated ones and, as such, have reduced incentive to engage in other mitigating activities, such as social distancing. On the other hand, non-vaccinated people now also face a different risk environment than before vaccination was rolled out, because some of the infected people they are likely to meet have reduced infectiousness as they are now vaccinated. This in turn reduces the incentives for susceptible non-vaccinated individuals to engage in costly mitigation. Kremer (1996) noted some of these possible effects and raised the possibility that the combined impact of imperfect vaccines and voluntary behavioral change could lead to detrimental effects on welfare. Gans (2023) embeds behavioral responses in a model otherwise similar to Francis (1997). Avery, Chen and McAdams (2024) consider how vaccination and social distancing interact, emphasizing that policies that encourage one form of protective behavior might blunt incentives to engage in other forms of protective behavior. These changes in incentives also have general equilibrium effects, because they influence the future path of aggregate behavior, and thus disease paths that influence present and future behavior.

Finally, we note that, to the extent that vaccination reduces the virulence of the disease, it alters the tradeoffs faced by the individual, with *ex ante* mitigation becoming less valuable, *ceteris paribus*. This effect on incentives mirrors the effect of decreased susceptibility due to imperfect vaccines, as discussed in Section 2.2.1.

3.6 Treatment

Once individuals become infected, it is sometimes possible to provide treatment with therapeutics. Treatment affects the welfare of the treated by changing the direct cost of being in the infected and infectious compartment. It can also reduce the ability of infectious individuals to infect others by reducing the production of pathogen or reducing the time that the individual is infectious. These features make therapeutics a potentially effective instrument for controlling infectious diseases, at individual and aggregate levels. Treatment has received less attention in the literature than vaccines and NPIs, yet for some diseases, such as syphilis, treatment is an important tool in addition to behavioral interventions (Aadland, Finnoff and Huang, 2013). Treatments have several effects. Depending on the disease being modeled, treatment can induce recovery back to the susceptible state for diseases without immunity, such as many bacterial and fungal infections, which are best modeled as SIS diseases. Treatment can also accelerate movement to the recovered and immune state, as in classical SIR type diseases. Because the controlled dynamics differ significantly across these two scenarios, we address them separately. For completeness, we note that the economic literature on treatment has largely considered the choice of treatment itself as the object of analysis, rather than how this decision interacts with possible mitigation efforts that the individual or the social planner may engage in (although see Makris and Toxvaerd 2020 for an exception). As we have seen in the positive economic analysis, the availability of treatment may reduce the ex ante incentives to engage in costly protective behavior ex ante.

3.6.1 Susceptible-Infected-Recovered

A model of treatment with perfect and permanent immunity for all recovered individuals (naturally or through treatment) can be characterized with a modified SIR model. Treatment is assumed to be costly, but people are taken from the infected state to the recovered state at an increased rate. There is a substantial literature within health economics on the demand for health care and for health insurance, but to our knowledge, there is no significant literature on private demand for treatment in the economic epidemiology literature, perhaps because medical treatment has been viewed in large part as a static and private problem.

Consider the model of Toxvaerd and Rowthorn (2022). They assume that for some treatment intensity $\tau(t) \in [0,1]$, the rate at which individuals transitions from $I(t)$ to $R(t)$ 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.[16] This modeling approach is similar to the way that vaccines are modeled, transferring people from the susceptible class to the recovered class. When considering the efficient level of treatment, the social planner takes into account the aggregate flow utility of individuals in the population. In addition, the planner factors in the direct costs of treatment, captured by the term $\tau(t)c_T$. This is the cost of increasing the treatment rate $\tau(t)$, multiplied by the cost of treatment per instant per treated individual, c_T.

The social planner solves

$$\max_{\tau(t)\in[0,1]} \int_0^\infty e^{-\rho t}[S(t)\pi_{\mathcal{S}} + I(t)\left(\pi_{\mathcal{I}} - \tau(t)c_T\right) + R(t)\pi_{\mathcal{R}}]dt \tag{36}$$

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

subject to

$$
\begin{aligned}
\dot{S}(t) &= -\beta I(t)S(t) && (37)\\
\dot{I}(t) &= I(t)\left[\beta S(t) - \alpha_T \tau(t) - \gamma\right] && (38)\\
\dot{R}(t) &= I(t)\left[\alpha_T \tau(t) + \gamma\right]. && (39)
\end{aligned}
$$

The value of treatment for the planner in this setting is best understood in terms of private and social effects. The private effect of treatment is that infected people experience more rapid recovery. This value is directly experienced by those undergoing treatment and, on this count, the individuals themselves are expected to value treatment as much as the social planner. The social effect is an additional benefit from treatment because it helps increase the number of recovered people, just as natural recovery does. In this sense, treatment has some features in common with vaccination in that it increases the number of recovered people. However, while vaccines take susceptible people and move them to the immune state (reducing infection and intertemporal externalities), treatment instead takes infected people and speeds up their recovery, thereby directly reducing disease prevalence (and potentially reducing matching externalities).

The wedge between the social and private value of treatment decreases throughout the pandemic and is tied to the number of remaining susceptibles. Heuristically, the measure of susceptibles in the population is a proxy for the magnitude of the positive external effects of treatment, because treatment of an individual benefits neither other infected individuals nor those who have already recovered. The upshot is that, in a socially optimal treatment policy, there can be at most one shift in regime from full treatment to no treatment. Except for uninteresting cases where the intrinsic effect does not outweigh the cost of treatment, the typical scenario is one with full treatment at early stages (when there are many susceptible people that benefit from the external effects), followed by a complete cessation of treatment, when the external effects of further intervention have eroded to become lower than treatment costs. Note that this limits attention to the most interesting case in which the cost of treatment outweighs the intrinsic, private, benefits of recovery, for otherwise it would be trivially optimal to always treat everyone who is infected immediately irrespective of the stage of the epidemic.

Under decentralized decision making, equilibrium treatment decisions for a given individual are disconnected from the current or future state of the epidemic. The infected person only compares the cost of treatment with the private net present value of recovering. Thus, equilibrium treatment is different in kind, and not just in degree, from the socially optimal treatment policy.

3.6.2 Susceptible-Infected-Susceptible

Treatment without immunity was first studied by Sanders (1971), with subsequent analysis that merges treatment and immunity by Lightwood and Goldman (1995), Sethi (1974), Goldman and Lightwood (2002), Gersovitz and Hammer (2004), and Rowthorn and Toxvaerd (2012). Using a simplified version of Toxvaerd (2009), the model of treatment without immunity is a modified SIS type disease, given by the following maximization problem for the social planner:

$$
\max_{\tau(t)\in[0,1]} \int_0^\infty e^{-\rho t}\left[\, I(t)(\pi_{\mathcal{I}} - \tau(t)c_T) + S(t)\pi_{\mathcal{S}}\right] dt, \tag{40}
$$

subject to,

$$\dot{S}(t) = -I(t)\left[\beta S(t) - \gamma - \alpha_T \tau(t)\right] \tag{41}$$
$$\dot{I}(t) = I(t)\left[\beta S(t) - \gamma - \alpha_T \tau(t)\right]. \tag{42}$$

Because $S(t) + I(t) = 1$ by assumption, the dynamics can be reduced to a single law of motion for disease prevalence.

This model is unusual in the economic epidemiology literature because the non-convexities of the model can yield multiple equilibria (under decentralized decision making) and path dependence of the optimal policy. This model shares many properties with so-called shallow lake systems, known from environmental economics (see Maler, Xepapadeas and De Zeeuw, 2003). The central driving force in this model is the existence of dynamic complementarities in the sense that more treatment in the present brings about conditions in the future that make further treatment even more desirable. Similarly, less treatment in the present makes future treatment less valuable, thereby reinforcing the lack of treatment. The value of treating an infected individual is that the individual recovers faster and, for some time, remains healthy until possible reinfection. However, the future probability of reinfection is a function of the future path of disease prevalence, which is a function of treatment.

Consider a policy of aggressive treatment of every infected person. Under this policy, disease incidence is suppressed, and in future periods there will be fewer infected people around who can (re)infect the susceptible. This means that the value of treatment increases, thereby creating a virtuous cycle that leads to a high treatment/low prevalence steady state. In contrast, under a policy of no treatment, infection increases over time, pushing up reinfection probabilities and thereby eroding the value of treatment. This vicious cycle leads to a low treatment/high prevalence steady state. In contrast to many other economic-epidemic models, this setting can create path dependence in the sense that the optimal path depends on initial conditions. Specifically, for sufficiently low initial disease prevalence, it is optimal to treat at the maximum possible level and thereby eradicate the disease. In contrast, for sufficiently high initial disease prevalence, it is optimal to do nothing and let the disease become endemic.

Under decentralized decision making the equilibrium path is qualitatively similar to the planner's path, but is not socially optimal. In addition, dynamic complementarities may take the form of strategic complementarities, raising the possibility of multiple self-confirming equilibria for a range of initial conditions (see Toxvaerd, 2009).

3.7 Interactions Between Economic Activity and Epidemics

The COVID-19 pandemic prompted an important extension of the economic epidemiology literature to address interactions between the epidemic and the macroeconomy. Work on the HIV+/AIDS epidemic largely focused on risky sexual behavior, which did not have substantial macroeconomic effects in most countries. However, morbidity and mortality due to HIV+/AIDS in, in particular, sub-Saharan Africa generated a literature on how economic growth is affected by disease.[17] Yet, scant attention was paid to higher-frequency

[17]See for example Bloom and Mahal (1997) on HIV disease in Africa, Acemoglu and Johnson (2007) and Chakraborty, Papageorgiou and Sebastian (2010) on disease and development more generally, or Jedwab, Johnson and Koyama (2022) on the macroeconomic effects of the Black Death. For a review of

effects of disease on the aggregate economy, such as the extent to which an epidemic might cause a recession. Private and public responses to COVID-19 involved large shocks to economic activity, and the literature rapidly adapted to address the resulting issues.

The costs of mitigation and the shortfall in utility caused by infection are abstract ways to capture the tradeoffs between mortality and morbidity caused by infection and the socioeconomic disruption brought about by disease mitigation. The advantage of this abstract formulation is that it allows the modeler to cleanly disentangle the effects resulting from the interaction between the disease dynamics and the different mitigation measures, mediated by the objective function that dictates how tradeoffs are resolved (Farboodi, Jarosch and Shimer, 2021). When the aim is to understand macroeconomic effects of epidemics, a fully developed economic environment is warranted, and in such cases, an integrated macroeconomic-epidemiological framework is needed.

The macroeconomic-epidemic literature features two major additions to the literature. First, social contacts are delineated according to whether they are fiscally relevant interactions, such as those at a workplace or a shopping mall, or social interactions which are not fiscal transactions.[18] Only changes in fiscal contacts affect the macroeconomy. Second, drawing on conventional macroeconomic modeling approaches, these models feature additional dynamic constraints that keep track of state variables in the macroeconomic model. These augmented models are complicated, featuring intertwined economic and epidemic dynamics. Despite these complications, this literature holds significant promise going forward, especially for quantitative and practical policy analysis during epidemics.

We briefly sketch the model of Eichenbaum, Rebelo and Trabandt (2021). Consider a representative consumer who chooses a consumption and labor sequence $\{c_t, n_t\}$ in order to maximize discounted lifetime utility,

$$\sum_{t=0}^{\infty} \psi^t U(c_t, n_t), \tag{43}$$

subject to the budget constraint

$$(1 + \mu_t)c_t = w_t n_t + \Gamma_t. \tag{44}$$

Here, ψ is the discount factor, μ_t is a consumption tax, Γ_t a lump-sum transfer, and w_t is the real wage rate. The model is closed with a competitive, constant returns supply side. In the absence of disease dynamics, the agent's problem is a standard intertemporal maximization problem whose solution is characterized by an intratemporal equation equalizing the returns to consumption and labor and an intertemporal Euler equation relating consumption at consecutive dates. Eichenbaum, Rebelo and Trabandt (2021) introduce disease dynamics by supposing that an SIR disease is spreading in the economy. In addition to becoming infected through non-economic interactions, agents in this economy may become infected while shopping or working. Under random mixing assumptions for shopping and work interactions, and omitting time indexing for clarity, the number of

the macroeconomics of epidemics following COVID-19, see Bloom, Kuhn and Prettner (2022). See also Goenka, Liu and Nguyen (2014) for an early contribution linking epidemics and economic growth.

[18]We use the idea of fiscal transactions, which some may call "economic," rather then economic because welfare enhancing interactions are also economic in the sense of welfare economics.

new infections at any time can be expressed as

$$\beta_1(SC^S)(IC^I) + \beta_2(SN^S)(IN^I) + \beta_3 SI, \tag{45}$$

where C^j denotes the consumption expenditures and N^j the work hours of each agent in health compartment j. Here, $\beta_1(SC^S)(IC^I)$ is the number of susceptibles who become infected while shopping, $\beta_2(SN^S)(IN^I)$ is the number who become infected while working, and $\beta_3 IS$ is the number who become infected through non-fiscal interactions. Compare with equation (8) to see that the model nests the special case of the canonical SIR model if $\beta_1 = \beta_2 = 0$. The authors assume that infectious individuals die with some probability, thereby bearing an opportunity cost of not recovering to a healthy and economically productive state. This creates the basic tradeoff in the model and individuals must balance the utility of economic activity with increased risk of infection and death.

Eichenbaum, Rebelo and Trabandt (2021) solve for the competitive equilibrium of this model and compare it to the disease-free benchmark. In addition, they consider the dynamics when a benevolent social planner chooses the consumption tax sequence $\{\mu_t\}_{t=0}^{\infty}$ to maximize social welfare, interpreting the tax as a containment measure since it reduces infections that occur through fiscal activity, and compare the outcome with the dynamics under direct social control of consumption and labor choices.

3.8 Implementation and Second-Best Policies

The analysis of unrestricted policy design is useful to uncover the working of different models, including welfare tradeoffs that impact different interventions on disease dynamics and individual behavior. Yet in practice, policy can be constrained in a number of important ways related to pre-existing market distortions that may significantly impact how and whether different considerations are included in the policy making process. This requires 'second-best' analysis (Lipsey and Lancaster, 1956). These pre-existing distortions and constraints include the ability to properly restrict policy interventions to the intended targets of the measures. Relatedly, simply characterizing an optimal mitigation policy does not mean that such policies can be implemented without additional measures such as (i) direct restrictions on people's behavior or (ii) other measures such as taxes or subsidies that align individual incentives with social goals.

Once behavior and incentives are taken explicitly into account, it is not sufficient to just characterize the policy or policies that would be socially desirable. Unless individuals are directly restricted in their choices or otherwise incentivized to act according to the prescriptions of health authorities, there is no reason to expect them to play their part and do what the planner finds desirable. Therefore, one needs to explicitly consider the implementation of health policies and to ensure that the outcomes which authorities wish to implement are consistent with the behavior that individuals find privately optimal or feasible. Authorities may help ensure compliance by either directly restricting behavior or by influencing payoffs through taxes and subsidies to specific health states or by influencing state transitions. Rowthorn and Toxvaerd (2012) show how to design optimal tax/subsidy schemes that implement first-best outcomes. These schemes are complicated objects, as the subsidies on offer to a given individual at any given point in time are not only a function of the health state of the individual, but also a function of the aggregate state of the system. This means that optimal subsidies change across the stages of the epidemic.

Once health authorities find themselves in a world of second-best policy making, a sensible check on policy is to ensure that it is not counterproductive, however well-intentioned the policy may be, as policies may backfire and have unintended consequences. An example of this is provided in Toxvaerd (2019), who shows that a policy of rolling out pre-exposure prophylaxis to reduce disease transmission may actually prompt individuals to increase risky behavior and actually reduce overall social welfare.

4 Important Economic Epidemiology Topics Without Microfoundations

In this review, we have largely focused on the integration of intentional behavior into epidemiological models. Yet, there are other topics that are important in the broad area of economic epidemiology that do not fit neatly into this framework. Here, we briefly review some of those issues: Mutation, antimicrobial resistance, and testing. While analysis of any of these could have behavioral components, to the best of our knowledge, these have yet to be fully developed.

4.1 Mutation and Anti-Microbial Resistance

Most models of disease dynamics and pharmaceutical interventions, and the economic-epidemic models based on them, assume that the effectiveness of the intervention remains unchanged across the planning horizon under consideration. In practice, vaccines and treatments may lose their effect over time as viral and bacterial infections change due to mutation. Antimicrobial resistance, where antibiotic treatment induces selective pressure on sensitive strains to the advantage of resistant strains, threatens to render ineffective some of the most important treatments in modern medicine. The formal economic analysis of such changes is still in its infancy. As became clear during the COVID-19 pandemic, vaccines that initially offered some protection against infection may later provide less protection in slowing transmission (although they remained useful in reducing virulence). A reason is that the viruses change over time, reducing vaccine efficacy and creating an arms race: Vaccine development needs to keep pace with the evolving virus.

The economic and welfare losses due to antimicrobial resistance have long been recognized and seem certain to grow larger with time.[19] The central economic issue with antimicrobial resistance is that antimicrobials have features akin to exhaustible resources in that the more they are used, the less effective they become. This issue implies that the use of antimicrobial drugs must be carefully stewarded and used over time and in combination with other drugs in a judicious manner. Laxminarayan and Brown (2001) consider a formal economic-epidemic model of treatment in an abehavioral SIS model with two strains and two treatments, whereas McAdams (2017) and McAdams et al. (2019) extend this type of analysis to consider the effects of rapid diagnostics that detect possible resistance, finding that with diagnostic tests at hand, the exhaustible resource nature of antimicrobial resistance may be partially reversed. Antimicrobial resistance is different from the other models we have considered in this paper because the main diseases in question are non-communicable (i.e. there is little person-to-person transmission) and dynamics

[19]See e.g. O'Neill (2016), and also Dall'Antonia et al. (2005) on bacterial infections and Lipsitch et al. (2007) on viral infections. Laxminarayan and Herrmann (2015) provides a review of the literature. Poudel et al. (2023) provide a meta-analysis of estimates of medical costs due to antibiotic resistance.

mostly pertain to the competition between strains of viruses. However, given the enormous health and economic consequences of antimicrobial resistance, the economic analysis is likely to grow significantly in the coming years, as will health policy and prescription patterns (Dubois and Gokkoca 2023).

In seeking treatment with antimicrobials, individuals may impose an externality on future patients if the drugs become less effective with increased use. This creates a role for public intervention, as the social planner would seek to more carefully steward the use of antimicrobials and stem the rise of resistance.

4.2 Diagnostic Testing

For pathogens in which the infection status of individuals is not *a priori* clear, diagnostic tests form a cornerstone of policies to manage diseases, whether or not infection is communicable. Except perhaps for reasons of purely scientific research interest, diagnostic tests are usually carried out when the results of the tests can be acted upon, and it is from the consequences of these actions that tests derive their value. In other words, the value of tests is the value of the information that they yield. Boozer and Philipson (2000, 1996), provide a simple model illustrating this argument. The field of medical decision making is broad and the testing literature has been extended in a number of different directions.[20] Philipson and Posner (1993, 1995) provide an extensive theoretical discussion of issues surrounding testing of couples in the context of HIV/AIDS, while Gersovitz (2011) also discusses practical issues surrounding the administration of tests and disease control.

Testing can be costly and time consuming. Building on an idea proposed by Dorfman (1943), for example, Lakdawalla et al. (2020), Gollier (2020), and many others proposed testing for COVID-19 infection in large populations by first testing groups, then if at least one member of a group tests positive, test every member of that group. The idea is that group tests of P that return a negative test result rule out P people with one test, which reduces costs. This led to a blossoming literature on test design in 2020 and 2021. However, much of that literature ignored the practical challenges faced by most laboratories, including legal, machinery, and timing barriers to pooled testing (Fenichel et al., 2021).

The value to an individual of diagnostic testing lies in the possibility that better information can cause them to change behavior. But since behavior can have external effects on others, the testing decision itself involves external effects, which potentially drives a wedge between equilibrium dynamics and socially optimal ones. Toxvaerd (2022) shows how mass testing can cause disease incidence to increase, thereby influencing aggregate disease dynamics and welfare.

5 Conclusions

The process of science is often a process of discovery and re-discovery. Although economists have been actively working on infectious diseases for at least half a century, it is perhaps understandable that the dramatic events that unfolded during the COVID-19 epidemic will be the lens through which many economists view the field of economic epidemiology.

[20]See detailed exposition in Felder and Mayrhofer (2022) for static analysis and Toxvaerd (2022) for a review of more recent literature applying these concepts to explicitly dynamic economic-epidemic models.

To some extent, economists prompted to study epidemiological issues by the COVID-19 epidemic did not avail themselves of the existing literature, and one of our goals in this survey is to highlight important ideas from the pre-COVID-19 economic epidemiology literature that may have been overlooked. The epidemic also led to a significant increase in work by economists and to substantial interactions between economists and non-economists.[21] Having outlined the central themes and issues of the field, we conclude the paper with a discussion of how economic epidemiology fared and expanded when confronted with COVID-19, and some suggestions for promising directions that this literature may take going forward.

The COVID-19 pandemic raised the profile of what economists have to offer public policy and science related to the spread of infectious diseases and public health. Many of these contributions draw on work in fields parallel to economic epidemiology, such as methods to value morbidity and mortality, drawn from health and welfare economics. Integrating these ideas with insights from economic epidemiology produced new ideas and policy insights. A majority of economists appear to have supported at least some form of public health interventions at different stages of the pandemic, based in part on their understanding of the economic benefits of active disease management.[22]

COVID-19 generated tremendous and unprecedented interdisciplinary efforts to rapidly make available international data on health, behavioral, and policy outcomes. High frequency national and, in many countries, sub–national data on COVID-19 cases, hospitalizations, and deaths were available by mid-2020, such as through the World Health Organization's COVID-19 dashboard, allowing researchers to track the pandemic across regions in near real–time. Detailed data on policy responses, presented as "stringency indices" summarizing overall responses and broken down by each type of policy, were collated and made available to the research community, such as the Oxford COVID-19 Government Response Tracker (OxCGRT), which reported data on 19 policy areas across more than 180 countries (Hale et al., 2021). Data on social mobility inferred from cell phone locations, such as the Google Community Mobility Reports, allowed researchers to observe behavioral changes over time (e.g. Hu et al. 2021). Armed with these various rich data sources, econometric methods could quickly be brought to bear to study endogenous behavioral responses and the effects of various policies. This work confirmed the predictions of economic epidemiological models that substantial behavioral response would occur even in the absence of policy interventions, and attempted to measure the intended and unintended consequences of policy interventions as they were introduced. As the pandemic progressed, econometric policy evaluation was particularly important due to the almost complete lack of randomized controlled trials on non-pharmaceutical interventions and the limited utility of clinical evidence in the presence of noncompliance and general equilibrium effects. The empirical literature evaluating policy responses to COVID-19 is large, but because the policies were not enacted randomly and were highly correlated, we have few firm conclusions in hand regarding policy efficacy. Given the acrimony surrounding recent vaccination campaigns, mask mandates, and lockdowns, it is incumbent upon researchers to continue to develop methods suitable for credible near real-time policy analysis.

[21]See Murray (2020) for an epidemiologist's perspective on these interactions, Darden et al. (2022) for an interdisciplinary perspective, and Dangerfield et al. (2022) for a perspective on the challenges of these interactions from economists.

[22]See for example Quiggin and Holden (2021).

Even limiting attention to economic processes, the economic epidemiology literature at the beginning of the COVID-19 epidemic had little to say about the macroeonomic effects of the pandemic or appropriate macroeconomic policy responses. The COVID-19 pandemic revealed that microeconomic behavior coupled with policy response can produce sizable aggregate effects, and macroeconomists were quick to draw on the existing literature to develop theoretical and empirical models in which the macroeconomy and the epidemic interact. The lesson here is that these ideas, and possible tools, are likely to have a much wider set of important applications than only within modern macroeconomics.

Although the economic epidemiology literature expanded greatly in response to COVID-19, fundamental issues still have received inadequate attention. One such issue is that a building block of most existing work in economic epidemiology is the notion of conditionally homogeneous mixing, an assumption that simplifies analysis but is not descriptively accurate. This state of affairs is odd in the sense that economists have heavily emphasized that how many contacts people have is endogenous to the epidemic, but have largely ignored the matter of who exactly those contacts are with. To see why this issue matters, consider the opposite extreme of perfectly assortive mixing: If contacts only occur within compartments, then new infections would immediately cease as there would be no potentially transmissive contacts between susceptible and infected people. In the realistic case between these extremes, susceptible people may endogenously exert effort to decrease the proportion of infected people with whom they come into contact, which in turn affects spread of disease and policy efficacy, but this form of endogeneity has received little scrutiny. Furthermore, even assuming away endogenous efforts to avoid matching with infected people, the economic epidemiology literature has largely focused on the case of conditionally random matching, even though both are simplifications. Finding analytically convenient ways to model more realistic structures and still integrate intentional behavior seems worthwhile and an important avenue for future research.

A practical issue that policy makers had to address during the COVID-19 pandemic was the underlying interdependencies between sectors and types of economic and social activity. For example, closing schools might have the unintended consequence of increasing cases because health care workers may then reduce the labor supply to increase child care (see Bayham and Fenichel 2020). Future research should focus on achieving a better understanding of such interactions. Coupled with an understanding of the most common pathways of disease spread, this type of knowledge would enable policy makers to design more targeted, less disruptive interventions.

In our choice of models and coverage of results, we emphasized broad themes and went into some detail on different modeling approaches and structural properties of the models. We have not done justice to certain parts of the literature, most notably the enormous post-COVID-19 literature, with its many innovations in macroeconomic-epidemic modeling and many important empirical studies of voluntary behavior and policy effectiveness. Our main aim in synthesizing the long arc of the economic epidemic literature is to make it accessible to economists interested in the economics of infectious diseases. We hope that this exposition will also serve researchers in other fields and will showcase the distinct contributions that economics has brought to the field of epidemiology.

References

Aadland, David, David C Finnoff, and Kevin XD Huang. 2013. "Syphilis cycles." *The BE Journal of Economic Analysis & Policy*, 13(1): 297–348.

Abaluck, Jason, Laura H. Kwong, Ashley Styczynski, Ashraful Haque, Md. Alamgir Kabir, Ellen Bates-Jefferys, Emily Crawford, Jade Benjamin-Chung, Shabib Raihan, Shadman Rahman, Salim Benhachmi, Neeti Zaman Bintee, Peter J. Winch, Maqsud Hossain, Hasan Mahmud Reza, Abdullah All Jaber, Shawkee Gulshan Momen, Aura Rahman, Faika Laz Banti, Tahrima Saiha Huq, Stephen P. Luby, and Ahmed Mushfiq Mobarak. 2022. "Impact of community masking on COVID-19: A cluster-randomized trial in Bangladesh." *Science*, 375(6577): eabi9069.

Acemoglu, Daron, and Simon Johnson. 2007. "Disease and development: the effect of life expectancy on economic growth." *Journal of Political Economy*, 115(6): 925–985.

Acemoglu, Daron, Victor Chernozhukov, Iván Werning, and Michael D Whinston. 2021. "Optimal targeted lockdowns in a multigroup SIR model." *American Economic Review: Insights*, 3(4): 487–502.

Ahituv, Avner, V Joseph Hotz, and Tomas Philipson. 1996. "The responsiveness of the demand for condoms to the local prevalence of AIDS." *Journal of Human Resources*, 869–897.

Allcott, Hunt, Levi Boxell, Jacob Conway, Matthew Gentzkow, Michael Thaler, and David Yang. 2020. "Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic." *Journal of Public economics*, 191: 104254.

Allen, Douglas W. 2022. "Covid-19 lockdown cost/benefits: A critical assessment of the literature." *International Journal of the Economics of Business*, 29(1): 1–32.

Anderson, Roy M, and Robert M May. 1991. *Infectious diseases of humans: dynamics and control.* Oxford university press.

Atkeson, Andrew G, Karen Kopecky, and Tao Zha. 2021. "Behavior and the Transmission of COVID-19." *AEA Papers and Proceedings*, 111: 356–360.

Auld, M Christopher. 2003. "Choices, beliefs, and infectious disease dynamics." *Journal of Health Economics*, 22(3): 361–377.

Auld, M. Christopher. 2006. "Estimating behavioral response to the AIDS epidemic." *Contributions in Economic Analysis and Policy*, 5.

Auld, M Christopher, and Flavio Toxvaerd. 2021. "The great Covid-19 vaccine rollout: Behavioural and policy responses." *National Institute Economic Review*, 257: 14–35.

Aum, Sangmin, Sang Yoon Lee, and Yongseok Shin. 2021. "COVID-19 doesn't need lockdowns to destroy jobs: The effect of local outbreaks in Korea." *Labour Economics*, 70: 101993.

Avery, Christopher, Frederick Chen, and David McAdams. 2024. "Steady-State Social Distancing and Vaccination." *American Economic Review: Insights*, 6(1): 1–19.

Avery, Christopher, William Bossert, Adam Clark, Glenn Ellison, and Sara Fisher Ellison. 2020. "An economist's guide to epidemiology models of infectious disease." *Journal of Economic Perspectives*, 34(4): 79–104.

Baril-Tremblay, Dominique, Chantal Marlats, and Lucie Ménager. 2021. "Self-isolation." *Journal of Mathematical Economics*, 93: 102483.

Barrett, Scott, and Michael Hoel. 2007. "Optimal disease eradication." *Environmental and Development Economics*, 12(5): 627–652.

Bayham, Jude, and Eli P Fenichel. 2020. "Impact of school closures for COVID-19 on the US health-care workforce and net mortality: a modelling study." *The Lancet Public Health*, 5(5): e271–e278.

Bayham, Jude, Nicolai V. Kuminoff, Quentin Gunn, and Eli P. Fenichel. 2015. "Measured voluntary avoidance behaviour during the 2009 A/H1N1 epidemic." *Proceedings of the Royal Society, London B.*, 282: 20150814.

Begon, Michael, Malcolm Bennett, Roger G Bowers, Nigel P French, SM Hazel, and Joseph Turner. 2002. "A clarification of transmission terms in host-microparasite models: numbers, densities and areas." *Epidemiology & Infection*, 129(1): 147–153.

Berry, Christopher R, Anthony Fowler, Tamara Glazer, Samantha Handel-Meyer, and Alec MacMillen. 2021. "Evaluating the effects of shelter-in-place policies during the COVID-19 pandemic." *Proceedings of the National Academy of Sciences*, 118(15).

Bisin, Alberto, and Andrea Moro. 2021. "Spatial-SIR with Network Structure and Behavior: Lockdown Rules and the Lucas Critique." National Bureau of Economic Research Working Paper 28932.

Bloom, David E., and Ajay S. Mahal. 1997. "Does the AIDS epidemic threaten economic growth?" *Journal of Econometrics*, 77(1): 105–124.

Bloom, David E, Michael Kuhn, and Klaus Prettner. 2022. "Modern infectious diseases: macroeconomic impacts and policy responses." *Journal of Economic Literature*, 60(1): 85–131.

Bobisud, LE. 1977. "Optimal control of a deterministic epidemic." *Mathematical Biosciences*, 35(1-2): 165–174.

Boozer, Michael A, and Tomas J Philipson. 2000. "The impact of public testing for human immunodeficiency virus." *Journal of Human Resources*, 419–446.

Boozer, Michael, and Tomas Philipson. 1996. "The Private Demand for Information and the Effects of Public Testing Programs: The Case of HIV." Discussion paper, Yale University.

Brito, Dagobert L, Eytan Sheshinski, and Michael D Intriligator. 1991. "Externalities and compulsary vaccinations." *Journal of Public Economics*, 45(1): 69–90.

Bürgi, C., and K. Wohlrabe. 2022. "The influence of Covid-19 on publications in economics: bibliometric evidence from five working paper series." *Scientometrics*, 127: 5175–5189.

Chakraborty, Shankha, Chris Papageorgiou, and Fidel Perez Sebastian. 2010. "Disease, infection dynamics, and development." *Journal of Monetary Economics*, 57: 859–872.

Chen, Frederick. 2012. "A mathematical analysis of public avoidance behavior during epidemics using game theory." *Journal of Theoretical Biology*, 302: 18–28.

Chen, Frederick, and Flavio Toxvaerd. 2014. "The economics of vaccination." *Journal of Theoretical Biology*, 363: 105–117.

Chernozhukov, Victor, Hiroyuki Kasahara, and Paul Schrimpf. 2021. "Causal impact of masks, policies, behavior on early COVID-19 pandemic in the US." *Journal of Econometrics*, 220(1): 23–62.

Chu, Derek, Elie Akl, Amena El-Harakeh, Antonio Bognanni, Tamara Lotf, Mark Loeb, Anisa Hajizadeh, Anna Bak, Ariel Izcovich, and Carlos A Cuello-Garcia. 2020. "Physical Distancing, Face Masks, and Eye Protection to Prevent Person-Person COVID-19 Transmission: A Systematic Review and Meta-Analysis." *Lancet*, 395(10242): 1973–1988.

Dall'Antonia, M, PG Coen, M Wilks, A Whiley, and M Millar. 2005. "Competition between methicillin-sensitive and-resistant Staphylococcus aureus in the anterior nares." *Journal of Hospital Infection*, 61(1): 62–67.

Dangerfield, Ciara, Eli P Fenichel, David Finnoff, Nick Hanley, Shaun Hargreaves Heap, Jason F Shogren, and Flavio Toxvaerd. 2022. "Challenges of integrating economics into epidemiological analysis of and policy responses to emerging infectious diseases." *Epidemics*, 39: 100585.

Darden, Michael E., David Dowdy, Lauren Gardner, Barton H. Hamilton, Karen Kopecky, Melissa Marx, Nicholas W. Papageorge, Daniel Polsky, Kimberly A. Powers, Elizabeth A. Stuart, and Matthew V. Zahn. 2022. "Modeling to inform economy-wide pandemic policy: Bringing epidemiologists and economists together." *Health Economics (United Kingdom)*, 31(7): 1291–1295.

Delavande, Adeline, and Hans-Peter Kohler. 2016. "HIV/AIDS-related expectations and risky sexual behaviour in Malawi." *The Review of Economic Studies*, 83(1): 118–164.

Diamond, Peter A, and Eric Maskin. 1979. "An equilibrium analysis of search and breach of contract, I: Steady states." *The Bell Journal of Economics*, 282–316.

Dorfman, Robert. 1943. "The detection of defective members of large populations." *The Annals of Mathematical Statistics*, 14(4): 436–440.

Dubois, Pierre, and Gokce Gokkoca. 2023. "Antibiotic Demand in the Presence of Antimicrobial Resistance." Toulouse School of Economics.

Eichenbaum, Martin S, Sergio Rebelo, and Mathias Trabandt. 2021. "The macroeconomics of epidemics." *The Review of Financial Studies*, 34(11): 5149–5187.

Ellison, Glenn. 2020. "Implications of Heterogeneous SIR Models for Analyses of COVID-19." NBER Working Paper W27373.

Engle, Samuel, Jussi Keppo, Marianna Kudlyak, Elena Quercioli, Lones Smith, and Andrea Wilson. 2021. "The Behavioral SIR Model, with Applications to the Swine Flu and COVID-19 Pandemics." University of Wisconsin Working Paper.

Eubank, Stephen, Hasan Guclu, VS Anil Kumar, Madhav V Marathe, Aravind Srinivasan, Zoltan Toroczkai, and Nan Wang. 2004. "Modelling disease outbreaks in realistic urban social networks." *Nature*, 429(6988): 180–184.

Eubank, Stephen, VS Anil Kumar, Madhav V Marathe, Aravind Srinivasan, and Nan Wang. 2006. "Structure of social contact networks and their impact on epidemics." *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 70: 181.

Farboodi, Maryam, Gregor Jarosch, and Robert Shimer. 2021. "Internal and external effects of social distancing in a pandemic." *Journal of Economic Theory*, 196: 105293.

Felder, Stefan, and Thomas Mayrhofer. 2022. "The Economics of Medical Decision Making." In *Medical Decision Making: A Health Economic Primer*. 209–231. Springer.

Fenichel, Eli P. 2013. "Economic considerations for social distancing and behavioral based policies during an epidemic." *Journal of Health Economics*, 32(2): 440–451.

Fenichel, Eli P, Carlos Castillo-Chavez, M Graziano Ceddia, Gerardo Chowell, Paula A Gonzalez Parra, Graham J Hickling, Garth Holloway, Richard Horan, Benjamin Morin, Charles Perrings, et al. 2011. "Adaptive human behavior in epidemiological models." *Proceedings of the National Academy of Sciences*, 108(15): 6306–6311.

Fenichel, Eli P., Nicolai V. Kuminoff, and Gerardo Chowell. 2013. "Skip the trip: Air travelers' behavioral responses to pandemic influenza." *PLoS ONE*, 8(3): e58249.

Fenichel, Eli P, R Tobias Koch, Anna Gilbert, Gregg Gonsalves, and Anne L Wyllie. 2021. "Understanding the barriers to pooled SARS-CoV-2 testing in the United States." *Microbiology Spectrum*, 9(1): 10.1128/spectrum. 00312–21.

Fenichel, E. P., and R. D. Horan. 2007. "Jointly-determined ecological thresholds and economic trade-offs in wildlife disease management." *Natural Resource Modeling*, 20(4): 511–547.

Ferguson, Neil. 2007. "Capturing human behaviour." *Nature*, 446(7137): 733–733.

Ferguson, Neil M, Daniel Laydon, Gemma Nedjati-Gilani, Natsuko Imai, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, Zulma Cucunubá, Gina Cuomo-Dannenburg, et al. 2020. "Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand." *Imperial College COVID-19 Response Team*, 20.

Fine, Paul, Ken Eames, and David L Heymann. 2011. ""Herd immunity": a rough guide." *Clinical infectious diseases*, 52(7): 911–916.

Flaxman, Seth, Swapnil Mishra, Axel Gandy, H. Juliette T. Unwin, Thomas A. Mellan, Helen Coupland, Charles Whittaker, Harrison Zhu, Tresnia Berah, Jeffrey W. Eaton, Mélodie Monod, Pablo N. Perez-Guzman, Nora Schmit, Lucia Cilloni, Kylie E. C. Ainslie, Marc Baguelin, Adhiratha Boonyasiri,

Olivia Boyd, Lorenzo Cattarino, Laura V. Cooper, Zulma Cucunubá, Gina Cuomo-Dannenburg, Amy Dighe, Bimandra Djaafara, Ilaria Dorigatti, Sabine L. van Elsland, Richard G. FitzJohn, Katy A. M. Gaythorpe, Lily Geidelberg, Nicholas C. Grassly, William D. Green, Timothy Hallett, Arran Hamlet, Wes Hinsley, Ben Jeffrey, Edward Knock, Daniel J. Laydon, Gemma Nedjati-Gilani, Pierre Nouvellet, Kris V. Parag, Igor Siveroni, Hayley A. Thompson, Robert Verity, Erik Volz, Caroline E. Walters, Haowei Wang, Yuanrong Wang, Oliver J. Watson, Peter Winskill, Xiaoyue Xi, Patrick G. T. Walker, Azra C. Ghani, Christl A. Donnelly, Steven M. Riley, Michaela A. C. Vollmer, Neil M. Ferguson, Lucy C. Okell, Samir Bhatt, and Covid-Response Team Imperial College. 2020. "Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe." *Nature*, 584(7820): 257–61.

Francis, Peter J. 1997. "Dynamic epidemiology and the market for vaccinations." *Journal of Public Economics*, 63(3): 383–406.

Funk, Sebastian, Marcel Salathé, and Vincent AA Jansen. 2010. "Modelling the influence of human behaviour on the spread of infectious diseases: a review." *Journal of the Royal Society Interface*, 7(50): 1247–1256.

Gans, Joshua S. 2023. "Vaccine Hesitancy, Passports, And The Demand For Vaccination." *International Economic Review*, 64(2): 641–652.

Geoffard, Pierre-Yves, and Tomas Philipson. 1996. "Rational epidemics and their public control." *International Economic Review*, 603–624.

Geoffard, Pierre-Yves, and Tomas Philipson. 1997. "Disease eradication: private versus public vaccination." *The American Economic Review*, 87(1): 222–230.

Gersovitz, Mark. 2011. "The economics of infection control." *Annu. Rev. Resour. Econ.*, 3(1): 277–296.

Gersovitz, Mark, and Jeffrey S Hammer. 2004. "The economical control of infectious diseases." *The Economic Journal*, 114(492): 1–27.

Giannitsarou, Chryssi, Stephen Kissler, and Flavio Toxvaerd. 2021. "Waning immunity and the second wave: Some projections for SARS-CoV-2." *American Economic Review: Insights*, 3(3): 321–338.

Glaeser, Edward L., Caitlin Gorback, and Stephen J. Redding. 2022. "How much does COVID-19 increase with mobility? Evidence from New York and four other U.S. cities." *Journal of Urban Economics*, 127: 103292.

Goenka, Aditya, Lin Liu, and Manh-Hung Nguyen. 2014. "Infectious diseases and economic growth." *Journal of Mathematical Economics*, 50: 34–53.

Goldman, Steven, and James Lightwood. 2002. "Cost optimization in the SIS model of infectious disease with treatment." *Topics in Economic Analysis & Policy*, 2(1): 1007.

Gollier, Christian. 2020. "Optimal Group Testing to Exit the Covid Confinement." Working Paper, Toulouse School of Economics.

Goolsbee, Austan, and Chad Syverson. 2021. "Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020." *Journal of Public Economics*, 193: 104311.

Grantz, Kyra H, Hannah R Meredith, Derek AT Cummings, C Jessica E Metcalf, Bryan T Grenfell, John R Giles, Shruti Mehta, Sunil Solomon, Alain Labrique, Nishant Kishore, et al. 2020. "The use of mobile phone data to inform analysis of COVID-19 pandemic epidemiology." *Nature Communications*, 11(1): 4961.

Greenwood, Jeremy, Philipp Kircher, Cezar Santos, and Michele Tertilt. 2019. "An equilibrium model of the African HIV/AIDS epidemic." *Econometrica*, 87(4): 1081–1113.

Grenfell, Bryan T, Ottar N Bjørnstad, and Jens Kappey. 2001. "Travelling waves and spatial hierarchies in measles epidemics." *Nature*, 414(6865): 716–723.

Grossman, Michael. 1972. "On the Concept of Health Capital and the Demand for Health." *The Journal of Political Economy*, 80(2): 223–255.

Gupta, Sumedha, Kosali Simon, and Coady Wing. 2020. "Mandated and voluntary social distancing during the COVID-19 epidemic." *Brookings Papers on Economic Activity*, 2020(2): 269–326.

Gupta, Sumedha, Thuy Nguyen, Shyam Raman, Byungkyu Lee, Felipe Lozano-Rojas, Ana Bento, Kosali Simon, and Coady Wing. 2021. "Tracking Public and Private Responses to the COVID-19 Epidemic." *American Journal of Health Economics*, 7(4): 361–404.

Haber, Noah A, Emma Clarke-Deelder, Avi Feller, Emily R Smith, Joshua A Salomon, Benjamin MacCormack-Gelles, Elizabeth M Stone, Clara Bolster-Foucault, Jamie R Daw, Laura Anne Hatfield, et al. 2022. "Problems with evidence assessment in COVID-19 health policy impact evaluation: a systematic review of study design and evidence strength." *British Medical Journal Open*, 12(1): e053820.

Hale, Thomas, Noam Angrist, Rafael Goldszmidt, Beatriz Kira, Anna Petherick, Toby Phillips, Samuel Webster, Emily Cameron-Blake, Laura Hallas, Saptarshi Majumdar, et al. 2021. "A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker)." *Nature Human Behaviour*, 1–10.

Hanski, Ilkka. 1999. *Metapopulation ecology.* Oxford University Press.

Hethcote, Herbert W. 2000. "The mathematics of infectious diseases." *SIAM review*, 42(4): 599–653.

Hochman, E., and D Zilberman. 1978. "Examination of environmental policies using production and pollution microparameter distributions." *Econometrica*, 46(4): 729–760.

Hoover, Darren, and Flavio Toxvaerd. 2022. "Epidemics in Space: Control, Targeting and Delegation." Centre for Economic Policy Research.

Hsiang, Solomon, Daniel Allen, Sébastien Annan-Phan, Kendon Bell, Ian Bolliger, Trinetta Chong, Hannah Druckenmiller, Luna Yue Huang, Andrew Hultgren, Emma Krasovich, Peiley Lau, Jaecheol Lee, Esther Rolf, Jeanette

Tseng, and Tiffany Wu. 2020. "The effect of large-scale anti-contagion policies on the COVID-19 pandemic." *Nature*, 584: 262–268.

Hu, Tao, Siqin Wang, Bing She, Mengxi Zhang, Xiao Huang, Yunhe Cui, Jacob Khuri, Yaxin Hu, Xiaokang Fu, Xiaoyue Wang, et al. 2021. "Human mobility data in the COVID-19 pandemic: characteristics, applications, and challenges." *International Journal of Digital Earth*, 14(9): 1126–1147.

Jackson, Matthew O, and Dunia López-Pintado. 2013. "Diffusion and contagion in networks with heterogeneous agents and homophily." *Network Science*, 1(1): 49–67.

Jedwab, Remi, Noel D. Johnson, and Mark Koyama. 2022. "The Economic Impact of the Black Death." *Journal of Economic Literature*, 60(1): 132–78.

Julliard, Christian, Ran Shi, and Kathy Yuan. 2023. "The spread of COVID-19 in London: Network effects and optimal lockdowns." *Journal of Econometrics*, 235: 2125–2154.

Kaufmann, Marc, Peter Andre, and Botond Kőszegi. 2024. "Understanding markets with socially responsible consumers." *The Quarterly Journal of Economics*, 139: 1989–2035.

Keeling, Matt J., and Pejman Rohani. 2008. *Modeling Infectious Diseases in Humans and Animals.* Princeton University Press.

Kerwin, Jason T. 2012. "'Rational fatalism': non-monotonic choices in response to risks." Working paper, Working Group in African Political Economy meeting, University of California, Berkeley, CA.

Kerwin, Jason T. 2020. "Scared straight or scared to death? Fatalism in response to disease risks." Working paper, University of Minnesota.

Kim, Dongwoo, and Young Jun Lee. 2022. "Vaccination strategies and transmission of COVID-19: Evidence across advanced countries." *Journal of Health Economics*, 82: 102589.

Korolev, Ivan. 2022. "On reduced form estimation of the effect of policy interventions on the COVID-19 pandemic." *The Econometrics Journal*, 25(3): 762–780.

Kremer, Michael. 1994. "Can having fewer partners increase prevalence of AIDS?" Working paper, National Bureau of Economic Research Cambridge, Mass., USA.

Kremer, Michael. 1996. "Integrating behavioral choice into epidemiological models of AIDS." *The Quarterly Journal of Economics*, 111(2): 549–573.

Kremer, Michael, and Charles Morcom. 1998. "The effect of changing sexual activity on HIV prevalence." *Mathematical biosciences*, 151(1): 99–122.

Lakdawalla, Darius, Emmett Keeler, Dana P Goldman, and Erin Trish. 2020. "Getting America Back to Work and Study With Pooled Testing." Schaeffer Center White Paper Series, University of Southern California.

Lakdawalla, Darius, Neeraj Sood, and Dana Goldman. 2006. "HIV breakthroughs and risky sexual behavior." *The Quarterly Journal of Economics*, 121(3): 1063–1102.

Laxminarayan, Ramanan, and Gardner M Brown. 2001. "Economics of antibiotic resistance: a theory of optimal use." *Journal of Environmental Economics and Management*, 42(2): 183–206.

Laxminarayan, Ramanan, and Markus Herrmann. 2015. "Biological resistance." In *Handbook on the Economics of Natural Resources*. 249–278. Edward Elgar Publishing.

Lightwood, James, and Steven M Goldman. 1995. "The SIS model of infectious disease with treatment." *Unpublished Manuscript*.

Lipsey, Richard G, and Kelvin Lancaster. 1956. "The general theory of second best." *The Review of Economic Studies*, 24(1): 11–32.

Lipsitch, Marc, Ted Cohen, Megan Murray, and Bruce R Levin. 2007. "Antiviral resistance and the control of pandemic influenza." *PLoS medicine*, 4(1): e15.

Li, Ying-Chun, Edward C Norton, and William H Dow. 2004. "Influenza and pneumococcal vaccination demand responses to changes in infectious disease mortality." *Health Services Research*, 39(4p1): 905–926.

Makridis, Christos, and Jonathan Rothwell. 2020. "The Real Cost of Political Polarization: Evidence from the COVID-19 Pandemic." *SSRN Electronic Journal*.

Makris, Miltiadis. 2024. "Covid and social distancing with a heterogenous population." *Economic Theory*, 77(1): 445–494.

Makris, Miltiadis, and Flavio Toxvaerd. 2020. "Great expectations: Social distancing in anticipation of pharmaceutical innovations." Faculty of Economics, University of Cambridge.

Maler, Karl-Goran, Anastasios Xepapadeas, and Aart De Zeeuw. 2003. "The economics of shallow lakes." *Environmental and Resource Economics*, 26: 603–624.

Maloney, William, and Temel Taskin. 2020. "Determinants of social distancing and economic activity during COVID-19: A global view." *Covid Economics*, 156.

Matthies, Konstantin, and Flavio Toxvaerd. 2023. "Rather doomed than uncertain: risk attitudes and transmissive behavior under asymptomatic infection." *Economic Theory*, 76(1): 1–44.

McAdams, David. 2017. "Resistance diagnosis and the changing epidemiology of antibiotic resistance." *Annals of the New York Academy of Sciences*, 1388(1): 5–17.

McAdams, David. 2021. "The blossoming of economic epidemiology." *Annual Review of Economics*, 13: 539–570.

McAdams, David, Yangbo Song, and Dihan Zou. 2023. "Equilibrium social activity during an epidemic." *Journal of Economic Theory*, 207: 105591.

McAdams, D, K Wollein Waldetoft, C Tedijanto, M Lipsitch, and SP Brown. 2019. "Resistance diagnostics as a public health tool to combat antibiotic resistance: A model-based evaluation." *PLoS Biolology*, 17.

McCallum, Hamish, Nigel Barlow, and Jim Hone. 2001. "How should pathogen transmission be modelled?" *Trends in ecology & evolution*, 16(6): 295–300.

McKusick, Leon, William Horstman, and Thomas J Coates. 1985. "AIDS and sexual behavior reported by gay men in San Francisco." *American Journal of Public Health*, 75(5): 493–496.

Mechoulan, Stéphane. 2007. "Risky sexual behavior, testing, and HIV treatments." *Forum for Health Economics & Policy*, 10.

Meriggi, Niccolò F, Maarten Voors, Madison Levine, Vasudha Ramakrishna, Desmond Maada Kangbai, Michael Rozelle, Ella Tyler, Sellu Kallon, Junisa Nabieu, Sarah Cundy, et al. 2024. "Last-mile delivery increases vaccine uptake in Sierra Leone." *Nature*, 627(8004): 612–619.

Morton, Richard, and Kenneth H Wickwire. 1974. "On the optimal control of a deterministic epidemic." *Advances in Applied Probability*, 6(4): 622–635.

Mullahy, John. 1999. "It'll only hurt a second? Microeconomic determinants of who gets flu shots." *Health Economics*, 8(1): 9–24.

Murray, Eleanor J. 2020. "Epidemiology's time of need: COVID-19 calls for epidemic-related economics." *Journal of Economic Perspectives*, 34(4): 105–120.

Newman, Mark EJ. 2002. "Spread of epidemic disease on networks." *Physical review E*, 66(1): 016128.

Okubo, Akira, Simon A Levin, et al. 2001. *Diffusion and ecological problems: modern perspectives.* Vol. 14, Springer.

O'Neill, Jim. 2016. "Tackling drug-resistant infections globally: final report and recommendations." Government of the United Kingdom.

Oster, Emily. 2012. "HIV and sexual behavior change: Why not Africa?" *Journal of Health Economics*, 31(1): 35–49.

Oster, Emily. 2018. "Does disease cause vaccination? Disease outbreaks and vaccination response." *Journal of Health Economics*, 57: 90–101.

Otto, Sarah P, and Troy Day. 2007. *A biologist's guide to mathematical modeling in ecology and evolution.* Princeton University Press.

Pastor-Satorras, Romualdo, Claudio Castellano, Piet Van Mieghem, and Alessandro Vespignani. 2015. "Epidemic processes in complex networks." *Reviews of Modern Physics*, 87(3): 925.

Paula, Áureo De, Gil Shapira, and Petra E Todd. 2014. "How beliefs about HIV status affect risky behaviors: Evidence from Malawi." *Journal of Applied Econometrics*, 29(6): 944–964.

Peltzman, Sam. 1975. "The effects of automobile safety regulation." *Journal of political Economy*, 83(4): 677–725.

Perrings, Charles, Carlos Castillo-Chavez, Gerardo Chowell, Peter Daszak, Eli P Fenichel, David Finnoff, Richard D Horan, A Marm Kilpatrick, Ann P Kinzig, Nicolai V Kuminoff, et al. 2014. "Merging economics and epidemiology to improve the prediction and management of infectious disease." *EcoHealth*, 11: 464–475.

Philipson, Tomas. 1995. "The welfare loss of disease and the theory of taxation." *Journal of Health Economics*, 14(3): 387–395.

Philipson, Tomas. 1996. "Private vaccination and public health: an empirical examination for US measles." *Journal of Human Resources*, 611–630.

Philipson, Tomas. 2000. "Economic epidemiology and infectious diseases." *Handbook of health economics*, 1: 1761–1799.

Philipson, Tomas J, and Richard A Posner. 1993. *Private choices and public health: The AIDS epidemic in an economic perspective.* Harvard University Press.

Philipson, Tomas J, and Richard A Posner. 1995. "A theoretical and empirical investigation of the effects of public health subsidies for STD testing." *The Quarterly Journal of Economics*, 110(2): 445–474.

Poudel, Ak Narayan, Shihua Zhu, Nicola Cooper, Paul Little, Carolyn Tarrant, Matthew Hickman, and Guiqing Yao. 2023. "The economic burden of antibiotic resistance: A systematic review and meta-analysis." *PLOS One*, 18(5): e0285170.

Prodanov, Dimiter. 2022. "Analytical solutions and parameter estimation of the SIR epidemic model." *Mathematical Analysis of Infectious Diseases*, 163–189.

Quaas, Martin F, Jasper Meya, Hanna Schenk, Björn Bos, Moritz A Drupp, and Till Requate. 2020. "The Social Cost of Contacts: Theory and Evidence for the COVID-19 Pandemic in Germany." *Available at SSRN 3606810.*

Quiggin, John, and Richard Holden. 2021. "Why most economists continue to back lockdowns."

Reluga, Timothy. 2010. "Game theory of social distancing in response to an epidmic." *PLOS Computational Biology*, 6.

Reluga, Timothy C., and Alison P. Galvani. 2011. "A general approach for population games with application to vaccination." *Mathematical Biosciences*, 230(2): 67–78.

Rowthorn, Bob RE, and Flavio Toxvaerd. 2012. "The optimal control of infectious diseases via prevention and treatment." CEPR Discussion Paper No. DP8925.

Sanders, Jerry L. 1971. "Quantitative guidelines for communicable disease control programs." *Biometrics*, 883–893.

Sears, James, J Miguel Villas-Boas, Sofia Berto Villas-Boas, and Vasco Villas-Boas. 2023. "Are we stayinghome to flatten the curve?" *American Journal of Health Economics*, 9(1): 71–95.

Sethi, Suresh P. 1974. "Quantitative guidelines for communicable disease control program: a complete synthesis." *Biometrics*, 681–691.

Sethi, Suresh P, and Preston W Staats. 1978. "Optimal control of some simple deterministic epidemic models." *Journal of the Operational Research Society*, 29(2): 129–136.

Springborn, Michael, Gerardo Chowell, Matthew MacLachlan, and Eli P Fenichel. 2015. "Accounting for behavioral responses during a flu epidemic using home television viewing." *BMC infectious diseases*, 15: 1–14.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics*, 225(2): 175–199. Themed Issue: Treatment Effect 1.

Talamàs, Eduard, and Rakesh Vohra. 2020. "Free and perfectly safe but only partially effective vaccines can harm everyone." *Games and economic behavior*, 122: 277–289.

Thornton, Rebecca L. 2008. "The Demand for, and Impact of, Learning HIV Status." *American Economic Review*, 98(5): 1829–63.

Toxvaerd, Flavio. 2009. "Recurrent infection and externalities in treatment." mimeo.

Toxvaerd, Flavio. 2017. "On the Dynamics of Beliefs and Risky Sexual Behavior." *SSRN Working Paper*. https://ssrn.com/abstract=3523662.

Toxvaerd, Flavio. 2019. "Rational disinhibition and externalities in prevention." *International Economic Review*, 60(4): 1737–1755.

Toxvaerd, Flavio. 2020. "Equilibrium social distancing." Faculty of Economics, University of Cambridge.

Toxvaerd, Flavio. 2022. "Silent Spreaders: Behavior and Equilibrium Under Asymptomatic Infection." Working paper, Centre for Economic Policy Research.

Toxvaerd, Flavio. 2024. "Contacts, altruism and competing externalities." *European Economic Review*, 167: 104794.

Toxvaerd, Flavio, and Robert Rowthorn. 2022. "On the management of population immunity." *Journal of Economic Theory*, 204: 105501.

Veliov, Vladimir M. 2005. "On the effect of population heterogeneity on dynamics of epidemic diseases." *Journal of Mathematical Biology*, 51: 123–143.

Venkatramanan, Srinivasan, Bryan Lewis, Jiangzhuo Chen, Dave Higdon, Anil Vullikanti, and Madhav Marathe. 2018. "Using data-driven agent-based models for forecasting emerging infectious diseases." *Epidemics*, 22: 43–49.

Wang, Peipei, Xinqi Zheng, and Haiyan Liu. 2022. "Simulation and forecasting models of COVID-19 taking into account spatio-temporal dynamic characteristics: A review." *Frontiers in public health*, 10: 1033432.

Yan, Youpei, Amyn A Malik, Jude Bayham, Eli P Fenichel, Chandra Couzens, and Saad B Omer. 2021*a*. "Measuring voluntary social distancing behavior during the COVID-19 pandemic." *Proceedings of the National Academy of Sciences*, 118(16).

Yan, Youpei, Jude Bayham, Eli P Fenichel, and Aaron Richter. 2021*b*. "Risk compensation and face mask mandates during the COVID-19 pandemic." *Scientific Reports*, 11(1).