THE SAM APPROACH TO EPIDEMIC MODELS

Pietro Garibaldi, Espen R. Moen
and Christopher A. Pissarides

ABSTRACT

We discuss the connections between epidemiology models and search and matching (SAM) approach and draw conclusions about modeling the trade-offs between lockdowns and disease spread. We review the pre-COVID epidemics literature, which was mainly by epidemiologists, and the post-COVID surge in economics papers that use meeting technologies to model the trade-offs. We argue that modeling the decentralized equilibrium with economic trade-offs gives rise to substantially different results from the earlier epidemics literature, but policy action is still welfare-improving because of several externalities.

Keywords: SIR models; matching models; COVID-19; social distancing; herd immunity; externalities

JEL codes: A12; I10; J18; D61-D62

1. INTRODUCTION

Models of infectious diseases share a key feature with a large number of economic models. This is that the utility (positive or negative) that an agent derives from their activities requires contact with another agent in "social space." In economics, this feature of exchange is stronger in frictional markets, in which participants need to search, find, and inspect alternatives before deciding to buy or sell, than in market-clearing neoclassical markets, in which goods are homogenous. A typical example is the labor market, in which participants need to search over alternatives before agreeing to a match. Another is the housing market, in which buyers need to search and inspect houses before buying.

50th Celebratory Volume
Research in Labor Economics, Volume 50, 1–23
Copyright © 2023 by Emerald Publishing Limited
All rights of reproduction in any form reserved
ISSN: 0147-9121/doi:10.1108/S0147-912120230000050001

Uncorrected Proof

A useful way to formalize these interactions is as a dynamic process during which the individual changes state, e.g., from unemployed to employed, or from a renter to a homeowner.[1] The transition from one state to another can then be modeled as in the "search and matching," or SAM, approach. This approach has been successfully applied initially to models of the labor market and subsequently to several others in which there is need to gather idiosyncratic information before exchange.[2]

The dynamics of epidemics have been modeled as transitions between states long before economists developed this particular methodology (Kermack & McKendrick, 1927). But the use of models of individual behavior driven by incentives, which motivated economists, was not applied to epidemiological models until much later. Our motivation for writing this paper is that we believe that epidemiological models in the spirit of Kermack and McKendrick (1927), have a lot to gain from the insights of the SAM approach, as developed by economists.[3] We critically survey models of infectious diseases that use matching processes to derive the transitions from a healthy (or "susceptible") state to an infectious one. As we will argue, variants of this approach have been used sparingly in the pre-COVID research but exploded during the COVID-19 pandemic.

In economic models the typical SAM situation is one in which contact is the result of some costly activity, such as search, and yields positive returns if successful, or nothing. In contrast to this typical situation, in an epidemic there is a positive probability that contact between two people would involve some cost: the transmission of the disease from an infected to a healthy individual. Putting the economic and epidemiological models together, a trade-off is created. The more activities you pursue in social space to increase your economic payoffs, the higher the risk of infection. Viewed from the lens of the economist, an epidemic introduces a cost to the economic interaction in social space, which shifts economic activities in favor of the ones that do not involve social interaction. For example, shopping activities shift from browsing in shops to buying online. Viewed from the lens of epidemiology, restricting social contact contains the disease but it puts a cost on society, the reduction in utility that could be derived from social interaction.

Before the outbreak of the COVID-19 pandemic in 2019, this trade-off led to a modest literature, mostly written by mathematical biologists, although there are also some papers by economists. Epidemiologists focused mainly on influenza outbreaks, whereas the best-known papers by economists are about HIV transmission. Following the outbreak of COVID-19, however, a very large economics literature emerged, covering all aspects of the pandemic and its economic costs. Our survey covers only a subfield of these literature strands, the one that models transitions between states by making use of a contact function, which could be a matching function as in SAM theory, or any other mathematical representation of contacts.

We summarize the way that this function has been modeled by epidemiologists and economists and show that its form influences the nature of equilibrium and its welfare properties. We argue that although in the labor literature a

Uncorrected Proof

linear-homogeneous function has received a lot of empirical support, in epidemiology it is more common to use a function with increasing returns to scale (usually the quadratic).[4] This distinction is plausible because in economics contacts are usually used for the exchange of a single good, e.g., a house or a service that only one person can enjoy, whereas in epidemiology a person carrying the virus can pass it on to a large number of contacts. This distinction turns out to be important because an action such as social distancing is much more effective in reducing contacts if the contact function is quadratic than if it is linear. So the policy recommendation due to the quadratic, or any other increasing returns function, is usually for more social distancing, but the exact recommendation depends also on other features of the model, such as heterogeneity or forward-looking behavior.[5]

We split the pre-COVID economics research into two main strands. First, we review the HIV research, working mainly with an SI (susceptible-infected) model. Second, we review models that use Kermack and McKendrick's (1927) susceptible-infected-removed (SIR) model. Key contributions to the HIV research on the SI model are Geoffard and Philipson (1996) and Kremer (1996). These models do not introduce explicitly an aggregate contact function that depends on individual incentives. Yet, they show that the flow from susceptible to infected individuals depends on the number of people currently infected and the prevailing sexual activity, which depends on the HIV infection risk. The SIR model pre-COVID was applied mainly to the modeling of influenza epidemics, and the question usually addressed was how much social distancing should the government legislate during an epidemic. Early models that introduced some kind of contact function in influenza models include Reluga (2010) and Chen, Miahoua, Rabidoux, and Robinson (2011).

With the outbreak of COVID-19, the economics literature exploded. A large number of papers focus on the optimal control of the pandemic by policy-makers, given a trade-off between deaths from the disease and GDP losses from lockdowns. Alvarez, Argente, and Lippi (2021) is an example of this class of models in the case of homogeneous agents while Acemoglu, Chernozhukov, Werning, and Whinston (2021) and Favero, Ichino, and Rustichini (2020) focus on heterogeneous agents. Acemoglu et al. (2021) use a SAM-type matching function to model contacts and assume increasing returns to contacts.

Several other models solve for equilibrium in the decentralized economy, in which optimizing agents respond to economic incentives or infection disincentives, depending on their state and information set. The main states which are influenced by individual decision-making are the susceptible and infected. In the susceptible state the decision of the individual whether to interact in social space or not influences their probability of infection, whereas in the infected state the influence is on the probability of infecting others. These models often obtain also the planning solution. Several of the papers in this class of models use a variant of the matching function to derive contacts between agents, which are usually restricted to be either linear-homogeneous or quadratic. We discuss models in this class in the main body of the paper.[6]

Uncorrected Proof

Section 2 discusses pre-COVID research, with emphasis on HIV and influenza epidemics. Section 3 focuses on the large COVID-19 literature that exploded in 2020, distinguishing between models with homogenous agents and models with heterogeneous agents. We then discuss the role of externalities between the decentralized and centralized solutions in both pre-COVID and COVID research. Section 4 concludes.

2. PRE-COVID RESEARCH

We distinguish between two contrasting strands in the pre-COVID research. In the first strand authors addressed transmissions of HIV, or other sexually transmitted diseases (STDs), by making use of a SI epidemiological model; in other words, a model in which there are no deaths or recoveries. The disease imposes some cost on the infected, which gives the reason people want to avoid it. In STD models, the contact that leads to the transmission is planned and usually restricted to two people. This contrasts with the second class of models, the study of influenza epidemics, which usually makes use of the original SIR model, in which there are recoveries from the disease (Kermack & McKendrick, 1927). In this class of models, disease transmissions can also arise after unplanned chance meetings between individuals.

In both types of models, contact between infected and susceptible individuals has features of the SAM approach. The typical labor situation with SAM is one in which two agents come into contact with a view to forming a productive relationship. The match takes place if both parties agree to it. This parallels contacts in STD situations virtually exactly.[7] In contrast, infections in the case of influenza can arise in a variety of situations in which people share space, giving rise to a different set of solutions. For example, one might plan a restaurant visit and get infected by someone else who happens to be in the restaurant.

The implications for the contact technology are that in STD situations the contact function can be approximated by a linear one. For example, suppose that in an HIV world a person engages in sexual contact with n other persons and contact with each involves a disease transmission risk β. The infection probability for this person is $p = 1 - (1 - \beta)^n$. If β (and consequently p) is small, this approximately satisfies $p = \beta n$, a linear transmission rate.

But in influenza situations, as in COVID-19, there is a proportionally bigger effect on infections if a person increases or restricts social behavior. As an approximation, the contact technology is typically assumed to be quadratic. To see the origins of this, suppose a person goes into social space n times during the week. Each time they goes out she gets sufficiently close to m other people that could infect her with probability β each. If m is fixed and independent of n, we could reason as in the case of HIV and proxy the infection probability by $p = \beta nm$. But m must depend on the number of times other agents go to social space. Let $\overline{n}$ be the average number of times other agents go to social space, then m is proportional to $\overline{n}$, $m = x\overline{n}$; e.g., if everyone else doubles the number of times they go to social space, the person who goes out n times will be twice as likely to

Uncorrected Proof

meet someone in that space. Therefore, the infection probability becomes $p = \beta x n \bar{n}$, and so in symmetric equilibrium, $p = \beta x n^2$.

In HIV situations, this "increasing returns" effect does not arise because if a person decides that they will have n partners, they can keep the n partners irrespective of how many partners others have and how many other "propositions" they get. But in influenza situations other people could affect the infections probability without the consent of the person in question.

We review each literature strand separately.

2.1 The SI Model and the Economics of HIV Epidemics

Early models (Geoffard & Philipson, 1996; Kremer, 1996) of STDs did not work with a general matching function in the spirit of SAM theory, but interpreted instead the standard incidence of the epidemiological model (Hetcote, 2000) through the lens of a random matching game. Of course, random matching is consistent with a simple matching function in which the aggregate probability of infection depends linearly on the stock of infected people, and this is the main transmission mechanism explored by these papers. The linear dependence of the transmission probability on the number of infected people is also a standard result in SAM models with linear technologies.

Consider the SI model studied by Toxvaerd (2019), who discusses the differences between the model without an economic dimension and the impact that economic incentives might have on it. In the former, a population $P = [0, 1]$ consists of a continuum of infinitely lived individuals who at each instant $t \geq 0$ can be in one of two states, namely, susceptible or infected. The set of infected individuals is denoted by $I(t)$ and has measure $I(t)$, whereas the set of susceptible individuals is denoted by $S(t)$ and has measure $S(t)$. In the absence of births and deaths the population size is normalized to unity, so these measures can be interpreted as fractions. $I(t)$ is referred to as the *disease prevalence*.[8]

At each instant, the population mixes homogeneously. This corresponds to *random matching*, where each individual has an equal chance of meeting any other individual, irrespective of the health status of the two matched individuals. A match between two infected individuals or two susceptible individuals does not create a new infection, but a match between an infected and a susceptible individual may do. In a continuous time model, the rate at which infection is transmitted in a match with a member of the $I(t)$ set is denoted by $\beta > 0$, so in a short interval of time dt the rate of getting the disease from an infected individual is βdt. This parameter captures the infectiousness of the disease. With random matching and large numbers, it follows that the probability that a susceptible person in this population is infected during the short interval dt is

$$\lambda(t) = \beta I(t) dt.$$

It follows that the average rate at which susceptible individuals become infected is

Uncorrected Proof

$$-\frac{dS(t)}{dt} = \beta I(t)S(t), \tag{1}$$

with $-\frac{dS(t)}{dt}$ giving the flow from the susceptible state to the infected. Thus, the rate of new infections, or *disease incidence*, is proportional to disease prevalence.

With this interpretation, the key transmission rate can be interpreted as the outcome of a SAM mechanism with linear technology (Diamond and Maskin (1979)). To turn the classical SI model into an economic model, Toxvaerd (2019) assumes that individuals earn a flow payoff $\pi_S > 0$ per instant while susceptible, a flow payoff $\pi_I < \pi_S$ per instant while infected, and that time is discounted at rate ρ. In Toxvaerd's notation $\pi = \pi_S - \pi_I > 0$ denotes the health premium. The health premium should be thought of broadly as the benefits of not being infected. To model the possibility of engaging in preventive behavior, assume that the individuals can affect the rate of infection by controlling the rate at which they expose themselves to a potential infection. In particular, at each instant $t \geq 0$, each susceptible individual $i \in S(t)$ noncooperatively chooses exposure level $\epsilon_i(t) \in [0, 1]$, at personal cost $(1 - \epsilon_i(t))c \geq 0$. Here, $\epsilon_i(t) = 0$ denotes complete shielding from social action whereas $\epsilon_i(t) = 1$ denotes no shielding at all, so c is the unit cost of shielding from social action. The introduction of shielding reduces the rate of infection to $\epsilon_i(t)\beta I(t)$. This formalization captures the notion that exposure is desirable, but shielding is pursued because it reduces the chance of an infection and the loss of the health premium. In a symmetric equilibrium, $\epsilon_i(t) = \epsilon(t)$ and the aggregate infection rate becomes (on the assumption that the infected population chooses full exposure),

$$-\frac{dS(t)}{dt} = \epsilon(t)\beta I(t)S(t).$$

Several new results follow from this general framework, since the choice of exposure depends both on the economic costs of contracting the disease and on the cost of shielding.

Returning now to HIV, in an early paper, Geoffard and Philipson (1996) show that in an optimizing model the aggregate transmission rate depends on incidence and derives some results about behavior from it. The model is SI; there is a population composed of agents who are either susceptible or infected, and engage in either protective or transmissive (exposed) activity. When susceptible, an agent can either become infected or remain susceptible; once infected, an agent remains infected. Agents continuously meet one another over time, and upon each meeting, they must decide whether to engage in transmissive or protective behavior. Susceptible agents who choose the former run the risk of contracting the disease, while susceptible agents who choose the latter run no such risk. Transmissive behavior is assumed to be desirable, and protective behavior costly. Since infection is an absorbing state, in the framework of Geoffard and Philipson (1996) no selfishly rational infected agent engages in protection and their problem is basically static. Even though individual behavior takes the form of a binary

Uncorrected Proof

decision, because of large numbers at the aggregate level the hazard function from susceptible to infected depends continuously on individual choices.

Geoffard and Philipson (1996) argue that "in standard mathematical epidemiology, this hazard rate is an increasing function of prevalence. In other words, the larger the fraction of infected people in the population, the larger is the fraction of uninfected people who become infected in the next period. This is because the larger the disease is, the larger is the chance that an individual who is still susceptible will meet an infected individual. This is true across a wide variety of epidemiological models, since they all share the feature that the demand for exposure does not respond to prevalence." In contrast, introducing economic considerations into the model implies that the hazard rate into infection may be a decreasing function of the prevalence of the disease because the individuals who are still susceptible face a larger risk of infection and increase protective behavior. Geoffard and Philipson (1996) show that the aggregate dynamics of the model can be written as

$$-\frac{dS(t)}{dt} = \beta G(I(t))Q(t)I(t). \tag{2}$$

As before, $S(t)$ and $I(t)$ denote the susceptible and infected population, respectively, and β denotes the infectiousness of the disease. $Q(t)$ denotes the probability that a susceptible agent who engages in transmissive activity (in the case of HIV, engages in sexual activity) during period $[0, t]$ is still susceptible at t, and $G(I(t))$ is an endogenous probability that keeps track of the share of susceptible people who engage in transmissive activity in t, as a function of prevalence $I(t)$. The function $G(I(t))$ picks up the disincentives that susceptible agents have; unlike the earlier epidemiological models without disincentives from prevalence, in which $G(.) = I(t)$, Geoffard and Philipson (1996) show that $G(.)$ is a decreasing function of prevalence $I(t)$. Note that it is only the present level of prevalence that matters, not expected future prevalence levels. To understand why, note that $I(t)$ is increasing in t as there are no deaths in the model. Furthermore, protective behavior reduces the probability of infection to zero. As a result, the individual decision of whether to protect oneself or not becomes like an optimal stopping problem, and the person stops transmissive behavior when the instantaneous gain from this behavior is exactly equal to the instantaneous cost associated with the risk of becoming ill. After that point, the person will never again engage in risky behavior. The future development of $I(t)$ is therefore irrelevant for the optimal stopping decision.[9]

A matching interpretation for this mechanism is one in which an individual is engaged in search sequentially but stops the search and enters an absorbing state when the cost of continuing rises up to the benefits of stopping. In this case the cost of continuing is the risk of infection, which rises during search because of the monotonic increase in infections. As in the simple matching model, there is heterogeneity across individuals which is not modeled explicitly and not identified a priori, but reflects the frictions inherent in the matching function (Pissarides, 2000, pp. 3–4). The friction here is the information about the health status of

Uncorrected Proof

other individuals in the market. When a susceptible individual meets another person, she has to decide if that person is a good match (susceptible) or a bad one (infected), and the more infected individuals there are, the more likely is the person to reject contacts. Of course, the absence of a priori information may lead to a failure early on, despite the low perceived probability $G(I_t)$.

Kremer (1996) – in a paper published in the same year as Geoffard and Philipson (1996) – also argues that most epidemiological models treat behavior as independent of prevalence. In contrast, he argues that his "analysis differs both from the traditional epidemiological analysis, which takes behavior as independent of prevalence, and from the few attempts to introduce behavioral considerations into epidemiology, which do not formally model how decisions about the rate of partner change depend on the composition of available partners." Kremer (1996) models the behavioral choice in the rate of partner change by writing

$$-\frac{dS(t)}{dt} = i(I(t))\beta I(t)S(t), \tag{3}$$

where the function $i()$ is the rate of partner change. The key contribution is thus similar to Geoffard and Philipson (1996), in the sense that the rate of partner change is the outcome of an optimizing decision that corresponds to partner selection. This rate depends on the stock of infection, referred to as incidence in this literature. The motivation given is different, and appealing in the context of HIV, in that partner change involves a potential cost, the possibility of getting infected from the new partner. In the context of social activity in epidemics that we discussed in our introduction, staying with the same partner in HIV epidemics has a similar impact on the spread of the disease as complete social distancing in epidemics such as COVID-19.[10]

Greenwood, Kircher, Santos, and Tertil (2019) introduce directed search ideas into the HIV model, by distinguishing three states in which agents may find themselves: healthy, infected, and infected with treatment.[11] The dynamics between the three states are simulated based on individual decisions that choose to participate in one of three alternative "markets," or "meeting places," respectively for single-partner long-term sex, casual sex with condoms, or casual unprotected sex. This segregation of market structures eliminates any complexities due to differences in the interests of partners: they have the same intentions when they enter the same market. Finding a partner generates utility from sexual behavior. Marriage has the additional benefit of continued interaction without the need to search again.

There are four types of status for each individual, and these are labeled abstinence, long-term sexual relationship, short-term unprotected sex, and short-term protected sex. In each of these states people can be susceptible to HIV, infected with no treatment and infected with treatment. The health status is only known to the individual and it cannot be observed by the sexual partners.[12] Infected and susceptible people choose rationally the submarket to enter by optimally choosing the odds of finding partners in each of the three sexual submarkets (π_l, π_p, π_u where subscript l, p, u refer respectively to the odds of finding

Uncorrected Proof

a long-term relationship, a short-term protected one, and a short-term unprotected one). Searching in a market has both a convex effort cost and the cost of possibly contracting HIV. The odds are chosen on the basis of convex cost functions $C_i(\pi_i)$ where $i \in \{l, p, u\}$. While π_l, π_p, π_u represent the probabilities of an individual finding a partner in the three submarkets, they also represent the fractions of people searching in each market that will find a partner, given the large size of the economy.

The transition rates as well as the infection rates are thus fully endogenous and respond to economic incentives, in line with the economic models outlined above. In equilibrium, each market is characterized by its riskiness and by a transfer that one partner could be making to the other. These transfers clear the market on the basis of the preferences of the potential partners. The equilibrium is characterized by an adverse selection problem: Individuals with a tendency toward risky sexual behavior enter the market for casual unprotected sex. As a consequence, this market tends to have a high rate of HIV incidence. Healthy individuals who do not have a strong preference for risky sexual behavior are further discouraged from entering this market because of the heavier concentration of infected individuals. This exacerbates the riskiness of the market for casual unprotected sex. The model is solved numerically and calibrated to match key moments of HIV and sexual behavior in Malawi. The calibration matches most targeted moments well, even though the model is sparse on gender differences.

2.2 Behavioral Influenza, SIR, and the Pre-COVID Contact Function

Before the emergence of COVID-19, the economics profession had only just began to propose behavioral SIR models, in which all three states, including that of recovery, were explicitly modeled and long-run equilibrium led to herd immunity, in line with Kermack and McKendrick (1927). These contributions were mostly published in journals specializing in theoretical and computational biology, so they remained unknown to the mainstream economics profession. Reluga (2010), Chen et al. (2011), and Chen (2012) are important contributions in the modeling of influenza epidemics with explicit reference to the influence of contact functions in the SIR framework. Although the marriage between these contact functions and the economic incentives of the SAM approach was still a long way off, papers that appeared on either side of the millennium explicitly modeled matching environments in product and labor markets.[13]

A typical example is Reluga (2010), who introduced the idea of an aggregate function that depended on endogenously determined social distancing. His primary objective was to derive the dynamics of epidemics by making the social distancing decision dependent on the state of the disease, very much along the lines of post-COVID research. He defines social distancing as the adoption of behavior by individuals in a community which reduces the risk of infection by either reducing contact with other individuals or reducing the transmission risk during each contact. The paper also recognizes that social distancing incurs some costs in term of individual behavior, although there is no connection between these costs and withdrawal from specific economic activities. The choice of social

Uncorrected Proof

distancing follows the tradition of symmetric Nash equilibrium games, with individuals choosing action c_s by taking as given the population average $\overline{c}_s$. The dynamics of epidemics depend on aggregate social distancing through a function $\sigma(\overline{c}_s)$, which is defined as the relative risk of infection given an investment $\overline{c}_s$ in social distancing. If there is no investment, the relative risk is $\sigma(0) = 1$. Reluga (2010) argues that the function $\sigma(.)$ features diminishing returns with increasing investment, so that $\sigma(.)$ is convex. The key transmission rate is thus modeled as,

$$-\frac{dS(t)}{dt} = \beta\sigma\left(\overline{c}_s\right)I(t)S(t) \tag{4}$$

with the same notation as before. Reluga (2010) solves for a Nash equilibrium in $\overline{c}_s$ and shows that social distancing reduces the epidemic peak and prolongs the epidemic, a result that has featured in a number of COVID-19 papers.

Chen et al. (2011) and Chen (2012) built on Reluga (2010)'s work and focus mainly on more elaborate game-theoretic solutions from which they derive the aggregate dynamics of the disease. They use an aggregate contact function which bears several similarities to the aggregate matching function of SAM theory. Chen (2012) considers an SIR model in discrete time with a continuum of agents. At each point in time, an agent can be in one of three health states: susceptible, infected, and recovered. An infected agent recovers at the end of any period with probability $\rho \in [0, 1]$. A recovered agent is fully immune and remains so for the rest of life.

Assume that there is no entry or exit of agents so that the population size is constant. The behavior of agents is specified as follows. In every period, agents choose how much time $\alpha \geq 0$ to spend outside the home during a period, referred to as the level of public activity. Individuals' choices of their public activity levels affect the rate at which contacts occur in the population and so the rate at which an infectious disease spreads. Assume that α belongs to the interval $[0, \overline{\alpha}]$ where $\overline{\alpha} = 1$ can be interpreted as agents' "normal" public activity level in the absence of the infectious disease. As is standard in the literature, agents are self-interested and seek to maximize their own payoff, without regard to the payoffs of other agents. In addition, assume that, all else being equal, an agent prefers less public avoidance over more. Since recovered agents cannot be infected, they have no incentive to engage in public avoidance behavior; thus recovered agents always choose $\alpha = 1$ for their level of public activity. Similarly, self-interested infected agents would not choose to adopt public avoidance behavior. However, the state of being infected can be sufficiently debilitating to cause some infected agents to have to stay home. Let us assume that, in any period, the fraction $1 - \gamma$ of infected agents, where $\gamma \in [0, 1]$ are too sick to engage in any public activities. The remaining infected agents – those who have only mild symptoms – fully participate in public activities, i.e., $\alpha = 1$ for these infected agents.

Let us now consider the contact structure and the disease transmission process. A susceptible agent who chooses public activity level α_t at time t has probability $\alpha_t\lambda_t$ of being infected in that period, where λ_t denotes the probability of infection in period t per unit public activity level.

Unedited Proof

This probability λ_t is a function of the disease prevalence p_t (which is equal to infections) as well as the public activity levels chosen by all other agents at time t. Now, letting r_t denote the time t fraction of recovered agents, the mean level of public activity in the population at time t is $n_t = \gamma p_t + r_t + \alpha_t(1 - p_t - r_t)$, provided all susceptible agents choose public activity level a_t in period t. Suppose λ_t is specified as follows:

$$\lambda_t = \lambda(p_t, r_t, \alpha_t) = m(n_t)\frac{\beta\gamma p_t}{n_t}$$

where $\beta \in [0, 1]$ is the transmission probability, and m is the *meeting or contact function* that specifies the rate at which a susceptible agent encounters other agents per unit of public activity. Note that $\beta\gamma p_t/n_t$ is the probability of becoming infected when meeting another agent. Chen (2012) assumes that the function m is positive increasing, since encounters with others are more likely when other people spend more time out in public, i.e., $m(n) \geq 0$ for all $n \in [0, 1]$; he also assumes that $m(n)$ is increasing. If $m(n) \leq 1$ is can be interpreted as a probability of meeting another agent per unit of public activity. Chen (2012) solves for a Nash equilibrium in α^*. As it will become clear when we discuss the post-COVID literature, these papers' specifications of contacts parallel the assumptions made by papers that adopt the SAM matching function with increasing returns to scale, justified by the assumption that when other agents take part in public activities, meeting probabilities increase even if there is no additional effort from a given agent.

3. COVID-19 RESEARCH

With the outbreak of COVID-19 economists turned in large numbers to research on the dynamics of infections, the economic costs of the disease versus the personal costs of infections, and the optimal policy that might be followed by governments in the face of macro trade-offs. As we mentioned in Section 2, the economic interpretation of the epidemiological model can be formulated in terms of a SAM model in which the disease spreads from meetings between susceptible and infected agents. The recent economics literature on COVID-19 that uses this or a related approach is too large to survey in any detail. We discuss here a selection of papers that are either particularly relevant in terms of approach or that explicitly make use of the properties of SAM models to derive their results.

3.1 Optimal Control and Optimizing Behavior in Macro Models

A number of papers use optimal control theory to derive optimal policies in the pandemic. Alvarez et al. (2021) is a good representative paper. They take a variant of the SIR model and formulate an optimal control problem by considering policies familiar from the pandemic, in particular, lockdowns (L_t), quarantine (Q_t), and testing (T_t). A central planner chooses optimally at each point in time the stock of lockdown individuals (L_t) as well as the stock of quarantined

Uncorrected Proof

people (Q_t). These stocks are chosen to minimize an aggregate social welfare cost. The key flow from the susceptible to the infected state is modeled as

$$-\frac{dS(t)}{dt} = \beta\tilde{S}(S_t, L_t)\tilde{I}(I_t, L_t, Q_t), \tag{5}$$

where the number of susceptible people in public places $\tilde{S}(.)$ is a function of all susceptible people and the number engaged in lockdown L_t, and the number of infected individuals $\tilde{I}$ is a function of both those in lockdown and those in quarantine. Alvarez et al. (2021) write

$$-\frac{dS(t)}{dt} = \beta[S_t(1-\theta L_t)][(I_t - Q_t)(1-\theta L_t)] \tag{6}$$

where $\theta \in (0, 1]$ is a measure of the lockdown effectiveness. In this framework the stock of quarantined agents follows the law of motion $\dot{Q}_t = T_t - \gamma Q_t$ where $T_t \leq \overline{T}$ denotes the flow per unit of time of agents that are traced, tested (positive), and placed into quarantine, and $\overline{T}$ is a capacity constraint on the number of agents that can be traced per unit of time. The policy prescribes a severe lockdown beginning a few weeks after the outbreak, covering almost 50% of the population after a month, with a total duration of approximately four months.

Eichenbaum, Rebelo, and Trabandt (2021) and Kaplan, Moll, and Violante (2020) propose behavioral SIR models in which the reproduction rate of the disease depends on aggregate consumption and hours of work, motivated by the fact that the more there is of each, the more likely it is that there will be social contacts and infections. The trade-off studied is between health and GDP losses; cutting down on consumption and work improves aggregate health outcomes through lower infections, but involves GDP losses. Eichenbaum et al. (2021) argue that people can become infected in three ways. First, consumption requires shopping in social space. The number of newly infected people that results from shopping activities is $\pi_1(S_tC_t^s)(I_tC_t^s)$. The terms $S_tC_t^s$, $I_tC_t^s$ represent total consumption expenditures by susceptible and infected people, respectively. The parameter π_1 reflects both the amount of time spent shopping and the probability of becoming infected as a result of that activity. Second, susceptible and infected people can meet at work. The number of newly infected people that results from interactions at work is given by $\pi_2(S_tN_t^s)(I_tN_t^i)$. The terms $S_tN_t^s$ and $I_tN_t^i$ represent total hours worked by susceptible and infected people, respectively. The parameter π_2 reflects the probability of becoming infected as a result of work interactions. Third, susceptible and infected people can meet in ways not directly related to consuming or working. The number of random meetings between infected and susceptible people is S_tI_t. These meetings result in $\pi_3S_tI_t$ newly infected people. The total number of newly infected people is given by

$$-\frac{dS(t)}{dt} = \pi_1\left(S_tC_t^s\right)\left(I_tC_t^s\right) + \pi_2\left(S_tN_t^s\right)\left(I_tN_t^i\right) + \pi_3S_tI_t \tag{7}$$

Uncorrected Proof

In this formulation, the corresponding equation of the classical Kermack and McKendrick (1927) can be interpreted as special case in which the propagation of the disease is not related to economic activity ($\pi_1 = 0, \pi_2 = 0$). We note, however, that this approach does not allow changes in shopping or work practices that offset, partially or wholly, a higher probability of attracting the virus; for example., online shopping replacing the trip to the shop or working from home instead of the office. In this sense, their approach is likely to give a higher GDP cost than would be observed in a world where there are different shopping and work possibilities. In a variant of this approach, Kaplan et al. (2020) argue that the reproduction number is $\beta_t = \beta(C_{st}, L_{wt}, t)$, where C_{st} is aggregate social consumption and L_{wt} is aggregate workplace hours.

3.2 COVID-19 and Optimizing Behavior with Search: Homogeneous Agents

We turn now to research with optimizing agents, which took center stage in COVID-19 related research. Unlike HIV epidemics, COVID-19 is more appropriately modeled within a traditional SIR model in the spirit of Kermack and McKendrick (1927), or one of its variants.[14] Economists contributed to the literature in at least three directions. First, in line with some of the early pre-COVID research surveyed in section 2, most papers model individuals as rational forward-looking optimizing agents. Second, the transition probabilities between different states, and notably between the susceptible and infected status, are explicitly modeled with the techniques traditionally used in SAM literature. Third, most models solve both a decentralized equilibrium and a central planning problem, and derive several types of externalities, due mostly to the fact that, as in the SAM literature, agents tend to ignore various external effects related to their search activities.

In this section, we discuss models that derive the equilibrium dynamics of the disease for homogenous agents, focusing on the implications of optimizing behavior during social interactions. We discuss the deviations between the decentralized and central planning solutions after we consider the role of heterogeneity.

Garibaldi, Moen, and Pissarides (2020b) work in discrete time and model, the case in which the infected people are, at least for an initial period, asymptomatic. The implication of this assumption is that the social action chosen by susceptible and infected agents prior to the arrival of symptoms is the same. Given this common action, which is denoted by $\overline{x}_t$, a single vulnerable agent who chooses social action x_t transits from the susceptible to the infected state between periods t and $t+1$ with probability,

$$p_{t+1} = p\left(x_t, \overline{x}_t, x_r, S_t, I_t, R_t\right). \tag{8}$$

x_r are the choices of recovered agents, and S_t, I_t, and R_t denote the mass in each respective state (informally referred to as the number of people in the state). A key assumption is

Unedited Proof

$$\begin{aligned} \frac{\partial p(x_t,.)}{\partial x_t} &\geq 0, \\ p(0,.) &= 0, \end{aligned} \tag{9}$$

where $p(0,.)$ is the transition to infection in the state of complete social distancing.

Garibaldi et al. (2020b) claim that the contact technology that yields the infection probability $p(x_t,.)$ parallels the matching function of labor economics but with some important differences. In the matching function of the labor literature, more workers looking for jobs reduce the success probability of a single worker because of congestion externalities in the application process. In epidemic models, more individuals coming out in the marketplace increase the chances of infection because a single exposed individual can infect many people; the infectious disease is "non-exhaustible," in the sense that many people could acquire it from a single person at the same time.

Garibaldi et al. (2020b) provide an intuitive derivation of the contact function. Suppose $\overline{x}_t$ stands for the number of trips outside the house that each person does in a single period. They assume that with $\overline{x}_t$ trips, each person experiences on average m_t contacts, defined by a well-behaved function $m_t = m(\overline{x}_t)$, with $m^{'}(\overline{x}_t) \geq 0$. The function $m(.)$ is similar to the matching function of labor economics, in the sense that it depends on the structure of the marketplace, including density of population, transportation facilities, types of establishments, etc. Note that the dependence of $m(.)$ on a single variable parallels the contact function used by Diamond (1982) in his "coconut" model. With respect to the m_t contacts in the epidemic context, some of these contacts are between susceptible and infected people, which lead to the infection of the susceptible agent with some positive probability that depends on the infectiousness of the disease.

In order to derive the optimal x_t for a single optimizing agent, Garibaldi et al. (2020b) follow the method used in search theory to choose the optimal search intensity (Pissarides (2000), chapter 5). With $m(\overline{x}_t)$ representing the total number of contacts for $\overline{x}_t$ outings, each outing on average generates $m(\overline{x}_t)/\overline{x}_t$ contacts. So if the individual chooses to go out of the home x_t times, her contacts are on average $x_t m(\overline{x}_t)/\overline{x}_t$. Of these, there is an infection with probability βI_t, where β as before is the infectiousness of the disease and I_t is the fraction of infected individuals. It follows that the transition probability from the susceptible to the infected state for the person who chooses x_t is,

$$p_{t+1} = \beta \frac{x_t m\left(\overline{x}_t\right)}{\overline{x}_t} I_t. \tag{10}$$

In an economy without an infectious disease, agents allocate their time between noninteractive activities and social activities x_t, given their utility functions and costs. During an epidemic, choosing social activities involves a probabilistic cost summarized in (10), which is the cost of an infection. In decentralized equilibrium, the typical agent therefore chooses a smaller x_t,

Uncorrected Proof

shifting activities to noninteractive pursuits, such as working at home, buying consumption goods online, and cooking at home instead of going to restaurants. The difference between the x_t chosen in the absence of an epidemic and the x_t chosen in an epidemic is defined as the social distancing that agents choose during the epidemic.

In moving from individual transitions to the average for a market where all agents optimize, Garibaldi et al. (2020b) assume a symmetric Nash equilibrium in which all agents choose the same policy, so $x_t = \overline{x}_t$. With S_t susceptible agents choosing social action $\overline{x}_t$, the transition from the susceptible to the infected state is therefore given by,

$$p_{t+1} S_t = \beta m\left(\overline{x}_t\right) I_t S_t. \tag{11}$$

A question that arises and which is important for the welfare analysis is the degree of homogeneity of the meeting technology. Some of the early literature assumed homogeneity less than one in large markets (Chen et al., 2011), whereas the labor literature converged on constant returns (Petrongolo & Pissarides, 2001). However, when there are unintended contacts in fixed social space, increasing returns are a more plausible assumption, driven by the fact that a given individual has no full control over chance meetings in social space (refer again to the discussion at the beginning of section 2). Intuitively, the justification for increasing returns is similar to the one used by Diamond (1982). In that paper, islanders possess a coconut which they acquire by climbing a tree but they cannot consume their own coconut. They have to find another islander with a coconut and swap nuts. Diamond's claim was that if the number of islanders climbing trees doubled, a passive islander was more likely to come out and climb a tree because the probability of finding a trade would be higher: a positive externality. Subsequent work did not find support for this claim in labor or goods markets, because as both buyers and sellers double in number they create congestion for each other and so many swaps are crowded out. In the context of an epidemic it is precisely this congestion that justifies the increasing returns because of the non-exhaustive nature of the disease. An infected agent can pass a disease to a very large number of people but in Diamond's example they can only give their coconut to one person. Diamond's intuition for increasing returns applies to this model much more than in a model of exchange.

Farboodi, Jarosch, and Shimer (2020) use a contact functions similar to Garibaldi, Moen, and Pissarides (2020a), and discuss the modeling of the contact function going back to Diamond and Maskin's (1979) seminal introduction of the distinction between a quadratic and a linear matching technology. With quadratic matching, additional social activity by others raises the likelihood of social contact and thus the disease transmission for all individuals. For example, with more individuals in parks, restaurants, and public transport, any given trip to a park/restaurant/subway is more likely to lead to disease. As we discuss also in Section 3.4, such a matching function has a search externality that traditionally is

Uncorrected Proof

viewed as positive (Diamond, 1982) but that turns negative in the context of disease.

3.3 COVID-19 and Optimizing Behavior with Search: Heterogeneous Agents

A well-documented feature of COVID-19 is the differential impact of the disease on individuals of different age or preexisting medical conditions. Given the importance of age, and the ease of observing it, several models worked out equilibrium outcomes with heterogeneous agents, distinguishing them across the age spectrum. Within the SAM tradition, Acemoglu et al. (2021) and Brotherhood, Kircher, Santos, and Terlit (2020) are key contributions.

Acemoglu et al. (2021) propose a multigroup SIR (MG-SIR) with three age groups, young, middle-aged, and old. The model is solved in the spirit of optimal control as proposed by Alvarez et al. (2021) and derives the optimal lockdown across the three groups. Let θ_j be the intensity of lockdown for each group, η_j the probability that an infected person fails to be isolated, and ρ_{jk} the contact rate across groups. As a result, infections for group j evolve according to,

$$-\frac{dS_j(t)}{dt} = \beta\left(1-\theta_j L_j\right) S_j \sum_k \rho_{jk}\eta_{jk}\left(1-\theta_k L_k\right) \tag{12}$$

In Acemoglu, Chernozhukov, Werning, and Whinston (2020)'s interpretation, (12) is the classic law of motion of SIR models, assuming a quadratic matching technology. In Acemoglu et al. (2021), a more general nonquadratic contact function M_j is considered so that

$$-\frac{dS_j(t)}{dt} = M_j\left(1-\theta_j L_j\right) S_j \sum_k \beta_{jk}\eta_{jk}\left(1-\theta_k L_k\right) \tag{13}$$

where

$$M_j(S,I,R,L) = \left(\sum_k \beta_{jk}\left[\left(S_k+\eta_k I_k+(1-\kappa_k)R_k\left(1-\theta_j L_k\right)\right)+\kappa_k R_k\right]\right)^{\alpha-2}$$

Note that if $\alpha = 2$ then $M = 1$ and also R_k drops out of the equation, so only S and I matter. In addition, with a single group the equation reduces to $-\frac{dS_j(t)}{dt} = \beta SI(1-\theta L)^2$, an expression identical to Alvarez et al. (2021). The quantitative analysis of optimal policies is applied to the United States. Acemoglu et al. (2021) find that optimal policies differentially targeting risk/age groups quantitatively outperform optimal uniform policies and most of the gains can be realized by having stricter lockdown policies on the oldest group. A strict and long lockdown for the most vulnerable group both reduces infections and enables less strict lockdown for the lower-risk groups, which reduces the economic cost. Favero et al. (2020), in a quantitative model applied to two Italian regions with a very large number of age differences, reach similar conclusions,

Uncorrected Proof

even though the modeling of individual behavior and transmission rates are only sketched.

Brotherhood et al. (2020) propose a model with age heterogeneity and focus on optimal policies. Age is indexed by $a \in \{y, o\}$ where y is young and o is old. The modeling of infection does not relate directly to a matching or contact function in the spirit of Garibaldi et al. (2020a), yet in Brotherhood et al. (2020) the key transmission risk $\Pi_t(a)$ is a time-varying function that depends on the number of infected people and how much time these people spend in social space. In equilibrium $n_t(j, \tilde{a})$ and $l_t(j, \tilde{a})$ denote the measure of agents of each type j of age $\tilde{a}$ who are outside the home for work or leisure. The aggregate probability of getting infected for a fraction of the period spent outside and given an exogenous SI transmission rate Π_0 (a parameter specific to the disease) is

$$\widehat{\Pi}_t = \Pi_0 \sum_{\tilde{a},j \in \{f_i,i,h\}} \left(n_t\left(j, \tilde{a}\right) + l_t\left(j, \tilde{a}\right) \right) M_t\left(j, \tilde{a}\right) \tag{14}$$

where each individual j can be in a state of infection (i), hospital (h) or fever and infected (f_i), and $M_t(.)$ is a law of motion that maps from the state vector, the equilibrium actions and the infection rates in period t to the number of agents of each type M_{t+1} in the next period.[15] Using this setting, Brotherhood et al. (2020) find that older individuals socially distance themselves substantially in equilibrium. Thus, the optimal lockdown is binding most for the young.

3.4 Externalities in the SI and SIR Epidemic Models

Transmission of a disease is an involuntary and often unobserved event, with no money transfers involved. It is usually the outcome of a contact initiated for other reasons. In such situations, it is reasonable to expect that the decentralized solution will suffer from externalities, e.g., because infected agents ignore the impact of their actions on others. Some of these externalities are associated with the SAM framework as such, while others are associated with the inherent dynamics of the epidemiological models. We survey here the externalities involved, by making the usual assumption of decentralized equilibrium models: that the agents in the economy do not have altruistic preferences but act to maximize their own utility. If the agents do have altruistic preferences, this may reduce the externalities, as the agents partly internalize the negative impact of their actions on fellow citizens.

In the context of pre-COVID research, Toxvaerd (2019) introduces policy dimensions, welfare and externalities in the context of the models proposed by Geoffard and Philipson (1996) and Kremer (1996). Toxvaerd (2019) shows that a permanent decrease in the infectiousness of the disease will prompt an increase in exposure and in steady-state disease prevalence. Although the decrease in infectiousness decreases the rate of transmission per exposure, the exposure itself increases sufficiently to lead to an increase in steady-state disease prevalence. This outcome is not pathological and holds both for a utilitarian social planner and for self-interested individuals.

Uncorrected Proof

Related to this, under certain conditions, a decrease in the infectiousness of the disease may lead to a decrease in social welfare. The reason is that individuals make different choices than those preferred by a social planner. There are two sources for this discrepancy. First, there is a pure externality effect that arises because individuals do not internalize the benefits to others that flow from the individual protection of themselves. As a consequence, aggregate equilibrium protection is too low, thus causing higher future disease prevalence. This implies that the equilibrium future path of infection faced by the individual is higher than the path preferred (and indeed chosen) by a social planner. The second discrepancy is related to the fact that each individual takes the path of aggregate infection as given, while the planner takes into account the fact that the aggregate path is made up of the sum of individual paths. These two externalities are in the spirit of what the COVID-19 literature called static and dynamic externalities, to which we turn next.

To simplify, let us consider a model with homogenous agents. Garibaldi et al. (2020b) classify externalities into four categories: Contact externalities, medical externalities, contagion externalities, and immunity externalities.[16] The contact externalities are most closely related to the SAM literature. A well-known result from the labor-search literature is that if there are increasing returns to scale in the matching function, this gives rise to search externalities and possibly multiple equilibria. The reason is that the search activities of different agents are *strategic complements*. Take the Diamond (1982) model as an example, where identical agents search to find trading partners. Due to increasing returns to scale in the meeting technology, the arrival rate of a trading partner per unit of search intensity of an individual depends positively on the average search intensity in the economy. Hence, if the other agents in the economy increase their search intensity, the incentives for a given agent to search increases, and this strategic complementarity leads to below-optimal social action and may also lead to multiple equilibria.

Similar mechanisms take place in the SIR-matching model. However, since attracting the virus has a negative impact on lifetime utility, increasing returns give rise to substitutability, not complementarity, in the agents' choice of social activity. To see this, suppose $m(x_t) = x_t^{\alpha}$, $\alpha > 0$. From (Eq. 10) we get that

$$p_{t+1} = \beta x_t \overline{x}_t^{\alpha - 1} I_t \tag{15}$$

Hence, for $\alpha > 1$, a higher average activity level will tend to increase the marginal effect of own activity level on the probability of attracting the virus. Everything else equal, this will tend to reduce the individual activity level. Hence, in contrast to the findings in Diamond (1982), increasing returns to scale in the contact function will not lead to multiple equilibria, as activity levels among the individuals become strategic substitutes rather than complements.

Still, increasing returns to scale gives rise to a negative meeting externality. If a (small group) of agents increase their activity, this increases the average activity level, and hence increases the probability that the other agents attract the virus. This negative effect will be internalized by the planner but not by the agents in the

Uncorrected Proof

economy. This is the essence of the contact externality and the best policy response is for the planner to order more social shielding.

The contact externality is related to, but different from, the contagion externality. The contagion externality is the "classical" externality that arises in models of infectious diseases: If a person is infected, they may infect others in subsequent periods, something a nonaltruistic person will not take into account.[17] This is a powerful externality, particularly when a vaccine is expected to arrive that will protect susceptible agents from the disease. In the pre-COVID literature, the efficiency role of a vaccine is highlighted in the influenza literature, and notably in Reluga (2010), even though he does not introduce an explicit welfare function. Conversely, the HIV economics literature that deals with welfare analysis (Greenwood et al., 2019; Toxvaerd, 2019) is less concerned with vaccination and focuses on policies aimed at reducing the risk of infection during sexual activity.

The medical externality naturally relates to the medical sector (Farboodi et al., 2020). With fixed capacity in this sector, at least in the short run, when the disease is spreading fast, the quality of treatment will fall as the number of agents who are infected and hospitalized increases and reaches hospital capacity levels. This will lead to poorer care, increased medical expenditures per person, and more importantly to higher death rates. Optimizing individuals take into account the individual cost of hospitalization and the fatality rate, but not the fact that by becoming ill, they increase the cost of hospitalization and the death risk for other people. A common slogan in the British lockdowns during the COVID-19 pandemic was "stay home, save the NHS," precisely trying to induce altruism with respect to medical congestion. A good policy is one that spreads infections more evenly over time, usually achieved by imposing strict lockdown in the beginning of an epidemic, when typically infections rise very fast.

Finally, the immunity externality, introduced by Garibaldi et al. (2020b) and further discussed by Brotherhood et al. (2020), is the most "hidden" of the externalities. It is, however, important in the absence of effective vaccines, in which case society has to rely on herd immunity to get back to normality. Herd immunity is achieved when a sufficiently large number of agents contract the virus and recover with immunity. In the absence of many remaining susceptible agents to visit, the virus is eradicated. It follows that there is a positive externality to catching the disease (assuming immunity on recovery): by becoming immune an agent helps society get closer to herd immunity, to the benefit of those who are still free of the disease.[18] This positive externality is ignored by individuals maximizing their own net worth. In contrast, a social planner will want to internalize it, by encouraging more social activity.[19]

Another way of motivating the immunity externality is in terms of a "rat race to shield." Forward-looking agents will know that in order to reach herd immunity, a certain fraction of the population will have to be infected. It is therefore optimal for someone to shield until someone else gets infected and helps drive the economy to herd immunity. But if all shield, herd immunity will never be reached. The optimal policy of the planner is to make sure there is no excessive shielding, and in some simulations this policy dominates the contagion

Uncorrected Proof

externality. But as this response is critically dependent on herd immunity as the only path to eradication of the disease, it collapses if another path to eradication is discovered, as in the case of an effective vaccine. In the latter case, it is socially optimal to shield more to save more people from the illness, until enough people get vaccinated to reach herd immunity that way.

Finally, consider briefly the implications of different social policies for people with different experiences of the disease. Recovered people with immunity are safe social contacts, and this is another positive externality that is ignored by them. If they know they cannot contract the disease, their optimal policy will be to maximize their utility with respect to the allocation of time to domestic and market activities, ignoring the fact that they don't infect other people in market activities. A social planner would take this externality into account and request more social contact from recovered people, e.g., provide incentives for them to take on extra social work on production of evidence of recovery.

4. CONCLUSIONS

Before the COVID-19 pandemic, economists paid limited attention to the modeling of epidemics, such as influenza or HIV. Epidemiologists and mathematical biologists created an extensive literature which addressed issues of transmission and herd immunity, as well as optimal government policy in response to the dynamic path of the disease, but paid limited attention to the modeling of individual incentives to shield and the economic costs involved. Of course, notable exceptions exist, and one of the purposes of this chapter was to review the pre-COVID literature that addressed economic trade-offs and optimal policy.

Economic modeling of epidemics took off, at unprecedented speed, soon after the COVID-19 pandemic took hold of countries worldwide. CEPR in its *Covid Economics* series published about 1,000 papers in less than two years. It is doubtful whether there has ever been an event that changed the economics literature so dramatically in such a short space of time. These papers addressed mainly the trade-off between the economic losses from social distancing and the gains in disease control. Our approach in this chapter focused on the review of models that used the ideas that were developed independently in the SAM literature about the meeting of agents in social space as a precondition for exchange. We have shown that these ideas lead to a neat modeling of the incentives that agents have for social distancing, given the new element that is introduced by an epidemic: the risk of infection, which translates into an economic cost in future periods. Put in these terms, the new trade-off in an epidemic is between a current cost when an agent refrains from social contact to avoid infection versus a future cost that involves, with positive probability, a medical cost and withdrawal from social contact as symptoms from a possible infection emerge.

We have shown that the introduction of economic incentives associated with the new trade-offs has a large impact on the dynamic evolution of the disease, both in decentralized equilibrium and in central planning solutions. Approaching

Uncorrected Proof

the epidemic problem in this way unearths different kinds of externalities that exist in these models and which require different policy responses depending on assumptions made about medical differences between agents, vaccinations, and the degree of infectiousness of the disease, among many others.

NOTES

1. Of course, there is also a lot of social interaction in economics that does not fit this description. For example, the provision of personal services, such as haircuts or restaurant meals.

2. See, for example, Pissarides (2000), Mortensen and Pissarides (1999) and Diamond and Maskin (1979).

3. For a discussion of more recent work and modeling of the mathematical equations, see Hetcote (2000) and Weiss (2013).

4. See Petrongolo and Pissarides (2001) for a survey of the matching function as used by labor economists.

5. Diamond and Maskin (1979) study the implications of linear and quadratic "search technologies" in economic interactions.

6. See, for example, Eichenbaum et al. (2021) and Farboodi et al. (2020), who borrow Diamond and Maskin (1979)'s "quadratic matching technology" or Garibaldi, Moen, and Pissarides (2000a) who invoke a more traditional aggregate matching function, with both linear-homogenous and quadratic forms as alternatives. We discuss several other references and the implications of their matching assumptions in the main text of the paper.

7. The parallel between employment contacts and marriage (more generally, mating) has been noted and explored in a number of papers. See, for example, Mortensen (1982) and Burdett and Coles (1997).

8. Note that in this basic model there are no deaths from the disease.

9. Chen (2004) extends the ideas of Geoffard and Philipson (1996) by introducing entry and exit of agents.

10. Despite the dynamic nature of Eq. (3), Kremer (1996) studies only the steady state of the epidemic.

11. Directed or competitive search was introduced by Moen (1997). See Wright, Kircher, Juline, and Guerrieri (2021) for a recent survey.

12. Greenwood et al. (2019) assume that in each infected status there is a probability of moving into a final stage of HIV where the health status becomes observable.

13. For a detailed exposition of Kermack and McKendrick (1927) and its dynamic behavior, see Hetcote (2000) and Weiss (2013). With respect to the dynamics of the model, Weiss (2013) writes "Mass action mixing assumes that the rate of encounter between susceptible and infected individuals is proportional to the product of population sizes. . ..This requires that the members of both populations are homogeneously distributed in space and thus . . .every person will encounter every other person per unit time with equal probability." He also argues that it is possible to formulate a stochastic analog of the SIR model as a Markov chain. The matching interpretation that we give to the model yields such an approach in response to the random meetings and the influence that economic incentives have on the infections rate.

14. We note that the insights related to SAM come from the transition from susceptible to infected, so the important states are SI. Results are broadly similar for models such as SIR, SIS, or other variants. In the case of COVID-19, although it was believed at first that recovery brought immunity, experience with the disease did not corroborate this belief, and an SIRS model seems to be more appropriate. See, for example, Giannitsarou, Kissler, and Toxvaerd (2021).

15. Brotherhood et al. (2020) write
$M_{t+1} = T(M_t, N_t, \Pi_t(o), \Pi_t(y))$
as the equilibrium allocation of the functions $n_t(j, \tilde{a})$ and $l_t(j, \tilde{a})$.

Uncorrected Proof

16. Brotherhood et al. (2020) and Farboodi et al. (2020) also discuss the first three externalities.

17. Some of the infected persons may be within the person's household. To the extent that this is the case, the assumption of altruistic preferences may be plausible.

18. In an early review of the epidemiological literature, Fine (1993) noted that herd immunity is "the indirect protection afforded to nonimmune individuals by the presence and proximity of others who are immune." This definition hints at the immunity externality as we defined it here, in the sense that "indirect protection" is ignored by nonaltruistic agents. The reference to "proximity" also hints at another positive externality of immunity that is briefly discussed at the end of this section. We are grateful to an anonymous referee for directing us to this reference.

19. There is some informal evidence that soon after the start of infections in Britain, this was the favored policy. For example, the Prime Minister's advisor at the time, Dominic Cummings, stated that herd immunity was government policy in the early stages of the pandemic but it quickly changed to shielding when the first deaths took place. See https://www.bbc.co.uk/news/57229390.

REFERENCES

Acemoglu, D., Chernozhukov, V., Werning, I., & Whinston, M. D. (2020). *A multi-risk sir model with optimally targeted lockdown*. NBER Working Paper N. 27102.

Acemoglu, D., Chernozhukov, V., Werning, I., & Whinston, M. D. (2021). A multi-risk sir model with optimally targeted lockdown. *The American Economic Review: Insights*, forthcoming.

Alvarez, F., Argente, D., & Lippi, F. (2021). A simple planning problem for covid-19 lockdown. *The American Economic Review: Insights*, forthcoming.

Brotherhood, L., Kircher, P., Santos, C., & Terlit, M. (2020). *An economic model of the Covid-19 epidemic: The importance of testing and age-specific policies*. CEPR Discussion Paper N. 14695.

Burdett, K., & Coles, M. (1997). Marriage and class. *Quarterly Journal of Economics*, *112*(1), 141–167.

Chen, F. (2004). Rational behavioral response and the transmission of STDs. *Theoretical Population Biology*, *66*, 307–316.

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

Chen, F., Miahoua, J., Rabidoux, S., & Robinson, S. (2011). Public avoidance and epidemics: Insights from an economic model. *Journal of Theoretical Biology*, *278*, 107–119.

Diamond, P. (1982). Aggregate demand management in search equilibrium. *Journal of Political Economy*, *90*(5), 881–994.

Diamond, P., & Maskin, E. (1979). An equilibrium analysis of search and breach of contracts, I: Steady states. *The Bell Journal of Economics*, *10*, 282–316.

Eichenbaum, M. S., Rebelo, S., & Trabandt, M. (2021). The macroeconomics of epidemics. *Review of Financial Studies*, *34*(11), 5149–5187.

Farboodi, M., Jarosch, G., & Shimer, R. (2020). *Internal and external effects of social distancing in a pandemic*. NBER Working Paper N. 27059.

Favero, C., Ichino, A., & Rustichini, A. (2020). *Restarting the economy while saving lives under COVID-19*. CEPR Discussion Paper N. 14695.

Fine, P. (1993). Herd immunity: History, theory, practice. *Epidemiological Reviews*, *15*(2), 265–302.

Garibaldi, P., Moen, E. R., & Pissarides, C. A. (2020a). Modelling contacts and transition in the SIR epidemic model. *Covid Economics: Vetted and Real Time Papers*, *5*, 1–20.

Garibaldi, P., Moen, E. R., & Pissarides, C. A. (2020b). *Static and dynamic inefficiencies in an optimizing model of the epidemic*. CEPR Working Paper N. 15439; IZA Discussion Paper 13844.

Geoffard, P.-Y., & Philipson, T. (1996). Rational epidemics and their public control. *International Economic Review*, *37*(3), 603–624.

Giannitsarou, C., Kissler, S., & Toxvaerd, F. (2021). Waning immunity and the second wave: Some projections for SARS-coV-2. *The American Economic Review: Insights*, *3*(3), 321–338.

Uncorrected Proof

Greenwood, J., Kircher, P., Santos, C., & Tertil, M. t. (2019). An equilibrium model of the African HIV/AIDS epidemic. *Econometrica*, *87*(4), 1081–1113.

Hetcote, H. W. (2000). The mathematics of infectious disease. *SIAM Review*, *44*(4), 599–653.

Kaplan, G., Moll, B., & Violante, G. (2020). *The great lockdown and the big stimulus: Tracing the pandemic possibility frontier for the U.S*. NBER Working Paper N. 27794.

Kermack, W. O., & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceeding of the Royal Society of London, Serie A Containing Papers of Mathematics and Physical Character*, *115*(772), 700–721.

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

Moen, E. (1997). Competitive search equilibrium. *Journal of Political Economy*, *105*, 385–411.

Mortensen, D. (1982). Property rights and efficiency in mating, racing and related games. *The American Economic Review*, *37*(5), 968–979.

Mortensen, D., & Pissarides, C. A. (1999). New developments in models of search in the labor market. In O. Ashenfelter & D. Card (Eds.), *Handbook for labor economics*. Amsterdam: Elsevier.

Petrongolo, B., & Pissarides, C. A. (2001). A survey of the matching. *Journal of Economic Literature*, *39*(2), 390–431.

Pissarides, C. A. (2000). *Equilibrium unemployment theory*. Cambridge, MA: MIT Press.

Reluga, T. C. (2010). Game theory of social distancing in response to an epidemic. *PLoS Computational Biology*, *6*(5), 1–9.

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

Weiss, H. (2013). The SIR model and the foundations of public health. *Materials Matematics*, *17*(5), 1–17.

Wright, R., Kircher, P., Juline, B., & Guerrieri, V. (2021). Directed search and competitive search equilibrium: A guided tour. *Journal of Economic Literature*, *59*(1), 90–148.

Uncorrected Proof

Uncorrected Proof

Author Query Form

Queries and/or remarks

[Q1]	Please check the edit(s) made in the sentence ‘First, there is a pure externality effect …’, and correct if necessary.
[Q2]	Please provide the expansion of "SIS"
[Q3]	Please advice if "SIRS" can be retained or changed to "SIR"
[Q4]	Please provide volume number and page range for the reference "Alvarez et al., 2021."

Uncorrected Proof

RLEC50 CH001

Non Print Items

Pietro Garibaldi
University of Torino, Italy

Espen R. Moen
Norwegian Business School, Norway

Christopher A. Pissarides
London School of Economics, UK

Uncorrected Proof