I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13844

Static and Dynamic Inefficiencies in an Optimizing Model of Epidemics

Pietro Garibaldi
Espen R. Moen
Christopher A. Pissarides

NOVEMBER 2020

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13844

Static and Dynamic Inefficiencies in an Optimizing Model of Epidemics

Pietro Garibaldi
Collegio Carlo Alberto, University of Torino, IZA and CEPR

Espen R. Moen
Norwegian Business School and CEPR

Christopher A. Pissarides
LSE, University of Cyprus, CEPR and IZA

NOVEMBER 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Static and Dynamic Inefficiencies in an Optimizing Model of Epidemics*

In an optimizing model of epidemics several externalities arise when agents shield to avoid infection. Optimizing behaviour delays herd immunity but also reduces overall infections to approximately the minimum consistent with herd immunity. For reasonable parameter values, and with no vaccine, we find that agents delay too much because of a "rat race to shield": they shield too much in the hope that others catch the disease and reach herd immunity. This and other externalities drive large wedges between private and social outcomes. The expectation of a vaccine reverses the effects, and agents shield too little.

JEL Classification: A12, I10, J18, D61, D62

Keywords: SIR models, matching model, COVID-19, social distancing, rat race, herd immunity

Corresponding author:
Pietro Garibaldi
Collegio Carlo Alberto
University of Torino
Via Real Collegio 30
10024 Moncalieri (Torino)
Italy
E-mail: pietro.garibaldi@carloalberto.org

* Research support from Collegio Carlo Alberto is gratefully acknowledged. We thank Per August Moen for excellent research assistance, and Espen Henriksen for advise on the computational methods. We thank online seminar participants at CEPR, Polytechnic of Torino, Collegio Carlo Alberto, Oslo Macro Group, University of Chicago Search Conference, Marseille, Norwegian Business School and the Frisch Centre.

1 Introduction

In this paper we model the transitions in an optimal forward-looking model of an epidemic, in the three-state SIR framework originally proposed by Kermack and McKendrick (1927). We focus on the consequences of optimal behaviour for infections and herd immunity, and the externalities that arise when private agents act optimally to shield themselves from the epidemic. Our model and results apply generally to the SIR model and its subsequent elaborations, and they are not specific to COVID-19.[1]

Following our derivation of the optimal transition rates, we make two comparisons. First, we compare the outcomes in the original non-optimizing SIR model and in our decentralized optimal model. Our second comparison is between the decentralized optimizing model and a social planning solution, with the social planner having access to the same information set as private agents. We obtain striking contrasts in both cases which we derive formally and illustrate with simulations.

In our formulation agents can be in one of four states, but by applying a modelling trick we collapse the model to one that is close to the original three-state SIR model. We work in discrete time. In the first state a mass S_t of agents are healthy but "susceptible" to the disease; in the second state a mass I_t of agents are "infected" but without symptoms, and they can pass the disease on to susceptible individuals with social contacts; in the third state the I_t infected individuals develop symptoms, which are costly in terms of lifetime utility but not time; they cross from the asymptomatic infected state in period t to recovery in period $t+1$ by bearing some cost which is a fraction of their lifetime utility. We associate this cost with "hospitalization" and refer to it as the medical cost. Recovered individuals become immune for the rest of their life.[2]

Our interest is in deriving the impact that agents have on the transitions across the SIR states. They respond optimally to the epidemic, with perfect foresight about the future. With the four states that we described there are three transitions. The last two, from the asymptomatic to the symptomatic infected state and from the symptomatic state to recovery, cannot be influenced by agents' actions. The first is a medical transition and the second depends on treatments which cannot be influenced by the patient. The transition on which we focus is the one from the susceptible to the asymptomatic infected state, which is influenced by contacts between the susceptible and the infected. It is influenced by decisions made by both sets of agents. The decisions are taken without information about their state (whether they are susceptible or asymptomatic) and with full knowledge of future transitions in the event of an infection. For

[1] In earlier work (Garibaldi Moen and Pissarides, 2020) we focused on optimal policy in the COVID-19 epidemic.

[2] We ignore deaths, but taking them into account would have no impact on our results. In some epidemics recovered individuals may not become immune for life, which we assume here. Relaxation of this assumption would have an impact on our results, at least quantitatively. In the case of COVID-19 it is still not certain that there is immunity after recovery.

simplicity of exposition we call individuals who are in either the susceptible state or the asymptomatic infected state, "vulnerable", so the mass of vulnerable individuals is $V_t = S_t + I_t$.

We borrow ideas from search and matching theory (Pissarides, 2000) and show that there is a well-defined solution in line with the solutions in the literature that followed Kermack and McKendrick's (1927) pioneering work. The main difference between our model and those in the epidemiological literature is that agents in our model are able to reduce their probability of infection by avoiding contacts with vulnerable individuals. We model behaviour in such a way that this difference is picked up by a single variable x_t, which we call social action. We normalize social action in the epidemiological literature model by $x_t \equiv 1$ in all periods of life. In our model we derive an optimal $x_t \leq 1$ which varies over time, depending on incentives. When $x_t < 1$ we say that there is *social distancing*, or *shielding*.

Returning now to the two main comparisons that we make in this paper, we first show that in our model the time that it takes to reach herd immunity is much longer than in the original SIR model; but the number of people that avoid infection altogether is much higher. With plausible parameter values we show that in the original SIR model herd immunity is reached after about 20 weeks whereas in our model the time is closer to 80 weeks. But in the traditional SIR model only 3.6% of the people avoid the disease whereas in our model as many as 40% avoid it, a number that is close to the maximum possible consistent with herd immunity. The simulations show that there is a clear trade-off between the length of time that it takes to reach herd immunity and the cumulative number of infections, and with the high costs of the disease that we consider plausible, optimizing agents choose much longer adjustment paths than would emerge from a standard SIR model.[3]

But the adjustment path that results from optimizing decentralized choices is not optimal. There are two consequences of private actions that are ignored by agents, one operating across space and the other over time. Across space, individuals ignore the fact that when they take social action, other people may randomly come into contact with them and be infected. This gives rise to a *static externality*. We show that under the plausible restriction of increasing returns in the contact function, defined here as an elasticity of contacts with respect to social action greater than one, static externalities make the social planner recommend more social distancing. But with constant returns (unit elasticity), the static externalities are internalized.

Over time, when private agents shield they reduce the future infected population and increase the future susceptible population. These relative population changes alter the flows across states: the lower infected pool lowers future infections and lowers congestion in hospitals, whereas the higher susceptible pool slows down the adjustment

[3]These trade-offs are derived for a recovery (medical) cost that is set equal to 10% of the remaining lifetime income. Including deaths would obviously be a major factor in the calculations of disease costs.

to herd immunity. Private agents ignore these secondary effects of their actions whereas the social planner takes them into account. We call these *dynamic externalities.* The result is a divergence between the decentralized optimum and the social optimum that can be large. Because of these externalities, the social planner may recommend more or less social distancing, depending on parameters.

By working in discrete time and deriving the optimal policies from Bellman equations, we are able to distinguish between the *contagion* externality, caused by the impact of private actions on the stock of infected individuals, the *medical congestion* externality, caused by the same stock of agents when they go through the hospitalization phase, and the *immunity* externality, caused by the impact of private actions on the stock of susceptible agents. The externalities interact with each other to produce deviations between the private and social outcomes that can go either way and can change in the course of the epidemic. We illustrate these interactions with two important cases of deviation, one that we call the *rat race to shield* and one that arises when there is an expectation with sufficiently high probability that an effective vaccine will be discovered.

Consider first decisions made in the absence of a vaccine, when herd immunity is the only path to eradication of the disease. Because of forward-looking behaviour, agents know that eventually the economy will reach the herd immunity state with a fraction of people who never experience the disease. With positive disease costs, they shield more than other considerations would imply, to increase the probability that they belong to the infection-free group when the disease is eradicated. This has the features of a rat race, because if, say, 60% of the people need to get the disease before herd immunity, an efficient mechanism would be to ignore its impact on social distancing and allocate the 60% randomly across the population. But instead of racing to be first as in the traditional rat race, here there is racing to be the last, by shielding. When medical costs are sufficiently high (as in our simulations of 10% of remaining lifetime utility), the "rat race to shield" drives a large wedge between the private and social optima, which dominates deviations due to other externalities. The social planner offsets the rat race to shield by recommending much more social activity than private agents choose to do. The contagion externality is of course still present, and for very small medical costs it could dominate the immunity externality. But in our simulations with plausible costs of the disease the rat race to shield emerges as a strong externality in the absence of a vaccine

If a vaccine is expected to arrive with sufficiently high probability, we show that both private agents and the social planner will want to shield more. When an effective vaccine arrives both future infections and herd immunity become irrelevant, so the immunity externality loses its power and the rat race to shield vanishes. In this case the contagion externality dominates, because it is still present in the interval between the present time and the actual arrival of the vaccine. With the expectation of a vaccine our simulations show that the social planner will want to impose more social

distancing than private agents choose to do.[4]

There has been a very large output of working papers on COVID-19, and more generally on epidemics, by economists since the outbreak of the pandemic. We cannot provide a survey here but we put our paper into context and mention some papers that address epidemics within models that have common features with ours. We believe that our main results about externalities have not appeared in any other papers but there are papers that independently model the trade-offs that we do, pursuing related objectives. In general, the externality that is frequently encountered in the literature is akin to our static externality that arises through random contacts.

First, we should note that although most economists became interested in epidemiological models because of Covid-19, a small earlier literature dealt with behavioural responses to epidemics in the framework of Kermack and McKendrick (1927). Most papers in this tradition focus on a negative contact externality similar to our static externality: too little social distancing by self-interested agents that do not internalize the costs of transmission to others.[5]

Post-Covid, several papers study the trade-off between the policies needed to contain the disease and the loss of economic activity that they necessitate. Alvarez, Argente, and Lippi (2020) solve an optimal control problem to find the efficient containment policy in the presence of Covid-19, but they do not model the decentralized solution. Eichenbaum, Rebelo, and Trabandt (2020a) solve computationally for the optimal containment policy in a model in which individuals face a consumption externality. Farboodi, Jarosch, and Shimer (2020) use a quadratic contact function and discuss the impact of the disease on social distancing and the resulting contact externality. They obtain the longer time needed to reach herd immunity when agents shield. Rachel (2020) solves for the optimal lockdown and highlights the role of fiscal policy and infection externalities. The latter has both a static and a dynamic dimension and has ambiguous overall effects on the planner's choice. His model, however, is not as general as ours, in the sense that he derives his results from a two-state social action, high and low. Nævdal (2020) shows that there are increasing social returns to scale to social distancing and other efforts to control the epidemic. Aspri, Beretta, Gandolfi and Wasmer (2020) study the trade off between mortality and the fall of GDP in a SEIARD model, and simulate various containment policies. Kapicka and Rupert (2020) model

[4]We should note that since private agents choose a path that yields a cumulative infection rate close to the minimum consistent with herd immunity, the deviations between the private and social optimum paths relates mainly to the adjustment to herd immunity and not to the final infectious outcome.

[5]See for example, Chen (2012), who like us invokes increasing returns to scale in contacts, and Rowthorn and Toxvaerd (2019), who use a contact function with fixed proportions and a linear cost of prevention borne by the individual. Quercioli and Smith (2006) go further and touch on one aspect of the dynamic externalities. In the pre-Covid research, economists have been working on integrating behavioural choice into epidemiological models of sexually transmitted diseases, and HIV in particular. Kremer (1996) and Greenwood, Kircher, Santos and Terlit (2019) are two such examples, but are not dealing within a SIR environment.

the labour market dimensions of the pandemic, and solve for the centralized program that takes into account an infection externality, as agents do not take into account that, once infected, the probability that others will get infected increases. The labour market response to the Covid-19 pandemic within a search environment is also studied by Gregory, Menzio and Wiczer (2020), but without reference to the planner's solution. The potential uncertainty about one's infection status, which plays an important role in our paper, and the role of testing are studied by Berger, Herkenhoff, and Mongey (2020) and von Thadden (2020).

Another branch of the recent literature considers heterogeneous agents. Heterogeneity is a feature of Covid-19 that we ignore in this paper. Models that deal with it include Acemoglu, Chernozhukov, Werning, and Whinston (2020), Kaplan, Moll, and Violante (2020), Alfaro, Faia, Lamersdorf, and Saidi (2020), Brotherhood, Kircher, Santos, Tertilt (2020) and Favero, Ichino, and Rustichini (2020)

Section 2 describes the model in more detail and derives the individual maximizing choices in a world were no vaccine is expected to arrive, so that the epidemic will not end before herd immunity is obtained. Section 3 simulates the epidemic equilibrium in our model against the standard SIR. Section 4 derives the welfare maximizing chocies of a central planner. Section 5 analyzes the model when a vaccine is expected to arrive with some probability. Section 6 simulates the model highlighting the key role of the static and dynamic externalities in driving a wedge between the decentralized and social optima. The last section concludes.

2 Decentralized Equilibrium in a Model of Epidemics

In this section we develop a model of transitions with forward-looking individual decision making. We work in discrete time and define the period to be the length of time that an infected individual has no symptoms. In terms of COVID-19, the length of the period is therefore about two weeks, although it could be longer. During this period infected individuals are unaware that they are contagious because of their infection. At the end of the period of infection symptoms arrive and the individual receives treatment overnight, which is costly in terms of lifetime utility, waking up in the next period fully recovered and immune.[6] We assume that no vaccine is expected to arrive.

The sequence of events that we study is as follows. In some initial state the entire population except for a very small number ε are susceptible, with mass S_0.[7] The ε

[6] As we mentioned before, full immunity is not a feature of all infections. In the case of COVID-19, there is some support that infection could leave undetermined long-term effects which can be quite costly. See for example, https://www.bbc.com/future/article/20200622-the-long-term-effects-of-covid-19-infection and https://news.umiamihealth.org/en/what-are-the-long-term-effects-of-covid-19/

[7] We use S, and later I and R, to denote both the state and the mass of individuals that belong to

individuals are infected with no symptoms, so they belong to state I. Transitions of some susceptible individuals from state S to state I the following period depend on contacts, which arise in a variety of situations, such as work, shopping and leisure activities. All individuals in state I in period t make a transition to recovery a period later and join state R, after undergoing treatment between periods t and $t+1$. The transition out of I and the treatment cost between periods depend only on medical conditions related to the disease that the individual cannot influence. In the absence of deaths, we make use of the convenient assumption that population is constant and all individuals in each period t belong to one of three states, S, I and R. We simplify further by assuming

$$S_t + I_t + R_t = 1 \qquad \forall t. \tag{1}$$

Agents have a utility function defined over two activities, one of which is done at home and one outside. Home activities include paid work at home, home production, online shopping and home leisure activities, such as watching TV. Social activities, include market work and consumption and leisure activities outside the home. Social contact results only from the second set of activities. We denote the first set of activities by x_h and the second by x_s and write the per-period utility function as,

$$u_t = u(x_{ht}, x_{st}). \tag{2}$$

This function is assumed to satisfy the Inada conditions.[8] The choice of x_{ht} and x_{st} is constrained by a cost function which we assume for simplicity that it is a convex utility cost $c(x_{ht}, x_{st})$. We define net utility from all activities by,

$$\hat{\phi}_t = \hat{\phi}(x_{ht}, x_{st}) = u(x_{ht}, x_{st}) - c(x_{ht}, x_{st}), \tag{3}$$

assumed to be single peaked.

To simplify further, we note that since home activities do not influence contacts, and are chosen optimally by each individual, there is an implicit policy function for x_{ht}, derived from,

$$\frac{\partial \hat{\phi}(x_{ht}, x_{st})}{\partial x_{ht}} = 0 \qquad \forall t. \tag{4}$$

We substitute back x_{ht} from this function into equation (3), to obtain an indirect net-of-cost utility function (suppressing the subscript s),

$$\phi_t = \phi(x_t). \tag{5}$$

it. In general, when reference is to the state there is no subscript but when reference is to the mass it is dated with a time subscript.

[8]Our utility function implies that going out yields utility that is independent of the actions of others. McAdams (2020) explores a situation in which the utility from social action depends on others taking social action too. This gives rise to multiple equilibria in social distancing.

This function has an inverted U-shape and we normalize x_t such that $\phi\prime(x_t) = 0$ at $x_t = 1$.

Agents in our framework perceive to be in one of two states: not recovered or recovered and immune. Not recovered individuals are either healthy but susceptible to the disease or infected without symptoms. An important assumption of our model is that there is no information available that can help agents distinguish between states S and I; in order to highlight this important assumption we introduce a new term for non-recovered individuals, *vulnerable*. All vulnerable individuals choose the same x_t that maximizes lifetime returns subject to the probabilities of belonging to state S or I. Testing the population not in the recovery state would reveal the information needed to distinguish between the susceptible and the infected agents, but it is assumed absent.

Vulnerable agents enjoy per-period net utility given by (5) and they perceive a risk of infection from contacts with infected individuals. In general, an agent who is in state S in period t enters state I in period $t + 1$ with probability,

$$p_{t+1} = p(x_t, \bar{x}_t, x_{r,} S_t, I_t, R_t), \tag{6}$$

where x_t is the choice of social activity made by the agent, $\bar{x}_t$ are the choices of other vulnerable agents, 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). We assume,

$$\begin{aligned} \frac{\partial p(x_t, .)}{\partial x_t} &\geq 0, \\ p(0, .) &= 0, \end{aligned} \tag{7}$$

where $p(0, .)$ is the transition to infection in the state of complete social distancing. Note that the assumptions made about timing in the transition in equation (6) imply that p_{t+1} is the probability that the susceptible agent becomes infected between periods t and $t + 1$ and so it is predetermined at the start of period $t + 1$.

We now introduce present discounted values associated with the three states. The NPV of utility of a recovered individual is denoted by W_t^R. Assuming infinite lives and immunity, the recovered individual chooses x_t to maximize the PDV of $\phi(x_t)$ without the risk of infection, so the solution is trivially the constant $x_r = 1$ that satisfies $\phi'(x_r) = 0$. With a discount factor β we write

$$W^R = \frac{\phi(x_r)}{1 - \beta} \qquad \phi'(x_r) = 0 \tag{8}$$

The NPV of utility of an infected individual is denoted by W_t^I, while the NPV of utility of a susceptible individual is denoted by W_t^S. Consider now a vulnerable individual in period t. She chooses x_t without knowing whether she is in the infected

state (with utility W_t^I) or in the susceptible state (with utility W_t^S). The expected NPV of utility of vulnerable individuals is denoted by W_t^V. The vulnerable person knows that she was susceptible in the previous period $t-1$, otherwise she would have been sick and recovered between $t-1$ and t, and she knows the probability p_t that she contracted the virus in that period, which is now predetermined. She chooses the present-period x_t by maximizing the expected value of the PDVs of being in the two vulnerable states, given her p_t :

$$W_t^V = \max_{x_t} \left\{ p_t W_t^I + (1-p_t) W_t^S \right\}. \tag{9}$$

In order to obtain the solution to this problem we need to specify the lifetime values W_t^I and W_t^S. Both infected and susceptible individuals obtain utility from activity $\phi(x_t)$. An individual who is infected in t will suffer overnight a utility loss from treatment but become recovered in $t+1$. We assume that the utility loss from the disease is proportional to the lifetime utility of a recovered person and write it as $\delta(I_t)W^R$. The cost depends on the number of infected people who are receiving treatment at the same time as the agent, which we call *medical congestion*. A low I_t implies that there is better health care. With fixed hospital space medical care workers (both of which are outside our model) give more attention to fewer patients; $\delta'(I_t) \geq 0$. It therefore follows that the NPV of utility of an infected individual, W_t^I, is,

$$W_t^I = \phi(x_t) + \beta(1-\delta(I_t))W^R. \tag{10}$$

Clearly, if the person knew that she was in this state she would have (selfishly) chosen the same social action x_t as a recovered person, shown in (8). But because of the absence of this information, the x_t that enters (10) is the one that is obtained from (9).

An individual who is susceptible may be in contact with an infected person during the period, contract the virus, and end up as infected next period. This happens with probability $p_{t+1} = p(x_t)$. The person continues to be susceptible with the complementary probability $1-p_{t+1}$. It follows that we can write

$$\begin{aligned} W_t^S &= \phi(x_t) + \beta \left[p_{t+1} W_{t+1}^I + (1-p_{t+1}) W_{t+1}^S \right] \\ &= \phi(x_t) + \beta W_{t+1}^V. \end{aligned} \tag{11}$$

Substituting (11) and (10) into the maximization problem of a vulnerable individual, (9), this problem becomes,

$$W_t^V = \max_{x_t} \left\{ \phi(x_t) + \beta p_t (1-\delta(I_t))W^R + \beta(1-p_t)W_{t+1}^V \right\}. \tag{12}$$

Differentiation of (12) for period $t+1$ yields,

$$\frac{\partial W_{t+1}^V}{\partial x_t} = -\beta p'(x_t)\left[W_{t+2}^V - (1-\delta(I_{t+1}))W^R\right]. \tag{13}$$

The first order condition for x_t that is obtained from problem (12) is therefore given by,

$$\phi'(x_t) = \beta^2(1-p_t)p'(x_t)\left[W_{t+2}^V - (1-\delta(I_{t+1}))W^R\right]. \tag{14}$$

It trivially follows from our assumptions that $\left[W_{t+2}^V - (1-\delta(I_{t+1}))W^R\right] > 0$, which amounts to saying that healthy individuals are better off than infected ones. It is clear from the first order conditions that in the case of an infectious disease healthy agents restrict their activities outside the home to avoid infection. Without an infectious disease the first order condition for activities outside the home would be $\phi'(x_t) = 0$, as chosen by recovered individuals, yielding a higher x_t than the solution in (14). We refer to this property as social distancing or shielding.

We now specify the contact technology that yields the infection probability $p(x_t, .)$. This parallels the matching function of labour economics (Petrongolo and Pissarides, 2001) but with some important differences. In the matching function of the labour literature, more workers looking for jobs reduces the success probability of a single worker because of congestion externalities in the application process. Here more individuals coming out in the marketplace increases 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.

To provide an intuitive derivation of our contact function suppose $\bar{x}_t$ stands for the number of trips outside the house that each person does.[9] Assume that with $\bar{x}_t$ trips, each person experiences on average m_t contacts, defined by a well-behaved function $m_t = m(\bar{x}_t)$, with $m'(\bar{x}_t) \geq 0$. The function $m(.)$ is similar to the matching function of labour economics in the sense that it depends on the structure of the marketplace, including density of population, transportation facilities, types of establishments etc.[10] 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.

Consider now the choices made by a single individual, who is able to choose the number of trips outside the home, denoted by x_t. Here we follow the method used in search theory to choose the optimal search intensity (Pissarides, 2000, chapter 5). With $m(\bar{x}_t)$ representing the total number of contacts for $\bar{x}_t$ outings, each outing on average

[9] Appendix 1 derives a special example of a contact function from the urn-ball game, which satifies the main properties of the general form discussed here.

[10] The dependence of $m(.)$ on a single variable parallels the contact function used by Peter Diamond (1992) in his famous "coconut" paper. He assumes that there are b agents with a coconut each coming into contact pairwise, for a contact technology $m(b)$ with $m'(b) > 0$. See section 4.2 for more discussion of the contact function.

generates $m(\bar{x}_t)/\bar{x}_t$ contacts. So if the individual chooses to go out of the home x_t times, her contacts are on average $x_t m(\bar{x}_t)/\bar{x}_t$. These are total contacts. We are interested in the contacts that can potentially lead to an infection, and these are contacts that involve a person from set I. Since the susceptible person cannot distinguish a priori who is in which state, on average the fraction of contacts that are infected is equal to the fraction of persons in set I in the population. With the normalization of the population size to unity, we obtain that the probability that a contact is with an infected person is simply I_t. Finally, suppose that the probability that a contact between a susceptible and an infected person leads to the infection of the susceptible person is a medical constant k. This medical constant is a measure of the infectiousness of the disease, with $k = 0$ indicating a non-contagious disease and $k = 1$ indicating a very contagious one.

The transition from the susceptible to the infected state for the person who chooses x_t becomes,[11]

$$p_{t+1} = k\frac{x_t m(\bar{x}_t)}{\bar{x}_t} I_t. \tag{15}$$

This expression satisfies the extreme properties that for a non-infectious disease ($k = 0$) or complete social isolation ($x_t = 0$), $p_{t+1} = 0$. It follows from this expression that p_t now depends on a smaller set of variables than in the general expression (6). Its partial derivative satisfies,

$$\frac{\partial p_{t+1}}{\partial x_t} = k\frac{m(\bar{x}_t)}{\bar{x}_t} I_t = \frac{p_{t+1}}{x_t}. \tag{16}$$

In moving from individual transitions to the average for a market where all agents optimize we assume a symmetric Nash equilibrium in which all agents choose the same policy, so $x_t = \bar{x}_t$. For notational simplicity we drop the bar from $\bar{x}_t$ and write the equilibrium p_{t+1} as,

$$p_{t+1} = km(x_t)I_t, \tag{17}$$

with x_t obtained as the solution to (14), under the restriction $\bar{x}_t = x_t$ and given all the value equations previously derived and the matching restrictions in (16).

This completes our specification and derivation of the solution equations for the agents in the model. It is noteworthy that when comparing with the epidemiological SIR model, our innovation is the insertion of x_t in the transition probability p_{t+1}, which picks up the disincentives that the susceptible individuals have when they go out of their homes. Some obvious properties of this choice, given our strong functional assumptions, can easily be derived. There is social distancing (lower x_t), for higher

[11] Another derivation of the probability of meeting at least one infected individual is to reason as follows. Since for each contact there is a probability $(1 - I_t)$ that the person does not meet an infected person, there is a probability $(1 - I_t)^{xm(\bar{x})/\bar{x}}$ that the person does not meet any infected persons in her x outings. If I is a small fraction of the population, this is approximately equal to $\exp\{Ixm(\bar{x})/\bar{x}\}$, so the probability of meeting an infected person is $1 - \exp\{.\}$ and for small transition probability this is approximately equal to the expression in the text.

k and higher I_t (more infectiousness of the disease or more infected people) and for higher unpleasantness from treatment (higher difference between the value of avoiding infection W_t^S and getting infected, W_t^I).

We now complete the description of the decentralized equilibrium by deriving the transitions implied by our individual models. With transition probability from state S to state I given by (17), the number of people in state S falls each period by the fraction in (17). This is also the number of people who join the I state, whereas a period later every infected individual joins the recovery state R. In discrete time, the dynamics of the system can be written as,

$$S_{t+1} = S_t - km(x_t)I_tS_t \tag{18}$$
$$I_{t+1} = km(x_t)I_tS_t \tag{19}$$
$$R_{t+1} = R_t + I_t \tag{20}$$

Because of our assumption that infected people recover in one period, our model implies that the "basic reproductive number" R_0 of the disease is simply $R_{0t} = km(x_t)$. From this we derive the effective reproductive number, or R-number, which is given by S_tR_{0t}. This number plays a key role in the dynamics of the disease. From (19), infections begin to fall when the R-number drops below 1.

We are now in a position to define our decentralized equilibrium.

Definition 1 *A decentralized epidemic equilibrium is a set of sequences of state variables* $\{S_t, I_t, R_t\}_{t=0}^{\infty}$*, a set of value functions* $\{W_t^V, W_t^S, W_t^I, W^R\}_{t=0}^{\infty}$*, and a set of sequence of probabilities and social contacts* $\{p_t, x_t\}_{t=0}^{\infty}$ *such that, for given initial conditions* $S_0 = 1 - \epsilon$*,* $I_0 = \epsilon$*,* $R_0 = 0$

1. S_t, I_t, R_t *solve equations (18)-(20)*
2. *The value functions* W^R, W_t^V, W_t^S, W_t^I*, solve equations (8), (12), (11), and (10)*
3. x_t *solves the first order condition (14)*
4. p_t *solves equation (17)*

3 Matching-SIR vs Original SIR: Simulations

3.1 Parameterization and the basic reproductive number

As we pointed out in the preceding section, our main innovation when compared with the standard epidemiological SIR model is the optimal choice of x_t, which reflects the disincentives that agents have to mix socially when there is a risk of infection. Before we examine in detail the properties of our decentralized equilibrium, we compare the dynamic behaviour of a parameterized version of our model with the standard

SIR model. We define the standard SIR model in the same way as our model, with the exception that all agents choose the first-best social activity level that satisfies $\phi'(x_t) = 0$. (In our model, this is chosen only by recovered and immune individuals.) This conforms with standard practice (see for example, Weiss, 2013) and it highlights the new elements of our model.

We work with a model that restricts our general framework of section 2 in simplifying ways that enable us to focus on the main differences of the dynamics between the two models. We regard the main differences as two: the length of time that it takes to reach herd immunity and the cumulative number of infections. We show that the optimal decentralized behaviour of x_t lengthens substantially the time it takes to reach herd immunity but it also reduces substantially the cumulative number of infections.

We use the following parameterization of the net semi-indirect payoff function, $\phi(x_t) = A + lnx_t - x_t$. Therefore the optimal x_t for all agents in the standard SIR model and for recovered agents only in our model satisfies the normalization $x_r = 1$. The PDV after recovery is $W^R = (A-1)/(1-\beta)$. We set $A = 1.5$ and $\beta = 0.998$. If a period is two weeks, this corresponds to an annual discount rate of 0.05. The other equations of the model are (12), (14), (18) and (19), with the cost of treatment assumed to be constant at $\delta_0 = 0.1$.

R_{0t} is the basic reproductive number in period t, defined as $km(x_t)$, where x_t is the social activity chosen by the vulnerable individuals who have no immunity. The meeting technology is assumed to have constant elasticity $m(x_t) = x_t^\alpha$, so $R_{0t} = kx_t^\alpha$, with $\alpha \geq 1$.[12] In the illustrative simulation below we select $\alpha = 1$ and $k = 2.4$, so for the standard SIR, $R_{0t} = 2.4\ \forall t$. For our model, $R_{0t} = 2.4$ when vulnerable agents select $x_t = 1$, which takes place before the epidemic starts. All parameters are listed in Table 1.

We note that in the standard model the dynamics of the effective reproductive number R_t are governed by the dynamics of the stock S_t, the number of susceptible people. With our parameterization, the effective R_t of the standard model is $2.4S_t$. But in our model R_t is $2.4S_tx_t$, assigning as big a role to the optimal choice of x_t as to S_t.

3.2 Trajectories

The model's solution, once it is shocked in the beginning by a very small number of infections $\varepsilon = 0.001$, is obtained with a shooting algorithm - a standard solution algorithm for systems of difference equations that are highly non-linear and feature both initial and terminal conditions (Sargent and Stachurski, 2020).

In a first comparison we plot the dynamic performance of our model and the standard SIR model in Figure 1. In each chart the red/broken line refers to the decentralized

[12]Note that we are reinterpreting k in this simulation. It is the product of the infection probability conditional on a contact and any constants that might belong to $m(x_t)$. We economize on notation by including all constants in k.

outcome while the blue/continuous line to the standard SIR. As in most figures that follow, the top chart shows the activity level $\{x_t\}$ in the decentralized equilibrium, the middle chart the mass of susceptible people, and the bottom chart the mass of infected people. Figure 1 shows the first 100 periods of the simulation (about four years).

Before the disease the two models have the same activity level, normalized to unity. But whereas in the standard SIR the activity level remains at that level throughout the epidemic, in the decentralized simulation it drops a little when the ε infections are first introduced and then drops dramatically and very quickly to a number below 0.5. This drop cuts the effective reproductive number in the decentralized equilibrium to $1.2S_t$. The middle panel plots the stock of susceptible people, initially normalised to one for both models. In the standard SIR model, with the high activity level continuing, the stock of susceptible people drops very quickly to the herd immunity level. After about ten periods (20 weeks) only 3.7% of the susceptible people avoid the infection, which is virtually a state of herd immunity, reached when 3.6% avoid it. Adjustments are much slower in the case of the optimizing solution. As people drop the level of activity, the stock of susceptible people falls smoothly and gradually throughout the epidemic. After 10 periods, 72.5% of the population are still disease-free. In the bottom panel of Figure 1 we plot the stock of infected people. While in the case of the standard SIR the dynamics follow the traditional hump shape, the growth of infected people in the case of the optimizing SIR is so mild that it is barely discernible, even after four years.

3.3 The state of herd immunity

To illustrate further the properties of the state of herd immunity, let S^∞ be the number of susceptible individuals in the new steady state equilibrium after herd immunity is obtained. Since in steady state $I_t = 0$, we have that $R^\infty = 1 - S^\infty$, where R^∞ is the number of recovered people (which is equal to the number who historically got the disease).

In steady state, the effective reproductive number $S^\infty R_0$ has to be less than or equal to 1. Hence an upper bound for S^∞, $S^{\max}$, is given by

$$S^{\max} = R_0^{-1}. \tag{21}$$

In the standard SIR model, which is similar to our model with constant x, the maximum number of infected individuals is obtained when $I_t \approx I_{t+1}$ (with equality in continuous time). Plugging $I_t = I_{t+1}$ into (19), gives that $S = R_0^{-1}$ $(= S^{\max})$. From this point on, the disease is on retreat, as the effective reproductive number falls below 1. However, it takes time before the disease burns out, and along the path many more people are infected. It can be shown (Weiss 2013) that the steady state value of S_t in the continuous time SIR model, denoted S^{SIR}, is given by the solution to the equation $\ln S^{SIR} = R_0(S^{SIR} - 1)$. This equation can be solved numerically, and for $R_0 > 1$ it gives that S^{SIR} is substantially lower than $S^{\max}$.

When the activity level is set by forward-looking individuals, this is no longer the case. At the point at which I_t reaches its maximum level, the individual probability of attracting the virus for a given activity level is the highest, and the agents respond by cutting back on social activity. The simulations show that the activity level, and hence the basic reproductive number, are at their lowest around the period in which I_t is at its highest. If we denote by x^I the decentralized equilibrium value of x at the point at which I reaches its maximum level, it follows that the stock of infected people at this point is approximately equal to $S^{\max}/x^I > S^{\max}$ (exactly equal in continuous time). Hence, at the peak, $S_t > S^{\max}$, and then it gradually falls and dips below $S^{\max}$ as society converges towards a new steady state with herd immunity.

In our simulations of the decentralized equilibrium, S^{∞} is almost equal to $S^{\max}$ ($S^{\max} = .417$ while S^{∞} is slightly below .4) , meaning that the stock of susceptible people converges to a value almost equal to the highest level consistent with herd immunity. The key result is that forward-looking agents will restrict activity so that herd immunity is reached at a point close to the lowest possible number of total infections, even though this action delays the arrival of herd immunity substantially.

The behaviour of herd immunity in the two models is shown in the phase diagram of Figure 2. The stock of susceptible people is on the horizontal axis and the stock of infected people is on the vertical axis. The starting point of the epidemic is 1 on the horizontal axis, when the entire population is susceptible. The arrows plot the joint dynamic path of $\{S_t(x_t)\}$ and $\{I_t(x_t)\}$, where the optimal x_t is obtained from the. simulations in Figure 1. The black arrows refer to the traditional SIR with $x_t = 1$, while the red/broken arrows refer to the optimizing SIR. The value of $S^{\max}$ is also clearly indicated by the vertical line in Figure 2. The black curve reaches the maximum of infected people at $S^{\max}$, and thereafter converges to a new steady state, but in that steady state a large number of individuals have attracted the virus relative to the number needed for herd immunity.[13] The phase diagram of the optimizing SIR is given by the red/broken curve and converges to a point close to $S^{\max}$. It is clear from the phase diagram that the dynamics of the optimizing SIR towards $S^{\max}$ is also hump-shaped.

[13]In Figure 2 the black curve increases beyond $S^{\max}$ because time is discrete.

4 Externalities and Deviations from Social Efficiency

4.1 Formal statement

As in other models of pairwise interaction, we would expect the decision strategies derived in the preceding section to be subject to externalities and inefficient outcomes. We derive the socially optimal strategies by assuming the existence of a social planner who chooses social activities for all agents. The information that the social planner possesses about agent identities and the future path of the economy coincides with that of private agents. As with private agents, the planner chooses the same x_t for all vulnerable agents and a separate x_{rt} for the recovered. However, since the recovered are immune, and they do not influence the transition rates of the susceptible or infected individuals, the social planner will choose $\phi'(x_{rt}) = 0 \ \forall t$. This matches private choices so it can be ignored and we can focus our analysis on the vulnerable agents only.

There are two channels that link social activity to market outcomes. First, the social planner takes into account the fact that the equilibrium is a symmetric Nash equilibrium. All vulnerable agents end up choosing the same action and when one person meets another the other person is also involved in a meeting. These facts are ignored by private agents in the decentralized equilibrium. We refer to any externalities that arise from this channel as *static externalities.*

Second, the planner is also aware that unlike the perceptions of private agents, with her actions today she influences the future dynamics of S_t and I_t, through equations (18)-(20). Future value functions are influenced by the measures S_t and I_t, partly by influencing infection probabilities and partly through the medical congestion externality. We refer to the externalities due to this channel as *dynamic.* Further below we argue that there are three types of dynamic externalities, which we label *medical congestion, contagion* and *immunity.* We assume that medical expenses are fully covered by the individual.

Because the PDV of expected returns for all vulnerable individuals is the same value function $W_t^V = W^V(x_t, S_t, I_t)$, we can derive the social optimum choice of x_t by having the social planner step into the shoes of the vulnerable agent and select x_t to maximize the social W_t^V. The relevant transition probability for the social planner is (17), in which the restriction that all agents choose the same x_t is imposed. The planner's controls are the activity levels $\{x_t\}_{t=0}^{\infty}$ and the objective the constrained maximization of the value function

$$\begin{aligned} W_t^V(S_t, I_t, p_t) &= \phi(x_t) + \beta p(x_{t-1}, I_{t-1}) W^R(1 - \delta(I_t)) \\ &\quad + \beta(1 - p(x_{t-1}, I_{t-1})) W_{t+1}^V(S_{t+1}, I_{t+1}, p_{t+1}), \end{aligned} \tag{22}$$

subject to the laws of motion (18)-(19).

Consider first the role of x_t in the contemporaneous value function W_t^V. It influences the utility $\phi(x_t)$ directly, but it does not influence the transition probability $p_t = p(x_{t-1}, I_{t-1})$. So the social planner cannot do anything to influence the second term on

the right-hand side of (22), which is the expected lifetime return of infected agents. The third term is the expected lifetime return of the vulnerable and the value of x_t that is chosen in t will influence the transition probabilities in period $t+1$ and from there and through the constraints it will influence the measures $S_{t'}$ and $I_{t'}$ in all future periods $t' > t$. Taking those into account we write the first-order condition for x_t,

$$-\phi'(x_t) = \beta\,(1-p_t)\left\{\frac{\partial W^V_{t+1}}{\partial p_{t+1}}\frac{\partial p_{t+1}}{\partial x_t} + \frac{\partial W^V_{t+1}}{\partial I_{t+1}}\frac{\partial I_{t+1}}{\partial x_t} + \frac{\partial W^V_{t+1}}{\partial S_{t+1}}\frac{\partial S_{t+1}}{\partial x_t}\right\}. \tag{23}$$

The left-hand side of (23) measures the utility cost of the social distancing in the current period, the deviation of $\phi'(x_t)$ from the unconditional optimum 0. The right-hand side gives the gains in expected lifetime returns from the social distancing, which accrue to the susceptible individuals (the fraction $1-p_t$ of all vulnerable), from the next period onwards, discounted to the present at β.

We show:

Proposition 2 *The social optimum level of social activity is the solution to the following equations:*

$$\begin{aligned} -\frac{\phi'(x_t)}{1-p_t} &= -\beta^2 k m'(x_t) I_t \left[W^V_{t+2} - (1-\delta(I_{t+1}))W^R\right] \\ &+ \beta\left(\frac{\partial W^V_{t+1}}{\partial I_{t+1}} - \frac{\partial W^V_{t+1}}{\partial S_{t+1}}\right) k m'(x_t) S_t I_t. \end{aligned} \tag{24}$$

$$\frac{\partial W^V(S_t, I_t, p_t)}{\partial I_t} = -\phi'(x_t)\frac{m(x_t)}{m'(x_t)I_t} - \beta p_t \delta'(I_t) W^R \tag{25}$$

$$\begin{aligned} \frac{\partial W^V(S_t, I_t, p_t)}{\partial S_t} &= \frac{1-p_{t+1}}{p'_{t+1}S_t}\left[\phi'(x_t) + \beta(1-p_t)p'_{t+1}\frac{\partial W^V_{t+1}}{\partial p_{t+1}}\right] \\ &+ \beta(1-p_t)\frac{\partial W^V_{t+1}}{\partial I_{t+1}} \end{aligned} \tag{26}$$

The proof of the Proposition uses the Benveniste-Scheinkman theorem and is given in Appendix 2. Here we will try to give some intuition for the proof and the expressions in the Proposition.

In (24), the first term on the right is the expected gain from social distancing. This term picks up the *static efficiency condition* and follows immediately from differentiation of (22). In (25) the last term on the right-hand side is the *medical congestion externality* and also follows from differentiation of (22). In the first term, one additional unit of I_t increases the number of infected people in the next period with the

same number as $\frac{m(x_t)}{m'(x_t)I_t}$ units of x_t.[14] The unit cost of reducing x_t is $m'(x_t)$. The marginal cost of increasing I_t due to the potential for more infections in period $t+1$ is the product of the two, which is the first term on the right-hand side of (25). In the proof in the Appendix this argument, which is the source of the *contagion externality,* is made rigorous.

Consider finally (26), which gives rise to the *immunity externality.* The immunity externality is associated with a decrease in S_t. For a given (optimal) sequence of activity levels $x_t, x_{t+1}, ...$, a unit decrease in S_t does not influence the pay-offs in period t, nor the probability p_{t+1} that a given person is infected in the next period, since $p_{t+1} = x_t k m(x_t) I_t$. However, S_t does influence the number of infected people in period $t+1$, $I_{t+1} = km(x_t)S_t I_t$, and hence also the probability of contracting the virus in that period, $p_{t+2} = I_{t+1} k m(x_{t+1})$, and in later periods as well. Hence we would expect the expression $\frac{\partial W^V(S_t, I_t, p_t)}{\partial S_t}$ to be negative. From (24), when I_{t+1} and S_{t+1} move in opposite directions, as they do when x_t changes, the contagion and immunity externalities have counteracting effects on the planner's social choice.

It is not so easy to give a precise intuition as to why when considered together with the other externalities, the immunity sexternality has exactly the form given in (26). Note however, that the planner can costlessly react to an increase dS_t by increasing x_t by $\frac{1-p(x_{t+1})}{p'(x_{t+1})S_t} dS_t$ units, thereby keeping S_{t+1} unchanged. But this increases I_{t+1} by one unit and p_{t+1} by $p'(x_t)\frac{1-p_{t+1}}{p'_{t+1}S_t} dS_t$ units. The per-unit gain to the planner of increasing x_t is $\phi'(x_t)$ units. Together this gives (26). This argument is also made rigorous in the Appendix. Note that $\frac{\partial W^V_{t+1}}{\partial I_{t+1}}$ is given by (25), and $\frac{\partial W^V_{t+1}}{\partial p_{t+1}}$ follows readily by taking the derivative of (22).

We can write the dynamic externalities as

$$\begin{aligned}\left(\frac{\partial W^V_{t+1}}{\partial I_{t+1}} - \frac{\partial W^V_{t+1}}{\partial S_{t+1}}\right) &= \frac{\partial W^V_{t+1}}{\partial I_t} - \beta(1-p_t)\frac{\partial W^V_{t+1}}{\partial I_{t+1}} \\ &\quad - \frac{1-p_{t+1}}{p'_{t+1}S_t}\left[\phi'(x_t) + \beta(1-p_t)p'_{t+1}\frac{\partial W^V_{t+1}}{\partial p_{t+1}}\right]\end{aligned} \tag{27}$$

Equations (24) and (27) fully define the first order condition for the optimal x_t.

4.2 Discussion of externalities

From an *a priori* perspective, it is not clear if the planner will want to implement a higher or a lower activity level than the equilibrium level in the decentralized problem. The static externalities, as we argue below, are likely to lead the social planner to reduce social activity but the dynamic ones might give a different outcome. Although the internalization of the contagion externality leads the planner to reduce social activity

[14] This follows from the fact that $dI_{t+1} = km(x_t)dI_t + km'(x_t)I_t dx_t$.

further, the immunity externality gives a push-back which might offset it. Private agents know that there is a herd immunity state in which some susceptible people do not catch the disease so if the cost of catching it is high, each individual will have an incentive to reduce social activity further, in order to avoid being among those who get ill before herd immunity is reached. This has the elements of a rat race, although instead of a race to be first it is a race to shield: stay home in the hope that you will be among the lucky ones that reach the herd immunity state without an infection. This is likely to become stronger the higher the medical costs to the individual and introduces a reason that the planner might want to increase social activity above the decentralized equilibrium.

Similar considerations arise in the dynamic setting for the medical congestion externality. In a static perspective, the medical externality leads to a negative externality that the planner internalizes by reducing social activity. However, the immunity effect includes *future* medical externalities, which may be bigger or smaller depending on whether or not we are on the increasing or decreasing part of I_t's trajectory. The medical congestion externality implies that the planner will want to avoid peaks in infections, and this will typically imply more social distancing early on when infections are rising fast.

We explore further these issues in two ways. In this section we give more intuition to the results and point to likely directions of the externalities by working with an approximation that gives clear results. In section 6 we simulate the full solution with reasonable parameter values.

In equation (24), we hold for the moment I_{t+1} and S_{t+1} constant, then the equation becomes,

$$\phi'(x_t^c) = \beta^2 (1 - p_t) p'(x_t^c) \left[W_{t+2}^V - (1 - \delta(I_{t+1})) W^R \right], \tag{28}$$

where now superscript c denotes the value of x_t that the social planner would choose if only the static externalities were present (since we zeroed the dynamic externalities by holding I_{t+1} and S_{t+1} constant). Superscript c stands for contacts, the only source of externalities in this example.

The planner's static efficiency condition has similar terms to the ones in the decentralized maximization, (14). It is immediate from the comparison of the two conditions that the decentralized and social solutions coincide if the partial derivative $p'(x_t^c)$ of the social problem coincides with the partial derivative of the decentralized problem. From (16) and (17) the condition required is a log-linear meeting technology:

$$\frac{x_t m'(x_t)}{m(x_t)} = 1. \tag{29}$$

This requirement parallels the familiar elasticity condition from matching theory, often referred to as the Hosios (1990) condition, which applies to situations of pairwise matching (see Pissarides, 2000, chapter 8). What does it mean in our context?

To show the significance of increasing returns in contacts, we differentiate the infection probability of a single agent, (15), with respect to other agents' actions, $\bar{x}_t$. With

constant returns in $m(\bar{x}_t)$, the partial is zero, so less social distancing by others in the market does not influence this person's infection probability. But with increasing returns, $\partial p_{t+1}/\partial \bar{x}_t > 0$; a person is more likely to be infected when other people in her community reduce their social distancing. This is the essence of the externality: it arises in situations in which a change in a typical individual's social distancing has an impact on other people's infections probability.[15]

This argument works conversely as well. When there are unintended contacts in fixed social space, increasing returns arise as a natural consequence. The justification is similar to the one used by Peter Diamond (1982). In that paper islanders posses 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 because as both buyers and sellers double in number they create congestion for each other and so many swaps are crowded out (Petrongolo and Pissarides, 2001). 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 she can only give her coconut to one person. Diamond's intuition for increasing returns applies to this model much more than in a model of exchange.[16]

To give intuition to the dynamic inefficiencies we follow a similar partial reasoning but now remove the static externalities. This can be done by assuming that the meeting technology is log-linear. The dynamic externalties D_{xt} are given by the last term in (24) multiplied with $1 - p_t$, hence

$$D_{xt} \equiv -\beta(1-p_t)km'(x_t)S_tI_t\left(\frac{\partial W^V_{t+1}}{\partial I_{t+1}} - \frac{\partial W^V_{t+1}}{\partial S_{t+1}}\right) \tag{30}$$

where the last factor is given by equation (27). Suppose now that in every subsequent period, the social solution is approximately equal to the decentralized solution. With a linear meeting technology, the last term in (27) is then close to zero (from the first order condition in the decentralized solution), so the dynamic externalities simplify to

$$D_{xt} \approx \beta(1-p_t)km'(x_t)S_tI_t\left(-\frac{\partial W^V_{t+1}}{\partial I_{t+1}} + \beta\left(1-p_{t+1}\right)\frac{\partial W^V_{t+2}}{\partial I_{t+2}}\right). \tag{31}$$

The intuition is now clear. When the deviation between the private and social solutions is not too large, the dynamic externalities are approximately given by the difference between the impact that more infections have on welfare next period, net

[15] See the pioneering work of Peter Diamond and Eric Maskin (1979), in which they explore meetings with linear and quadratic "search technologies."

[16] The example in the Appendix satisfies this property for given number of social spaces

of the impact that more infections have a period later. The reason for the change of sign is that a fall in infections this period, due to more social distancing, raises the susceptible population next period. A higher susceptible population next period raises the infections the period after, working against the containment of infections next period. The first term on the right of (31) picks up the contagion and medical externalities and if it acted alone it would require more social distancing. The second term is the immunity externality. The social planner will want to increase social activity to reduce the susceptible population faster towards the target needed for the herd immunity.

It is not possible to sign the net effect without quantitative restrictions, even for this approximation. Without it the last term in (27) that we cancelled out could be substantial, in which case signing the overall effect would be even more difficult. Although it would appear that in (31), and if the two marginal effects of higher I_{t+1} and I_{t+2} are approximately the same, the negative one would dominate because the positive one is discounted by $\beta(1-p_{t+1})$, the simulations in section 6 show that this could easily be reversed.

5 Vaccination

In this section we introduce the possibility of a sudden end to the epidemic, which we attribute to the discovery of an effective vaccine. Clearly, the possibility of an end to infections will influence the behaviour of forward-looking agents. If an individual succeeds to remain susceptible when the epidemic then she would have avoided an infection for good. We should therefore expect that for a given vector of state variables, the prospect of a vaccine makes agents cut back on social activity, as they have more to gain by waiting. For the same reason it also reduces the planner's preferred activity level. In addition, for the planner, the prospect of a vaccine reduces the positive immunity externalities associated with a higher activity level, so the planner might increase social distancing by more than forward-looking agents do in response to the vaccine.

We assume that the introduction of a vaccine follows a Poisson process. The probability that a vaccine arrives between two consecutive periods is denoted by a constant λ. The vaccine is 100 percent efficient. Individuals already infected have no use for the vaccine, as it is no cure.

Denote by $\overline{W}_t^{VV}$ the net present value of utility of a vulnerable agent in period t if a vaccine arrives between $t-1$ and t, and W_t^{VN} the utility if it does not arrive. Define $W_t^{VV} \equiv \lambda \overline{W}_t^{VV} + (1-\lambda) W_t^{VN}$. It follows that

$$\overline{W}_t^{VV} = \max_{x_t} \left\{ \phi(x_t) + \beta(1-p_t) W^R + \beta p_t (1-\delta(I_t)) W^R \right\} \tag{32}$$

Both in the market solution and the planner's solution optimality obtains at the highest

level of x_t, $\phi'(x_t) = 0$. This yields, noting that $\phi(x_r) = (1-\beta)W^R$,

$$\overline{W}_t^{VV} = [1 - \beta p_t \delta(I_t)]\, W^R. \tag{33}$$

Now suppose the vaccine has not yet arrived in period t. With predetermined probability p_t, the individual is already infected, and her continuation value is $(1-\delta(I_t))W^R$, independently of whether or not the vaccine arrives. With the complementary probability, the agent is not infected, and given that a vaccine might arrive next period, obtains a continuation pay-off of W_{t+1}^{VV}. It follows that

$$W_t^{VN} = \max_{x_t} \left\{ \phi(x_t) + \beta p_t \left(1 - \delta(I_t)\right) W^R + \beta(1-p_t) W_{t+1}^{VV} \right\} \tag{34}$$

An increase in λ increases W^{VV}, and hence increases the utility loss associated with getting the disease.

5.1 The decentralized solution

Since $W_t^{VV} \equiv \lambda \overline{W}_t^{VV} + (1-\lambda) W_t^{VN}$, it follows from (33) and (34) that

$$\begin{aligned} \frac{\partial W_t^{VV}}{\partial p_t} &= -\lambda\beta\delta(I_t)W^R - (1-\lambda)\beta\left[W_{t+1}^{VV} - (1-\delta(I_t))W^R\right] \\ &= -\beta\left[W_{t+1}^{VV} - (1-\delta(I_t))W^R\right] - \lambda\beta(W^R - W_{t+1}^{VV}). \end{aligned} \tag{35}$$

It follows that the first order condition for the maximum of the agents in the market is (from 34),

$$\phi'(x_t) = (1-p_t)\beta^2 p'(x_t) \left\{ \left[W_{t+2}^{VV} - (1-\delta(I_{t+1}))W^R\right] + \lambda(W^R - W_{t+2}^{VV}) \right\} \tag{36}$$

The first term inside the brackets on the right corresponds to the term we had earlier, and reflects the cost of being infected if a vaccine does not materialize before the next period. The last term reflects the gain to the susceptible if a vaccine materializes, which the infected will not take part in.

The decentralized equilibrium can be defined analogously with the definition with no vaccine, given by Definition 1, with W_t^V replaced by W^{VV} and $\overline{W}^{VV}$, and with the first order condition for x_t, (14), replaced by (36).

5.2 The planner

Before a vaccine arrives, the planner maximizes the net present value of utility of the representative vulnerable person W^{VN} defined by (34). The planner, in contrast

with the agents in the decentralized solution, considers the laws of motion (18)-(20) as endogenous. Parallel to (37), the first order condition for x_t reads

$$-\frac{\phi'(x_t)}{1-p_t} = \beta\left\{\frac{\partial W^{VV}_{t+1}}{\partial p_{t+1}}\frac{\partial p_{t+1}}{\partial x_t} + \frac{\partial W^{VV}_{t+1}}{\partial I_{t+1}}\frac{\partial I_{t+1}}{\partial x_t} + \frac{\partial W^{VV}_{t+1}}{\partial S_{t+1}}\frac{\partial S_{t+1}}{\partial x_t}\right\} \tag{37}$$

The factor of the first term inside the brackets, $\frac{\partial W^{VV}_{t+1}}{\partial p_{t+1}}$, is given by (35). The two last terms in (37) can be derived in exactly the same way as when there was no vaccine, taking into account that the probability that there is no vaccine in the next period is $1-\lambda$, and two periods ahead is $(1-\lambda)^2$. It follows that

$$\begin{aligned}
\frac{\partial W^{VV}(S_t, I_t, p_t)}{\partial I_t} &= -(1-\lambda)\phi'(x_t)\frac{m(x_t)}{m'(x_t)I_t} - \beta p_t \delta'(I_t) W^R \qquad (38)\\
\frac{\partial W^{S}(S_t, I_t, p_t)}{\partial S_t} &= \frac{(1-\lambda)(1-p_{t+1})}{p'_{t+1}S_t}\left[\phi'(x_t) + \beta(1-p_t)p'_{t+1}\frac{\partial W^{VV}_{t+1}}{\partial p_{t+1}}\right]\\
&\quad +(1-\lambda)\beta(1-p_t)\frac{\partial W^{VV}_{t+1}}{\partial I_{t+1}}
\end{aligned}$$

We have already noted that since the prospect of a vaccine increases W^{VV}, it makes the agents more cautious in the decentralized equilibrium, and therefore increases social distancing. The prospect of a vaccine reduces the dynamic externalities by a factor $1-\lambda$, except for the medical congestion externality, which depends on the number of infected agents this period, before the vaccine arrives. But if a vaccine arrives between this period and the next, it will make both the contagion effect of a higher I_t in period $t+1$ and the immunity externality of the lower S_{t+1} in $t+2$ irrelevant. The immunity externality is reduced by a factor $(1-\lambda)^2$, because it arrives two periods later. This explains our conjecture that the prospect of obtaining a vaccine will reduce the planner's optimal activity level more than the activity level of private agents in the decentralized equilibrium and it will reduce the race to shield and the prospect of the social planner increasing social distancing.

6 Simulations of Planner's Problem and Externalities

In section 4 we showed how the decentralized solution with optimizing agents differs from the standard SIR model without optimization. In this section we use similar parameters to highlight the main differences between the decentralized solution and the planner's social optimum. We focus on three factors that drive a wedge between the two sets of solutions. First the role of medical costs and their implications for the dynamic externalities, second the role of a vaccine and finally the role of increasing returns in matching that drive a wedge between the static solutions of the two models.

6.1 The importance of medical costs

Catching the disease involves some cost to the individual. In the absence of deaths, we modelled the cost as a proportion of lifetime returns after recovery, which is borne instantaneously, between the infectious and recovered states. In general the medical cost will depend on how many other people are sick at the same time, because of limited medical facilities. This dependence, which is ignored by private agents, drives a wedge between the decentralized and planning solution, which we called the medical congestion externality. But the overall medical cost, independently of any dependence on the number of infected people, influences the dynamic externalities which also drive a wedge between the decentralized and planning solutions.

Let the medical costs in general be $\delta_0 \exp(\delta_1 I_t)$. The parameter δ_1 picks up the medical congestion cost and we set it equal to zero for the moment, to focus on the impact of the overall cost. The total medical cost is therefore δ_0 and we simulate the solution for two levels of δ_0, a low one of 0.02 and a high one of 0.1. The other parameters are similar to the ones that we used in section 3 and they are listed in Table 2. As before the utility function is $\phi(x) = A + \log(x) - x$, and transitions are governed by $km(x_t) = R_0 x_t$, with basic reproductive number $R_0 = 2.4$. With no medical congestion externality and constant returns to scale in contacts there are two externalities that drive a wedge between the two solutions, the contagion externality and the immunity one. The contagion one arises because more infected people next period give rise to more infections the period after, something ignored by private agents. If this externality dominates, the social planner will want to impose more social distancing. The immunity externality arises because by shielding, private agents slow down the transition to herd immunity. If this externality dominates, the social planner will want to reduce social distancing.

We simulated the solutions for two levels of costs because for low costs the contagion externality dominates but for high costs the immunity one dominates. Figure 3 shows the net deviation between the two dynamic paths with $\delta = 0.02$. As in the simulations of section 3, the dotted (red) line refers to the decentralized outcome while the continuous (blue) line refers to the central planner solution. The top chart shows the activity level $\{x_t\}$, the middle chart the mass of susceptible agents and the bottom chart the mass of infected people.

From Figure 3 we see that the planner tends to impose more social distancing than the decentralized equilibrium, particularly around the peak of the disease (when I_t is at its maximum) and in the beginning of the recovery (Table 2). Figure 4 separates the immunity and contagion externalities associated with a one unit increase in x_t around the optimal solution, measured in units of utility. To give some idea of the magnitude of the externalities, the per period utility in the first-best solution with no virus present, $\phi(x_r)$, is equal to 3/2. So a net externality of 0.5 is equal to 1/3 of the per period utility in the first best solution. The Figures show that the immunity and the contagion externalities are similar in the first few periods, but in later periods the

negative contagion externality dominates.

This contrasts sharply with the impact of higher medical costs. For a fixed cost $\delta_0 = 0.1$, Figure 5 shows that the immunity externality clearly dominates. In the midst of the epidemic (except at the very top) the planner now wants to raise the social activity level chosen in the decentralized solution, because of a strong immunity externality. Private agents shield too much when the medical costs are high, delaying the transition to herd immunity. The reason is that forward-looking agents know that when the disease is eradicated there will be a substantial proportion of the population who will never catch the disease. In our simulations, this proportion is around 40%. With high costs on the 60% who do catch it, private agents withdraw from the market in the hope that the disease will be eradicated by infecting others. But if all agents did the same there would be no gain in terms of avoidance of the disease, only a delay in reaching herd immunity. This is what we termed the rat race to shield. It arises because private agents ignore the impact of their shielding on the mass of susceptible agents. The social planner will want to get rid of this rat race by forcing lower social distancing. It is clear from Figure 6 that the main difference from the case of low medical cost is the large rise in the immunity externality caused by the high medical cost. Note that the planner imposes more social distancing in the very first periods of the epidemic and towards the end of the epidemic. The latter reflects the fact that the planner wants to reduce the total number of infected people before herd immunity is reached, and obtains this by reducing the last elements of social distancing slowly. As a result, the total number of infected people is lower than in the decentralized solution, but only marginally because it is already close to the maximum allowable in the decentralized solution.

Figure 7 shows phase diagrams for the optimal path, the decentralized equilibrium path, and the equilibrium in the original SIR model with x_t fixed at 1 in all periods and $\delta_0 = 0.1$ in all cases. We see that the phase diagram for the optimal solution moves more quickly than the decentralized in the early stages of the epidemic. It eventually converges to a steady state with herd immunity at a slightly higher level of disease-free agents. As in the earlier simulations, the SIR trajectory converges to an equilibrium with a much lower number of susceptible individuals than both the market solution and the optimal solution do.

We finally consider the impact of medical congestion externalities. For fixed medical costs, the planner will want to increase social activity, which leads to a peak of infections early on (Figure 5). But with a medical congestion externality present, this will not be optimal. A peak in infections overcrowds hospitals and raises the medical costs. The social planner will now want to flatten the infections curve, so as to avoid high peaks of medical costs. This is shown in Figure 8. We parameterize $\delta(I)$ by writing $\delta(I) = \delta_0 e^{\delta_1 I_t}$. We set $\delta_0 = 0.07$ and $\delta_1 = 6$, so that medical costs increase fast with I_t and the mean overall cost between $I_t = 0$ and I_t at its maximum in the absence of the medical externality is approximately 0.1, as in the case of fixed medical costs.

Figure 8 shows that the trajectories for the planner and the decentralized solution

are now similar. The similarity is a coincidence with the particular set of parameters that we selected for the simulations. In the absence of medical congestion externalities, agents shield more than the social planner because of the rat race to shield, and in the presence of the medical congestion externality, they shield less because they ignore the impact of their illness on the medical costs. The two happen to offset each other for the medical costs $\delta_0 = 0.07$ and $\delta_1 = 6$. Other sets of medical cost parameters could yield different results; in general, we conjecture that a higher δ_0 will have a bigger impact on private shielding and a higher δ_1 will have a bigger impact on social shielding.

6.2 Implications of a vaccine

In an epidemic there is always a hope that a vaccine will be discovered that will end the disease before herd immunity is reached. In the simulations that follow we investigate how the dynamic paths that we derived so far are affected by this expectation before the vaccine actually arrives. As we explained in the section preceding this one, the expectation of a vaccine is likely to increase shielding by both planner and private agents, as they forgo social activities until medical research comes up with a solution. Importantly for the externalities that we simulated in the first part of this section, the social planner will not have as strong an incentive to increase social activity to reach immunity faster. The immunity externality pushes more people to infection in the short run, to avoid the long delay to herd immunity that the rat race to shield yields.

In Figure 9 and Table 3 we assume that the per period probability of obtaining a vaccine is .05. With this probability the mean duration of time before a vaccine arrives is 20 periods, or approximately 40 weeks. The lines shown in the Figure are calculated conditional on the vaccine not arriving. As the model predicts, the activity level in both the optimal solution and the decentralized solution shifts down relative to the situation without the prospect of a vaccine arriving. However, there is a dramatic difference in how the two lines change. In Figure 9 the planner is now reducing social distancing much more than the agents in the decentralized solution. This is due to the big fall in the wedge caused by the immunity externality. With the prospect of a vaccine arriving within a year, the contagion externality is the main one that drives a wedge between the planning and the decentralized solution. This explains the planner's much more cautious approach to social distancing in this version of the model than in the one without the prospect of a vaccine. The interaction between the two externalities is shown in Figure 10, which should be contrasted with Figure 6. The prospect of a vaccine virtually eliminates the immunity externality, while it increases the absolute value of the contagion externality

Of course, with this very cautious policy, in the absence of a vaccine, herd immunity is much harder to achieve. Even after 100 periods without a vaccine, the fraction of susceptible individuals in the optimal solution is around .8, and society is still far from herd immunity.

6.3 Increasing returns to scale in contacts

So far the simulations of this section focused on the dynamic externalities, the medical, contagion and immunity. The static externalities were shut out by imposing constant returns to scale in the contact technology. We now relax this assumption and derive the implications of increasing returns, which as we argued in section 4, we consider to be *a priori* more plausible.

In the theory section it was possible to isolate the impact of increasing returns by shutting out the dynamic externalities. In the simulations, however, this cannot be done. As inspection of equation (28) shows, the solution to be simulated depends on future present-discounted values, which are influenced by the dynamic externalities. So we approach the problem as follows. For each problem, the decentralized maximization by agents and the social optimum of the planner, we simulate two equilibrium paths, one with linear technology as before and one with quadratic technology; i.e., we now write the contact function: $km(x_t) = R_0 x_t^2$. As in the simulation of fixed medical externalities in this section, we set $\lambda = 0$ (no vaccine) and $\delta(I) \equiv .1$ (no medical congestion externalities).

With a quadratic contact function it is straightforward to derive, by substituting the quadratic expression for $m(x_t)$ into (16) and (17), that for private agents $p'(x_t^p) = R_0 x_t^p I_t$ whereas for the planner, $p'(x_t^s) = 2R_0 x_t^s I_t$. We argued that because of this difference and ignoring the dynamic externalities, the planner will want to impose more social distancing. But dynamic externalities are present, and applying the quadratic contact technology in the dynamic expressions (38) shows that both are reduced by a factor of 2.

In Figure 11 we compare the paths of the decentralized equilibrium with $\alpha = 1$ and $\alpha = 2$. Because with $\alpha = 2$ the lower x_t^p reduces the probability of infection by more ($R_0 x_t^p I_t$ versus $R_0 I_t$), the private agent reduces social activity by less when activity is already low but this reverses when activity rises towards unity. Also, with $\alpha = 2$ the impact of social distancing on the effective reproductive number is bigger and so the peak reached by infections is lower than with $\alpha = 1$. Perhaps the most interesting result in this comparison, however, is the herd immunity state. Despite these changes in private activity, the herd immunity state is reached with some delay with $\alpha = 2$ but the terminal condition, indicating the cumulative number of infections, is the same as in the case $\alpha = 1$. Overall, although increasing returns to scale alter some of the dynamics of the decentralized equilibrium, the equilibrium outcomes do not differ in important ways from the much simpler case of constant returns. Partly this can be explained by the fact that even for $\alpha = 1$, herd immunity is obtained with close to the highest possible number of susceptible individuals.

There is a more substantial impact of increasing returns on the social solution, shown in Figure 12. Although the impact on activity is qualitatively the same as in the private solution, there is enough quantitative difference to lead to a substantial difference in the number of susceptible individuals for very many periods. With $\alpha = 2$,

the effect of reducing x_t below 1 in terms of reduced number of contacts is twice as high as when $\alpha = 1$. This induces the planner to set x_t somewhat below 1 for an extensive number of periods. As a result, herd immunity (which obtains for $S_t < R_0^{-1} = 0.417$) is not reached after 100 periods, at which more than 50% of the population is still susceptible. With $\alpha = 1$, the new steady state is reached more quickly, and herd immunity is obtained within 100 periods. As before, adjustment to the herd immunity state is much faster in the social planning solution in the relatively early stages of the epidemic, because of the race to shield and the immunity externality, which are strong with the fixed medical cost of $\delta_0 = 0.1$.

7 Conclusions

We have shown that the differences between three alternative paths from the onset of an epidemic to its eradication can be large and complex. The first path, commonly found in the epidemiological literature, exposes large numbers of individuals to the disease and reaches herd immunity quickly. An important result of our paper is that the other two paths, chosen by private agents in a decentralized equilibrium or by a social planner, favour much longer adjustment paths by restricting the number of infections. By limiting per-period infections the other two paths achieve herd immunity close to the maximum number of susceptible agents (who remain free of the disease) consistent with herd immunity.

But such a large number of susceptible individuals at herd immunity produces a "race to shield," in which agents shield too much at the midst of the disease to increase the chances that they will be the lucky ones who will avoid the disease altogether. Obviously not everyone can succeed in this race, so a social planner wants to avoid it by choosing more social activity.

The expectation of a vaccine makes a large difference to the results. When private agents expect that a vaccine will arrive they substantially reduce the race to shield. But they still ignore the fact that if they contract the disease they will infect others in the future. The social planner takes this contagion externality into account, and imposes more social distancing, to avoid more infections before the vaccine arrives. The ranking result that we derived with a race to shield reverses.

We illustrate our results with simulations but as more data become available the model should be taken to the data. Our model applies generally to epidemiological models in the SIR tradition of Kermack and McKendrik (1927) and shows that quantitatively some parameters make a large difference to the simulated paths, such as the lifetime cost of the disease, the elasticity with which contacts respond to changes in social action and the vaccine arrival probability. Information on their relative magnitudes is still scant and may be disease-dependent. As our assumptions are consistent with many with the features of COVID-19, data for this epidemic could shed light on these magnitudes. Another feature that needs to be taken to the data is the economic

cost of shielding. In this paper we assume that it is a simple convex function of a single variable, our measure of social action, but as many authors have shown, it depends on a multitude of features of occupations and economic structures.[17]

Finally, our analysis is based on the assumption that recovered individuals stay immune permanently. If recovered individuals may lose their immunity as time goes by, this will change our model in several ways. Herd immunity will never be obtained, and the new steady state will be characterized by equal flows of individuals moving out of and into the susceptible state. We conjecture that without a vaccine the activity level never recovers to its first best level. Generally, when immunity is temporary, social distancing will be larger. In addition the immunity externality weakens, so the planner increases social distancing more than the agents do in the decentralized solution. These are topics that should explored in future work.

[17]Some pertinent references are given in our introduction to this paper.

Table 1: Parameters used in simulation (decentralized equilibrium)

Parameter	Notation	Value
Preferences		
Discount Factor	β	0.998
Constant	A	1.500
Infection Characteristics		
Basic Reproductive Rate	R_0	2.400
Infection cost	δ_0	0.100
Initial infection	ϵ	0.001
Contact Technology		
Log-Linear	α	1.000
Main activity results		
Activity at the outbreak of the epidemic	x_0	0.868
Activity when infections are at maximum	x^I	0.430
Activity of the recovered	x_r	1.000
In standard SIR $x = x_r$ throughout the epidemic.		

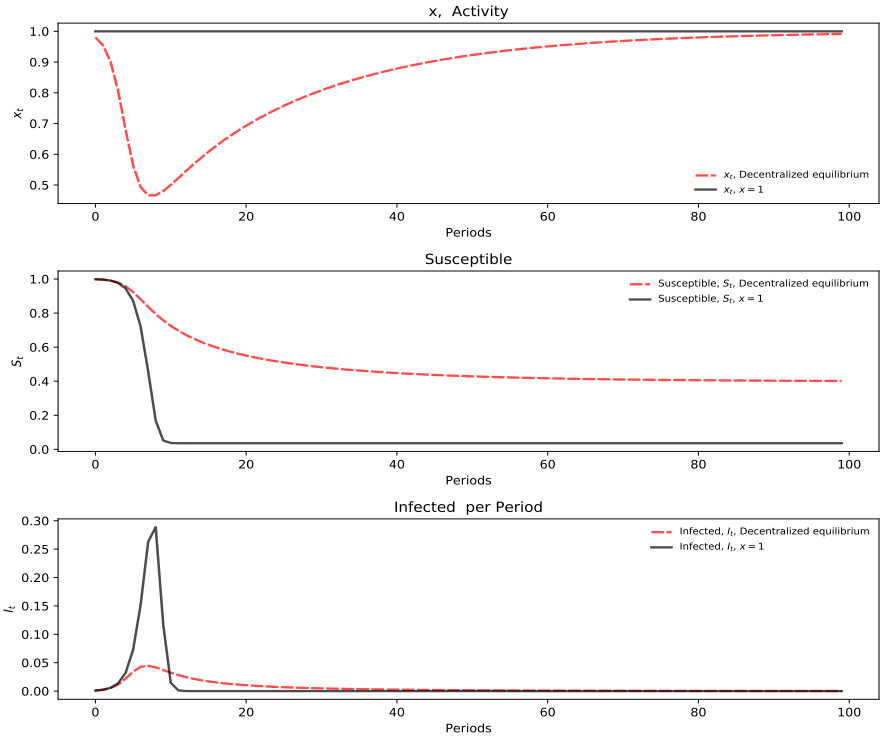
Table 2: Basic Central Planner Simulations

Parameter		Notation	low δ_0	high δ_0
Preferences				
Discount Factor		β	0.998	0.998
Utility constant		A	1.500	1.500
Infection Characteristics				
Basic Reproductive Rate		R_0	2.400	2.400
Infection cost		δ_0	0.020	0.100
Medical congestion		δ_1	0.000	0.000
Probability of vaccine		λ	0.000	0.000
Initial infection		ϵ	0.001	0.001
Contact Technology				
		α	1.000	1.000
Activity				
Activity at the outbreak of the epidemic		x_0	0.989	0.959
Activity when infection is maximal		$x_{I_{\max}}$	0.531	0.520
Activity of the recovered		x^r	1.000	1.000
Contagion Externality		X_I		
at $t = 0$	level		-0.015	-0.048
	% deviation		-0.006	-0.020%
at extremum	level		-0.914	-1.116
	% deviation		-0.370	-0.467%
Immunity Externality		X_S		
at $t = 0$	level		0.007	0.026
	% deviation		0.003	0.011%
at extremum	level		0.713	3.944
	% deviation		0.289	1.665%
Net Externality		$X_I + X_S$		
at $t = 0$	level		-0.008	-0.021
	% deviation		-0.003	-0.009%
at extremum	level		-0.517	3.017
	% deviation		-0.208	1.274%

Table 3: Planner Simulation with Vaccine

Parameter		Notation	Value
Preferences			
Discount Factor		β	0.998
Utility Constant		A	1.500
Infection Characteristics			
Basic Reproductive Rate		R_0	2.400
Infection cost		δ_0	0.100
Medical congestion		δ_1	0.000
Probability of vaccine		λ	0.050
Initial infection		ϵ	0.001
Contact Technology			
		α	1.000
Activity			
Activity at the outbreak of the epidemic		x_0	0.509
Activity when infection is maximal		$x_{I_{\max}}$	0.433
Activity of the recovered		x^r	1.000
Contagion Externality		X_I	
at $t = 0$	level		-0.940
	% deviation		-0.385%
at extremum	level		-1.242
	% deviation		-0.509%
Immunity Externality		X_S	
at $t = 0$	level		0.022
	% deviation		0.009%
at extremum	level		0.062
	% deviation		0.025%
Net Externality		$X_I + X_S$	
at $t = 0$	level		-0.918
	% deviation		-0.376%
at extremum	level		-1.180
	% deviation		-0.484%

Figure 1: Dynamics of the Epidemic in Optimizing SIR: 100 periods



.

Figure 2: Joint Phase Diagram for Optimizing SIR and Standard SIR

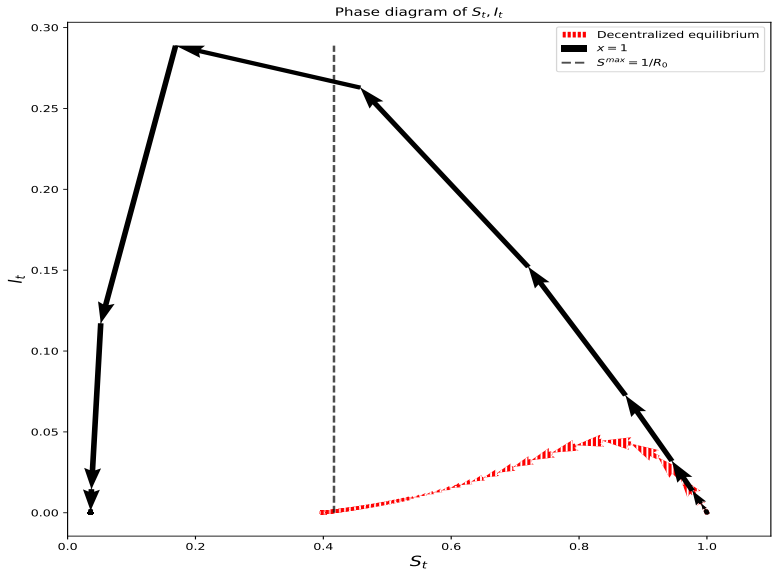


Figure 3: Matching-SIR: Planner vs market, low δ (0.02)

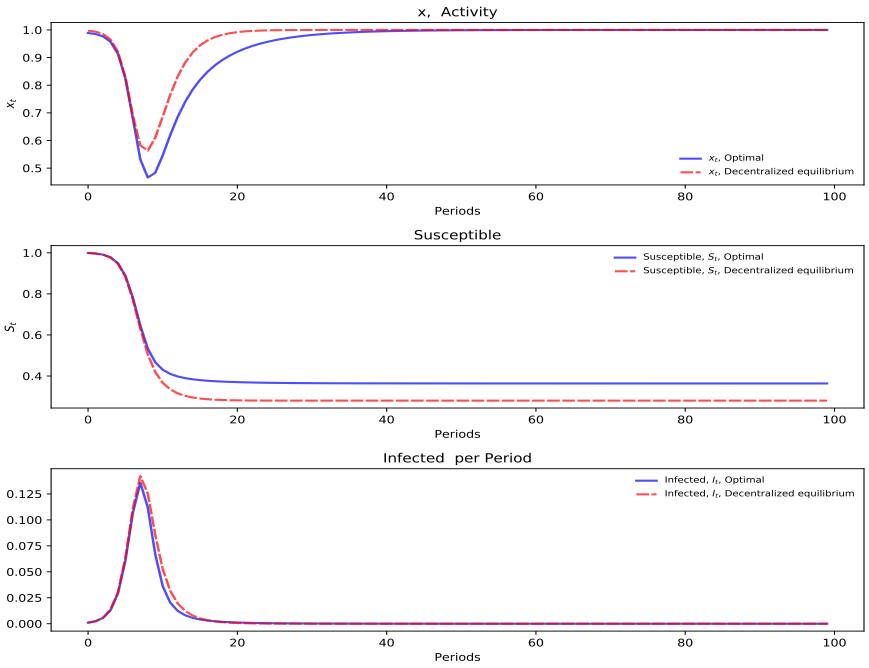


Figure 4: Externalities per unit of x at planner solution, low δ (0.02)

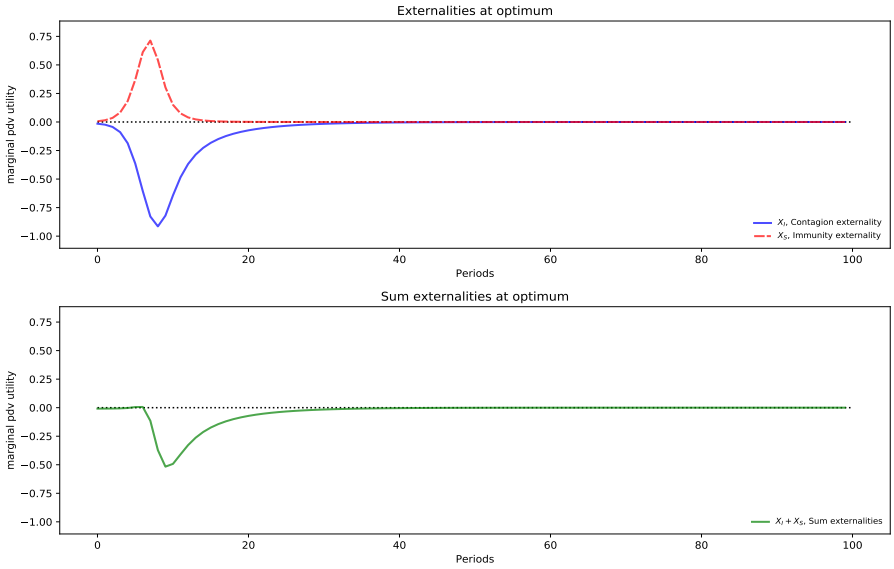


Figure 5: Matching-SIR: Planner vs market, high δ (0.1)

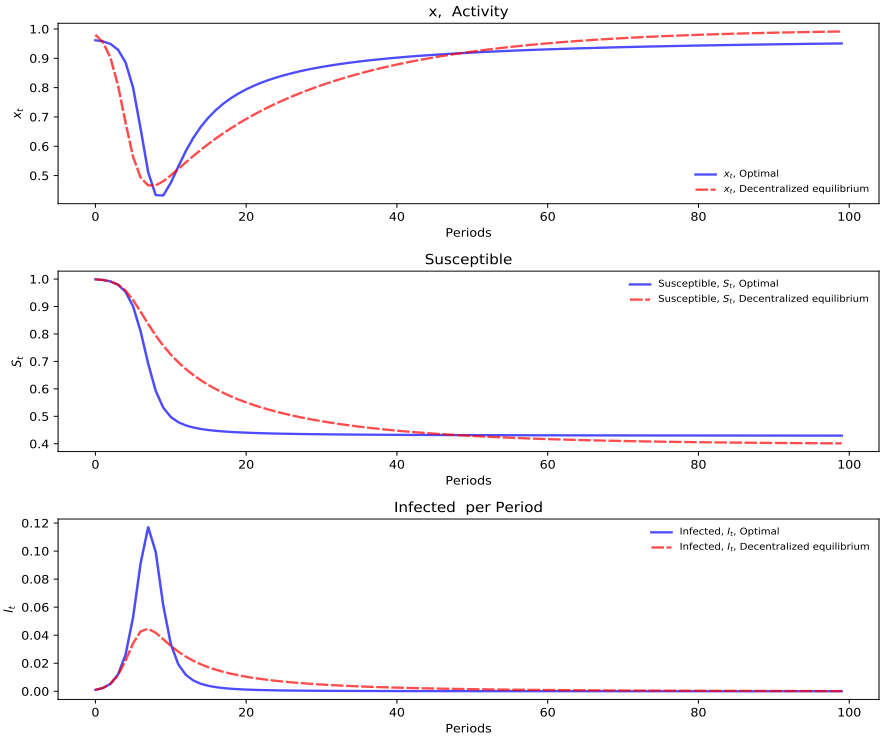


Figure 6: Externalities per unit of x at planner solution, high δ

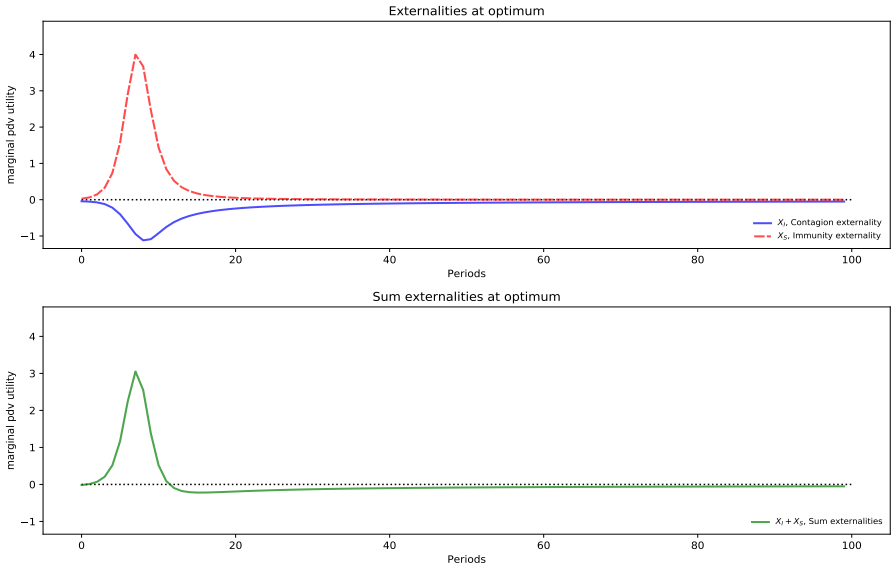


Figure 7: Phase diagram, high δ

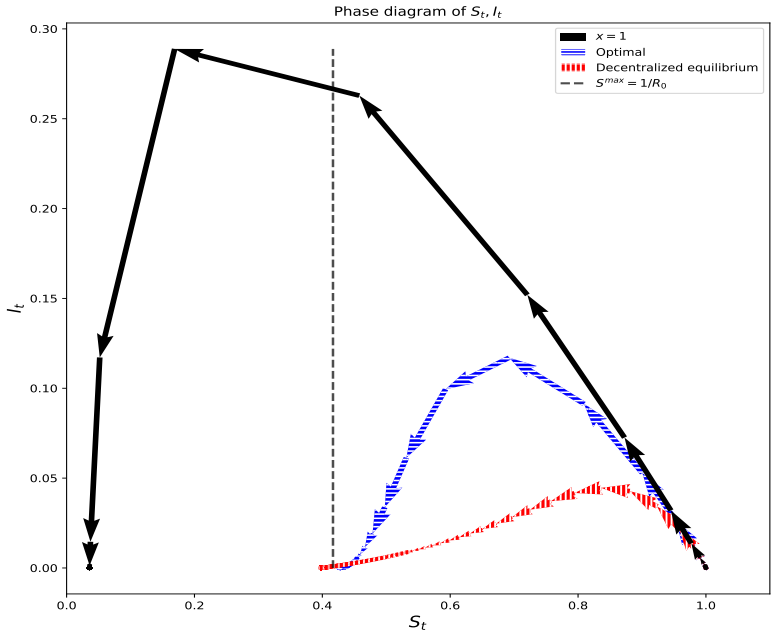


Figure 8: Planner vs market with medical externalities

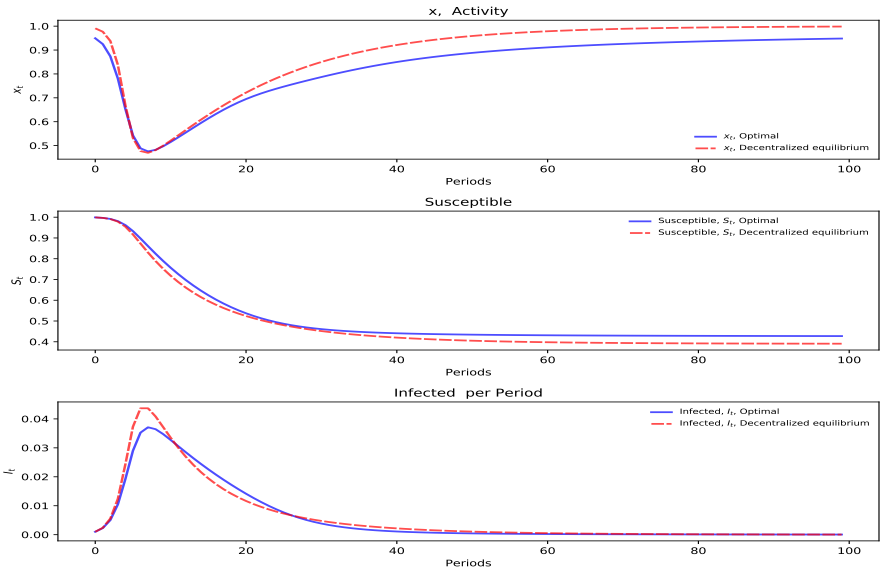


Figure 9: Planner vs market with vaccine, $\lambda = 0.05$

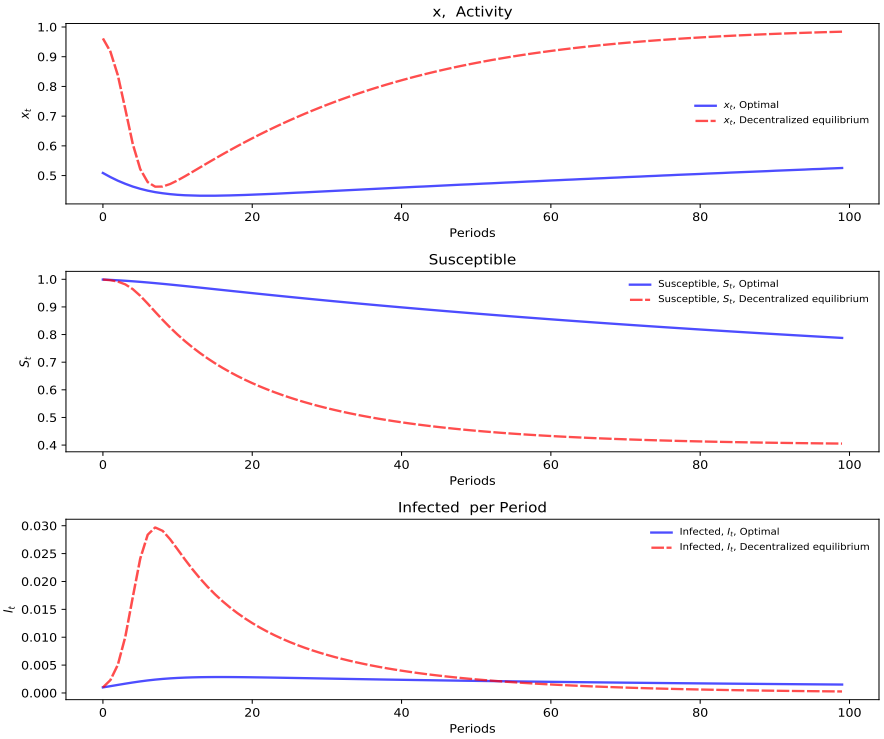


Figure 10: Externalities per unit of x at planner solution with vaccine, $\lambda = 0.05$.

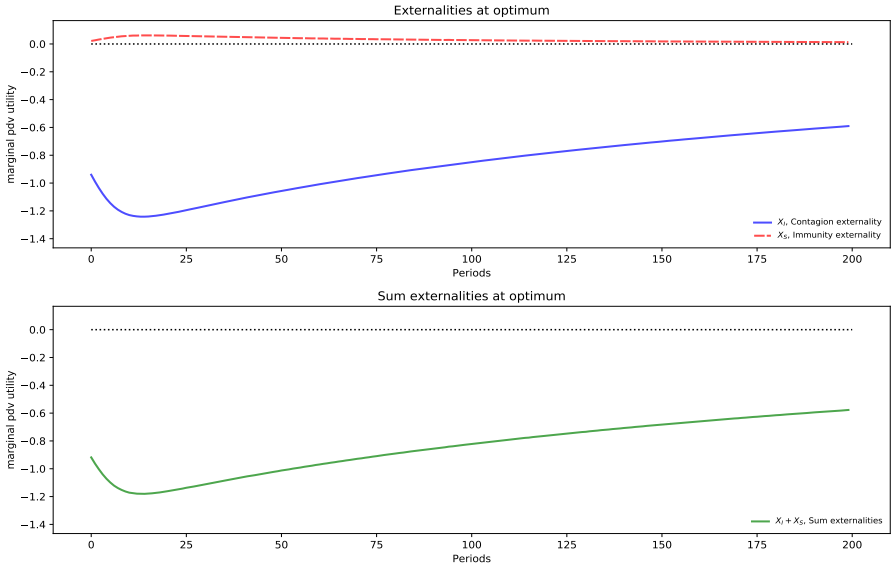


Figure 11: Decentralized equilibrium with constant and increasing returns to scale, high δ, no vaccine or medical externalities

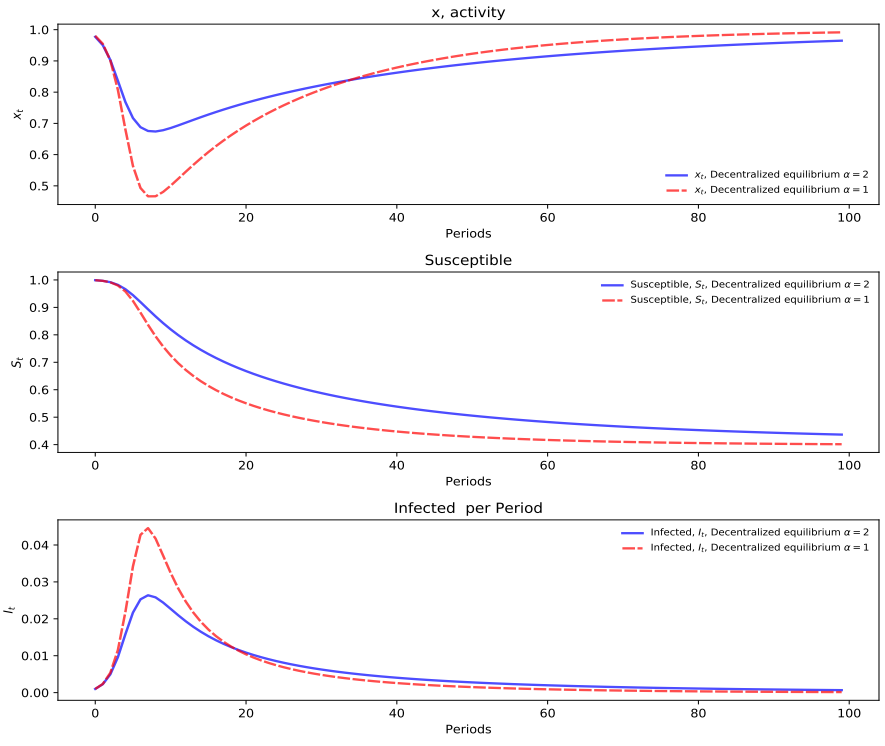
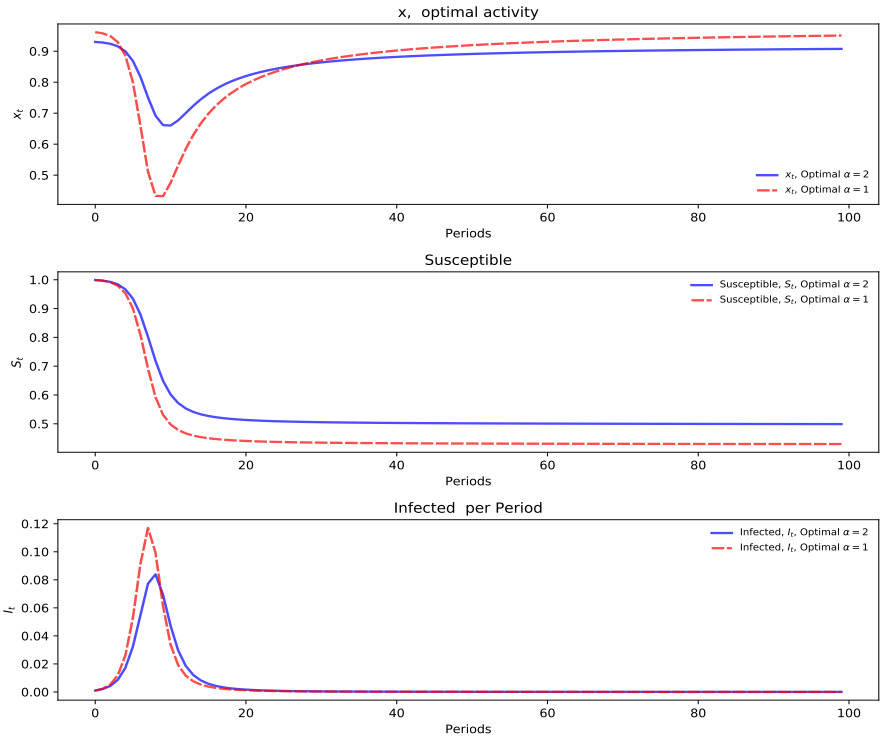


Figure 12: Central Planner equilibrium with constant and increasing returns to scale



References

[1] Acemoglu, Daron, Victor Chernozhukov, Ivan Werning, and Michael D. Whinston (2020), "A Multi-Risk SIR Model with Optimally Targeted Lockdown", *NBER Working Paper N. 27102.*

[2] Alvarez, Fernando, David Argente, and Francesco Lippi (2020), "A Simple Planning Problem for COVID-19 Lockdown", *NBER Working Paper N. 26981.*

[3] Alfaro, Laura, Faia, Ester, Lamersdorf, Nora, and Farzad Saidi (2020), "Social Interactions in Pandemics: Fear, Altruism and Reciprocity", *NBER Working Paper N. 27134.*

[4] Aspri, Andrea, Beretta, Elena, Gandolfi, Alberto, and Etienne Wasmer (2020), "Mortality Containment vs. Economics Opening: Optimal policies in a SEIARD model", *Covid Economics: Vetted and Real Time Papers* , 25, 3 June, 55-89.

[5] Berger, David W, Kyle F Herkenhoff, and Simon Mongey (2020), "An SEIR Infectious Disease Model with Testing and Conditional Quarantine", *NBER Working Paper N. 26901.*

[6] Brotherhood, Luiz, Kircher, Philipp, Santos, Cezar, and Michéle Tertilt (2020), "An Economic Model of the COVID-19 Epidemic: The Importance of Testing and Age-Specific Policies ", *CEPR Discussion Paper N. 14695.*

[7] Chen, Frederick (2012), "A Mathematical Analysis of Public Avoidance Behavior During Epidemics Using Game Theory", *Journal of Theoretical Biology*, 302: 18-28.

[8] Diamond Peter, and Eric Maskin (1979), "An Equilibrium Analysis of Search and Breach of Contracts, I: Steady States", *Bell Journal of Economics*, 10 :282-316.

[9] Diamond, Peter A. (1982), "Aggregate Demand Management in Search Equilibrium", *Journal of Political Economy*, Vol. 90, No. 5, pp. 881-894.

[10] Eichenbaum, Martin S., Rebelo, Sergio, and Mathias Trabandt (2020), "The Macroeconomics of Epidemics" *NBER Working Paper No. 26882.*

[11] Farboodi, Maryam, Gregor Jarosch, and Robert Shimer (2020), "Internal and External Effects of Social Distancing in a Pandemic", *NBER Working Paper N. 27059.*

[12] Favero, Carlo, Andrea Ichino, and Aldo Rustichini (2020), "Restarting the Economy while Saving Lives Under COVID-19", *Technical Report 3580606, SSRN.*

[13] Garibaldi, Pietro, Espen R. Moen, and Christopher A. Pissarides (2020), "Modelling Contacts and Transition in the SIR Epidemic Model", *Covid Economics: Vetted and Real Time Papers*, 5, 16 April, pp. 1-20.

[14] Greenwood, Jeremy, Philipp Kircher, Cezar Santos and Michéle Tertilt (2019), "An Equilibrium Model of the African HIV/AIDS Epidemic ", *Econometrica*, 87(4): 1081-1113.

[15] Gregory, Victoria, Menzio, Guido, and David Wiczer (2020), "Pandemic Recession: L or V Shaped?", *Quarterly Review*, 40: 1-28.

[16] Hosios, Arthur J. (1990), "On the Efficiency of Matching and Related Models of Search and Unemployment ", *The Review of Economic Studies*, 57(2): 279-298.

[17] Kapicka, Marek, and Peter Rupert (2020), "Labor Markets During Pandemics", Working Paper.

[18] Kaplan, Greg, Benjamin Moll, and Gianluca Violante (2020), "The Great Lockdown and the Big Stimulus: Tracing the Pandemic Possibility Frontier for the U.S. ", Working Paper.

[19] Kermack, William O., and Anderson G., McKendrick (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*, Volume 115, Issue 772. 700-721.

[20] Kremer, Michael (1996), "Integrating Behavioral Choice into Epidemiological Models of AIDS", *The Quarterly Journal of Economics*, 111(2): 549-573.

[21] McAdams, David (2020), "Nash SIR: An Economic-Epidemiological Model of Strategic Behavior During a Viral Epidemic ", Working Paper.

[22] Nævdal, Eric (2020) "Epidemics and Increasing Returns to Scale on Social Distancing", *Covid Economics: Vetted and Real Time Papers*, 39, July, 139-148.

[23] Petrongolo, Barbara and Christopher A., Pissarides (2001), "Looking into the Black Box: A Survey of the Matching Function", *Journal of Economic Literature*, Vol. 39, No. 2, (pp. 390-431).

[24] Quercioli, Elena and Lones Smith (2006), "Contagious Matching Games", Working Paper.

[25] Pissarides, Christopher A. (2000), *Equilibrium Unemployment Theory*, MIT Press, Cambridge, Mass., USA.

[26] Rowthorn, Robert and Flavio Toxvaerd (2012), "The Optimal Control of Infectious Diseases via Prevention and Treatment ", *CEPR Discussion Paper No. DP8925.*

[27] Rachel, Lukasz (2020), "An Analytical Model of Covid-19 Lockdowns", Working Paper.

[28] Sargent, Thomas J., and Stachurski John (2020), *Advanced Quantitative Economics with Python*, available on line at https://python.quantecon.org/

[29] von Thadden, Ernst-Ludwig (2020), "A Simple, Non-Recursive Model of the Spread of COVID-19 with Applications to Policy", *Technical Report 3580767, SSRN.*

[30] Weiss, H. H. (2013). "The SIR Model and the Foundations of Public Health". *Materials Matematics*,0001-17 Volume 2013, no. 3, 17 pp. ISSN: 1887-1097.

Appendix 1: The infections technology

In this Appendix we illustrate the technology of infections with a particular example that is derived from first principles and satisfies the main properties of our general function in the text. It is based on the urn-ball game that has been used in labour theory and it brings out the contrast between the labour matching function and the epidemiology transmission function.[18]

Suppose that a social activity is performed in one of N social spaces. By social space we mean a place where performing an activity requires contact with at least one other person. For a given population size $S + I + R$, N is a measure of the density of the community, with smaller N indicating a more dense community. Social distancing is also related to this measure: bigger N makes social distancing easier to achieve. A contact in social space between a susceptible and an infected individual infects the susceptible individual with probability $k \in [0, 1]$.

We now interpret social spaces as urns. Infected individuals hold white balls and susceptible individuals hold black balls. They all place $\bar{x}_t$ balls each in randomly selected urns. To simplify the exposition we assume that there is no memory of where a previous ball was placed, so each person places each one of its balls in a randomly selected urn out of the N available. A susceptible person gets infected with probability k if any one of the urns that she selected for her black balls contains one or more white balls.[19] We are interested in deriving the probability that there will be a white ball in at least one of the x_t urns selected by the $i \in S$ individual, given the $\bar{x}_t$ selected by the infected individuals.

Our assumption of no memory makes the problem equivalent to placing $\bar{x}_t I_t$ white balls at random in urns. The probability that an urn avoids a given white ball is $1 - 1/N$, so, since there are $\bar{x}_t I_t$ white balls, the probability that an urn contains no white balls after all have been placed is,

$$\begin{aligned} h_t &= \left(1 - \frac{1}{N}\right)^{\bar{x}_t I_t} \\ &= e^{-\bar{x}_t I_t / N}. \end{aligned} \tag{39}$$

Because of large numbers, h_t is also the fraction of social spaces that are infection-free (healthy).

We consider now how the choices of a single susceptible agent influence the probability that the agent will get infected. The agent selects x_t urns to place black balls. The probability that a single ball avoids an urn containing a white ball is given by

[18] See Petrongolo and Pissarides, 2001, p. 401-2, for discussion of the use of this game in labour theory

[19] Here is the biggest contrast with the labour matching function. A job vacancy (read infected person) can remove at most one unemployed worker from the pool of unemployment (read, susceptible). An infected person can remove any number that comes into contact with them.

(39). So the probability that all x_t black balls avoid an urn containing a white ball is

$$h_{it} = \left(e^{-\bar{x}_t I_t/N}\right)^{x_t}. \tag{40}$$

It follows that $1-h_{it}$ is the probability that a single agent meets an infected individual and so the probability that this person gets infected in period t is,

$$p_{t+1} = k\left(1 - e^{-x_t \bar{x}_t I_t/N}\right). \tag{41}$$

Differentiation of p_{t+1} yields,

$$\frac{\partial p_{t+1}}{\partial x_t}\frac{x_t}{p_{t+1}} = \frac{x_t \bar{x}_t I_t}{N}\frac{e^{-x_t \bar{x}_t I_t/N}}{1 - e^{-x_t \bar{x}_t I_t/N}}. \tag{42}$$

This is a number less than 1, in contrast to the formulation in the text, which gives unit elasticity. For small x_t and I_t it is approximately equal to 1.[20]

From (41), and since there are S_t susceptible individuals who choose $x_t = \bar{x}_t$, the aggregate infections function is

$$M_{t+1} = kS_t\left(1 - e^{-x_t^2 I_t/N}\right). \tag{43}$$

For fixed N, this function exhibits increasing returns to scale in S_t and I_t. This result is saying that bigger social spaces that have proportionally more susceptible and more infectious individuals do not have a higher infections rate, but a proportional increase in susceptible and infectious individuals in a given social space does lead to proportionally more infections. Analogous to this result is the role of N. As it measures density, a lower N indicates a more dense community and a higher infections rate for given S_t and I_t.

Another property of significance is the dependence of the aggregate infections rate on the square of the social activities of susceptible and infected individuals. For small numbers of $x_t I_t/N$, as satisfied by our model, the elasticity of infections with respect to social action is approximately 2 but for larger numbers it is lower.

[20] An alternative formulation replicates the method used in the text exactly. From (39) the fraction of black balls (susceptible people) which are placed in urns that contain at least one white ball is approximately $\bar{x}_t S_t(1-\exp(\bar{x}_t I_t/N))$. A single individual supplies a fraction $x_t/\bar{x}_t S_t$ of black balls, so the probability that this person is infected is $kx_t(1-\exp(\bar{x}_t I_t/N))$, giving the proportionality between p_{t+1} and x_t.

Appendix 2

Proof of Proposition 2

In an arbitrary period t, let (S_t^o, I_t^o, p_t^o) be an arbitrary, feasible triple of values of S_t, I_t, and p_t, with $p_t^o = \frac{I_t^o}{S_t^o + I_t^o}$. From this arbitrary starting point, let the sequence $\{S_z^o, I_z^o, p_z^o, x_z^o\}_{z=t}^{\infty}$ solve the planner's problem.

Consider first an increase in the initial condition I_t from I_t^o to $I_t^o + \rho$, while S_t and p_t stay fixed at S_t^o and p_t^o, respectively.[21]

Define the function $x^I(\rho)$ implicitly by the function $km(x^I(\rho))(I_t^o + \rho)S_t^o = I_{t+1}^o$. This function exists on an interval $[-\bar{\rho}_I, \bar{\rho}_I]$ for some $\bar{\rho}_I > 0$. Clearly $x^I(0) = x_t^o$. It follows that if the economy starts at $I_t^o + \rho, S_t^o$, and the activity level is $x^I(\rho)$, then $S_{t+1} = S_{t+1}^o$, $I_{t+1} = I_{t+1}^o$, and $p_{t+1} = p_{t+1}^o$. Furthermore,

$$\frac{dx^I(0)}{d\rho} = -\frac{m(x_t^o)}{m'(x_t^o)I_t^o} = \frac{p_{t+1}^o}{p^{o\prime}S_t} \tag{44}$$

Now define the function $\tilde{W}^V(S_t^o, I_t, p_t^o)$ (where S_t^o and p_t^o are fixed, so this is a function of I_t only) for $I^t \in (I_t^o - \bar{\rho}_I, I_t^o + \bar{\rho}_I)$ as

$$\begin{aligned} \tilde{W}^V(S_t^o, I_t, p_t^o) &= \phi(x^I(I_t - I_t^o)) + \beta p_t^o W^R(1 - \delta(I_t)) \\ &\quad + \beta(1 - p_t^o)W_{t+1}^V(S_{t+1}^o, I_{t+1}, p_{t+1}^o) \end{aligned} \tag{45}$$

Since $W^V(S_t^o, I_t, p_t^o)$ is the pay-off in optimum, and $x^I(0) = x_t^o$, it follows that

$$\begin{aligned} \tilde{W}^V(S_t^o, I_t, p_t^o) &\leq W^V(S_t^o, I_t, p_t^o) \quad \forall I_t \in (I_t^o - \bar{\rho}_I, I_t^o + \bar{\rho}_I) \\ \tilde{W}^V(S_t^o, I_t^o, p_t^o) &= W^V(S_t^o, I_t^o, p_t^o) \end{aligned} \tag{46}$$

Hence the Benveniste-Scheinkman theorem applies, and we know that

$$\frac{\partial \tilde{W}^V(S_t^o, I_t^o, p_t^o)}{\partial I_t} = \frac{\partial W^V(S_t^o, I_t^o, p_t^o)}{\partial I_t} \tag{47}$$

From this equation, and (44) and (45), it follows that

[21] Along any path, we must have that $p_t = \frac{I_t}{S_t + I_t}$ for all t. However, mathematically, the planner's maximization problem is well defined for also for initial values S_t, I_t, p_t such that $p_t \neq \frac{I_t}{S_t + I_t}$. The dynamic equations ensure that $p_z = \frac{I_z}{S_z + I_z}$ for any $z > t$. Recall further that the effect of a change in I_t and S_t through p_t is captured by the term $\frac{\partial W^V}{\partial p_t}$.

$$\frac{\partial W^V(S_t^o, I_t^o, p_t^o)}{\partial I_t} = -\phi'(x_t^o)\frac{m(x_t^o)}{m'(x_t^o)I_t^o} - \beta p_t \delta'(I_t^o)V^R \tag{48}$$

Since the starting point S_t^o, I_t^o, p_t^o is arbitrary, this shows (25).

Next, consider an increase in S_t from S_t^o to $S_t^o + \rho$, while I_t and p_t stay fixed at I_t^o, p_t^o. Define $x^S(\rho)$ implicitly by the equation $(S_t^o + \rho) - km(x^S(\rho))I_t^o(S_t^o + \rho) = S_{t+1}^o$. This function exists on an interval $(-\bar{\rho}_S, \bar{\rho}_S)$ for some $\bar{\rho}_S > 0$.

By definition it follows that $S_{t+1} = S_{t+1}^o$, $I_{t+1} = I_{t+1}^o + \rho$, and $p_{t+1} = km(x^S(\rho)))I_t^o$. Furthermore, we have that

$$\frac{dx^S(0)}{d\rho} = \frac{1 - kI_t m(x_t^o)}{km'(x_t^o)S_t^o I_t^o} = \frac{1 - p_{t+1}^o}{p_{t+1}^{o\prime} S_t^o} \tag{49}$$

Now define the function $\hat{W}^V(S_t, I_t^o, p_t^o)$ for $S_t \in (S_t^o - \bar{\rho}_S, S_t^o + \bar{\rho}_S)$ as

$$\begin{aligned} \hat{W}^V(S_t, I_t^o, p_t^o) &= \phi(x^S(S_t - S_t^o)) + \beta p_t W^R(1 - \delta(I_t^o)) \\ &\quad + \beta(1 - p_t^o)W_{t+1}^V(S_{t+1}^o, I_{t+1}^o + S_t - S_t^o, km(x^S(S_t - S_t^o))I_t^o) \end{aligned} \tag{50}$$

Now I_t^o and p_t^o are fixed, so this is a function of S_t only. Again it follows by construction that

$$\begin{aligned} \hat{W}_t^V(S_t, I_t^o, p_t^o) &\leq W_t^V(S_t, I_t^o, p_t^o) \quad \forall S_t \in (S_t^o - \bar{\rho}_S, S_t^o + \bar{\rho}_S) \\ \hat{W}_t^V(S_t^o, I_t^o) &= W_t^V(S_t^o, I_t^o) \end{aligned} \tag{51}$$

so that the Benveniste-Scheinkman theorem applies. Hence

$$\frac{\partial \hat{W}^V(S_t^o, I_t^o, p_t^o)}{\partial S_t} = \frac{\partial W^V(S_t^o, I_t^o, p_t^o)}{\partial S_t} \tag{52}$$

Using this and (50) we get that

$$\begin{aligned} \frac{\partial W^V(S_t^o, I_t^o, p_t^o)}{\partial S_t} &= \frac{dx^S(0)}{d\rho}\phi'(x_t^o) \\ &\quad + \beta(1 - p_t^o)\frac{\delta W_{t+1}^V}{\delta I_{t+1}} + \beta(1 - p_t^o)p_{t+1}^{o\prime}\frac{dx^S(0)}{d\rho}\frac{\delta W_{t+1}^V}{\delta p_{t+1}} \end{aligned} \tag{53}$$

Together with (49) this shows (26).

Derivation of dynamic externalities with vaccine, equation (38)

Consider first an increase in I_t. If a vaccine has arrived, no-one is infected, hence from equation (33) it follows that

$$\frac{\partial \overline{W}^{VV}}{\partial I_t} = -\beta p_t \delta'(I_t) W^R \tag{54}$$

This happens with probability λ. With the complementary probability, a vaccine does not arrive. In this case we apply the exact same argument as when deriving $\frac{\partial W^V}{\partial I_t}$ above, Hence, parallel with (48)we get that

$$\frac{\partial W^{VN}(S_t, I_t, p_t)}{\partial I_t} = -\phi'(x_t)\frac{m(x_t)}{m'(x_t)I_t} - \beta p_t \delta'(I_t) V^R \tag{55}$$

Here, and below, x_t refers to optimal activity levels (as x_t^o in the proofs above). It thus follows that

$$\frac{\partial W^{VV}(S_t, I_t, p_t)}{\partial I_t} = -(1-\lambda)\phi'(x_t)\frac{m(x_t)}{m'(x_t)I_t} - \beta p_t \delta'(I_t) W^R \tag{56}$$

Then consider $\frac{\partial W_t^{VV}}{\partial S_t}$. First note that $\bar{W}_t^{VV}$ is independent of S_t: If a vaccine arrives between period $t-1$ and t, no-one is infected from period t and onwards, and S_t is irrelevant. If a vaccine does not arrive, we can use the exact same procedure as in the case with no vaccine in order to find the effect on W^{VN}. Parallel with (53) we have that

$$\begin{aligned}\frac{\partial W^{VV}(S_t, I_t, p_t)}{\partial S_t} \quad &= (1-\lambda)\frac{dx^S(0)}{d\rho}\phi'(x_t) \\ &+(1-\lambda)\left\{\beta(1-p_t)\frac{\delta W_{t+1}^{VV}}{\delta I_{t+1}} + \beta(1-p_t)p_{t+1}^{o\prime}\frac{dx^S(0)}{d\rho}\frac{\delta W_{t+1}^{VV}}{\delta p_{t+1}}\right\}\end{aligned} \tag{57}$$

where $x^S(0)$ is defined as in equation (53). Inserting this expression gives (38).