Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 196 (2021) 105293

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Internal and external effects of social distancing in a pandemic ☆

Maryam Farboodi [a], Gregor Jarosch [b], Robert Shimer [c,*]

[a] *MIT, United States of America*
[b] *Princeton University, United States of America*
[c] *University of Chicago, United States of America*

Received 2 March 2021; final version received 24 May 2021; accepted 5 June 2021
Available online 25 June 2021

Abstract

We develop a quantitative framework for exploring how individuals trade off the utility benefit of social activity against the internal and external health risks that come with social interactions during a pandemic. We calibrate the model to external targets and then compare its predictions with daily data on social activity, fatalities, and the estimated effective reproduction number $R(t)$ from the COVID-19 pandemic in 2020. While the laissez-faire equilibrium is consistent with much of the decline in social activity in March in the US before any formal stay-at-home orders, optimal policy further imposes immediate and highly persistent social distancing. The expected cost of COVID-19 in the US is substantial, $12,700 in the laissez-faire equilibrium and $8,100 per person under an optimal policy. Optimal policy generates this large welfare gain by shifting the composition of costs from fatalities to persistent social distancing that largely suppresses the outbreak.

JEL classification: E1; I12; H0

Keywords: COVID-19; Social activity; Reproduction number; Economic epidemiology; Optimal control; Externalities

☆ We thank SafeGraph for making their data freely available to the research community. We are grateful for comments from Fernando Alvarez, Eric Budish, Luis Bettencourt, Charles I. Jones, Pete Klenow, Michael Kremer, Benjamin Moll, Simon Mongey, and Laura Pilossoph, as well as numerous seminar participants, the editors Tilman Börgers and Hanming Fang, and two anonymous referees.
* Corresponding author.
E-mail addresses: farboodi@mit.edu (M. Farboodi), gjarosch@princeton.edu (G. Jarosch), shimer@uchicago.edu (R. Shimer).

https://doi.org/10.1016/j.jet.2021.105293
0022-0531/

1. Introduction

A key parameter in workhorse models of disease transmission is the rate at which sick people infect susceptible people. During the COVID-19 outbreak in 2020, all policy measures against the disease were, in one way or another, aimed at reducing this rate. Our starting point is the observation that the disease is primarily transmitted through social interactions, which in turn depend on individual social activity, such as using public transportation, going to the gym, meeting with friends, and dining at restaurants. These social activities always have costs and benefits, but the emergence of COVID-19 created a significant new cost, an elevated risk of disease transmission. We argue that standard economic tools are well-suited for modeling the behavioral response to COVID-19 and for assessing the costs and benefits of restrictions on social activity, such as widespread social distancing.

In this paper, we develop a simple model where individuals value social activity but face a risk of getting sick when they have social interactions. In a laissez-faire setting, forward-looking and rational individuals understand the risk of getting sick and the cost of the disease. This creates an internal benefit to social distancing, and individuals curtail their activity accordingly. Self-interested people do not, however, internalize the health risk to others brought about by their social activity, an external benefit of social distancing.

What shape do mobility restrictions take? A key quantitative finding is that both laissez-faire and optimal policy suppress social activity but do not wipe out the disease. Instead, upon reaching the peak infection rate, the stock of infections is kept stable, slowly moving the population towards herd immunity if a cure or vaccine is never found. Policy determines the level of infections and thus the speed of adjustment towards herd immunity, with optimal policy targeting a much more suppressed level and slower adjustment.

We start the paper by using cell phone tracking data from Google to depict the evolution of social activity. Using several different metrics, we show that individuals in the US and elsewhere substantially reduced their social activity before state and local governments imposed the first "stay-at-home" or "shelter-in-place" restrictions. Conversely, as the level of infection either passed its peak or remained relatively low, social activity increased before those restrictions were relaxed. What followed was a slow recovery of social activity which remained depressed in the fall 2020.

This is consistent with our hypothesis that individuals' desire to avoid getting sick is an important determinant of their social activity. Nevertheless, we also find that mandatory social distancing reduced activity below the laissez-faire level. Our results are remarkably similar across metrics, data sets, and locations.

We then offer a simple continuous-time model which integrates both individual behavior and policy analysis into an otherwise standard epidemiological model (Kermack and McKendrick, 1927). Living individuals are either susceptible to COVID-19, currently infected by it, or recovered. Susceptible individuals may catch the disease from infected individuals through the interaction between their social activity. We capture an important feature of COVID-19, that much of the transmission appears to be either pre-symptomatic or asymptomatic, through an assumption that individuals do not know when they are infected. Otherwise individuals are forward-looking and rational. They understand how their past behavior affects the probability that they are currently infected. They also understand how other people behave, the prevalence of the disease, and the risk of dying from it. In short, they know the costs and benefits of their own social activity. Finally, we assume that a cure which eliminates the health risk associated with the disease arrives stochastically, creating an incentive to delay getting sick.

Modeling overall social activity, rather than narrowly-defined economic variables such as consumption and labor supply, allows us to capture the full cost of self-protection and lockdown policies, which have disrupted all facets of life. It also allows us to connect with new, extremely rich, and high frequency micro-data measuring human activity, such as the cell phone tracking data we use in this paper. And it allows us to connect with the actual driver of disease transmission, which is human interaction broadly defined, rather than just consumption or hours worked, as modeled and measured in standard economic applications.

We use optimal control theory to derive two ordinary differential equations (ODEs) which capture the internal cost of being susceptible and infected, along with an optimality condition which determines the level of social activity. Together with two standard differential equations from the epidemiological model, the resulting system of four ODEs fully summarizes the model and can easily be solved on a computer.

We also derive the corresponding system of ODEs which characterize the symmetric Pareto optimal allocation. Optimal policy chooses a time path for the amount of social activity, recognizing the internal and external health consequences of social activity.

We then calibrate the model to US data. The calibration targets various epidemiological findings such as the initial growth rate of the disease, the duration of infectivity, and the fatality rate, along with the value of a statistical life. We use the quantitative model for two primary purposes. First, we highlight key observations about the laissez-faire economy and optimal policy which are robust to a wide range of values for the model parameters. Second, we quantify the cost of the pandemic and show how its level and decomposition into health and economic cost varies with policy choices.

We highlight several observations. First, the laissez-faire reduction in social activity due to the internalized cost of infection is strong, yet individuals reduce activity too late, namely once the risk of infection becomes non-negligible. In contrast, optimal social distancing starts as soon as the disease emerges, discontinuously imposing social distancing. Second, optimal social distancing is persistent, remaining in place for years if no cure or vaccine is found. Laissez-faire social distancing is significantly less persistent. Third, if people and policy-makers become aware of the disease sufficiently early, the effective reproduction number $R(t)$ never falls far below 1. This is equivalent to saying that both laissez-faire and optimal policy keep the stock of infections roughly constant once it reaches its peak. Of course, this stock is far lower under the optimal policy, but we highlight that optimal policy does not attempt to wipe out the disease via social distancing.

To understand why the effective reproduction number $R(t)$ never falls far below 1, note that the model has two state variables, the susceptible and infected rates. Just like in the data, the fraction of the population susceptible to the disease is moving very slowly. As a consequence, optimal policy seeks a stable, optimal value for the stock of infections. Since the growth rate of infections is proportional to $R(t) - 1$, this means that once the effective reproduction number $R(t)$ reaches 1, it stays close to it.

Using our calibrated model, we also derive quantitative results about the expected cost of COVID-19 and the benefits of policy. We first show that both the time path of the cost and its composition differ sharply between laissez-faire and optimal policy. Under laissez-faire, almost 90 percent of the costs associated with the pandemic are pure health costs, while optimal policy substantially reduces health costs, largely suppressing infections, but comes with much higher costs from social distancing. Further, under laissez-faire most of the costs are incurred during the first year of the pandemic. In contrast, optimal policy delays infections and hence has higher costs further in the future if a cure is not found.

Overall, the expected cost of COVID-19 for the US is drastic. Under a laissez-faire path, the expected total cost amounts to 28 percent of one year's GDP. While large, this cost is still far below the cost of a benchmark where individuals do not respond to the disease, highlighting that individual behavior is a powerful force in curtailing the fallout from the disease. Under the optimal policy, the expected total cost is significantly smaller but still substantial, 18 percent of one year's GDP. The expected discounted flow cost per capita under the optimal policy starts at $13 per person per day and is still at $4 per day two years later. In light of what is thus far known about the economic and health costs brought about by the pandemic, these numbers appear plausible.

Finally, we validate the model by contrasting it with data along various dimensions. We first show that it accounts well for the overall decline and subsequent recovery in social activity, which we measure using cell phone tracking data.

Second, we contrast the aforementioned prediction that $R(t)$ stays close to 1 after peak infections with an independent measure of $R(t)$, provided by `rt.live`. We show that the prediction bears out remarkably well in the data throughout the summer of 2020 across US states. Following peak infection, the stock of infected individuals has been slowly rolled over in most locations, keeping $R(t)$ close to 1.

Third, we consider the evolution of fatalities in the US. We show that total US fatalities have stayed substantially below what would have occurred under laissez-faire. At the same time, they have been vastly higher than what would have happened if the US had adopted the optimal policy—largely suppressing the outbreak via social distancing—starting on March 13, 2020.

We note that we largely restrict our empirical work and the comparison of model and data to the period of spring through fall of 2020. The model has time-invariant parameters and is as such not directly capable of picking up the second wave of infections in the summer of 2020 and the much larger third wave that started in the fall of 2020. We suspect the former has much do with the relaxation of temporary restrictions on activity while the latter has much to do with seasonality and the appearance of new mutations. In Appendix D, we demonstrate how a shock to the model's parameters can capture seasonality and mutations, and how this can give rise to a later wave as witnessed in the data.

Overall, the model's predictions align well with the data along various dimensions in the early phase of the pandemic. It is therefore a natural laboratory for quantitatively evaluating both policies that we have observed, like widespread social distancing, and policies that have been proposed but not yet implemented, including massive testing, test-trace-and-quarantine, and antibody testing. We leave this for future research.

The paper proceeds as follows: Section 2 relates our paper to the literature; Section 3 documents the evolution of social activity; Section 4 lays out the model and defines laissez-faire equilibrium as well as the optimum; Section 5 calibrates the model and presents our main qualitative and quantitative results; Section 6 compares our model to data; Section 7 discusses the role of several key assumptions and modeling choices; and Section 8 concludes.

2. Related work

Our basic approach builds on the susceptible-infected-recovered (SIR) model (Kermack and McKendrick, 1927). There exists a small body of pre-COVID-19 work which aims at integrating optimal control of economic or social activity with a model of disease dynamics (e.g. Philipson and Posner, 1993, 1995; Philipson, 2000; Kremer, 1996; Fenichel, 2013; Greenwood et al., 2019; Toxvaerd, 2019). Prior to 2020, quantitative evaluations of epidemiological models with standard

economic features have, to the best of our knowledge, been limited to HIV/AIDS, a very different disease than COVID-19 with its short infectious period and high recovery rate.

There is now a rapidly growing body of work which applies a mix of the basic epidemiological model and dynamic optimization tools from economics to think about various aspects of COVID-19.[1] Relative to these papers, we focus on writing down the simplest possible model that can capture the costs and benefits of widespread social distancing in a manner that is quantitative and in line with available evidence. Our quantitative objective means that we focus throughout on the fact that the disease can be transmitted through many types of social interactions, only some of which are captured by consumption or labor supply. Fortuitously, this allows us to map our model into newly available high-frequency data from Google and SafeGraph on social activity. This type of data only became available to researchers after the onset of COVID-19 and this paper was among the first to utilize it. Our focus on social activity also allows us to connect our model with the policy that was actually employed, social distancing.

Arguably what is most distinctive about our approach is that we offer a host of evidence on various aspects of both individual behavior and disease characteristics which we use to quantitatively evaluate the performance of the model. We discuss the evidence not only on social distancing but also on mortality and on the effective reproduction number $R(t)$. We also take advantage of the model's tractability to provide intuition for why the laissez-faire and optimal policies give rise to the outcomes that we observe.

Because of the significant overlap in the models, our paper leads to broadly similar qualitative conclusions as other papers in the rapidly growing literature on COVID-19. Individual optimality alone leads to social distancing, yet not enough. Optimal policy imposes immediate social distancing but allows the disease to spread until we reach herd immunity. Relative to these papers, we develop a unified and tractable approach to study both equilibrium and optimum, offer new evidence on equilibrium behavior in the early stages of the COVID-19 outbreak, and give a systematic assessment of the model's quantitative performance using a host of data on various aspects of the epidemic. Our quantitative prediction that optimal policy is persistent but not overly restrictive appears to be novel in this literature. In particular, we believe we are the first to connect these properties of optimal policy to the effective reproduction number $R(t)$, a fundamental concept in epidemiology, and show that it stabilizes close to 1.

At a more fine-grained level, our paper differs from most others in its information structure. We assume that newly-infected individuals do not know their infection status and so spread the disease unknowingly. This means that policy cannot discriminate between susceptible and infected people, which in turn means that social distancing policies must be wide-spread, rather than targeted on infected individuals. The aforementioned papers typically make different assumptions about individuals' information sets. On a related point, other papers focus on a broader set of policy tools. We deliberately focus on widespread social distancing because it was ubiquitous in the early stages of the COVID-19 outbreak and is likely to be proposed again for future outbreaks. Moreover, targeted policies like test-trace-and-quarantine still appear unavailable or dysfunctional many months after the onset of the outbreak.

[1] An incomplete list includes Acemoglu et al. (2020), Alvarez et al. (2020), Atkeson (2020), Abel and Panageas (2020), Barro et al. (2020), Bethune and Korinek (2020), Bognanni et al. (2020), Brotherhood et al. (2020), Budish (2020), Dewatripont et al. (2020), Eichenbaum et al. (2020), Garibaldi et al. (2020), Glover et al., Hall et al. (2020), Jones et al. (2020), Kaplan et al., Keppo et al., Krueger et al. (2020), Piguillem and Shi (2020), Rowthorn and Toxvaerd (2020), Toxvaerd (2020), and Rachel (2020).

The paper is also related to an older literature on social externalities, including Diamond and Maskin (1979) and Kremer and Morcom (1998). In particular, Diamond and Maskin (1979) introduce the distinction between a *quadratic* and a *linear matching technology*. With quadratic matching, additional social activity by others raises the likelihood of social contact and thus disease transmission for all individuals. For example, with more individuals in parks, restaurants, and public transit, any given trip to a park/restaurant/subway is more likely to lead to disease. Such a matching function has a search externality that, traditionally, is viewed as positive (Diamond, 1982), but that turns negative in the context of disease. It stands in contrast to a linear search technology where an individual's total social contacts merely depend on her own social activity and not on those of others. We believe that such a technology applies to cases where social contacts are explicitly sought out. We therefore argue that the quadratic technology is appropriate to model the dynamics of COVID-19, while a linear technology might be the right tool to model an epidemic like HIV.

3. Social activity

In this section, we use newly available data on social activity to document substantial social distancing across the US and elsewhere prior to any mandatory restrictions on social activity during the initial outbreak of COVID-19 in the spring of 2020. We also discuss the slow recovery in social activity that occurred throughout the summer and the second wave of social distancing that occurred in the winter 2020–21.

We use Google's publicly available Community Mobility Reports and provide similar evidence in micro-data from SafeGraph in the Appendix. Both data sources are based on cell phone tracking data. Importantly, since we will model the choice of social activity directly, our setup connects directly with this type of data.

3.1. Data on social activity

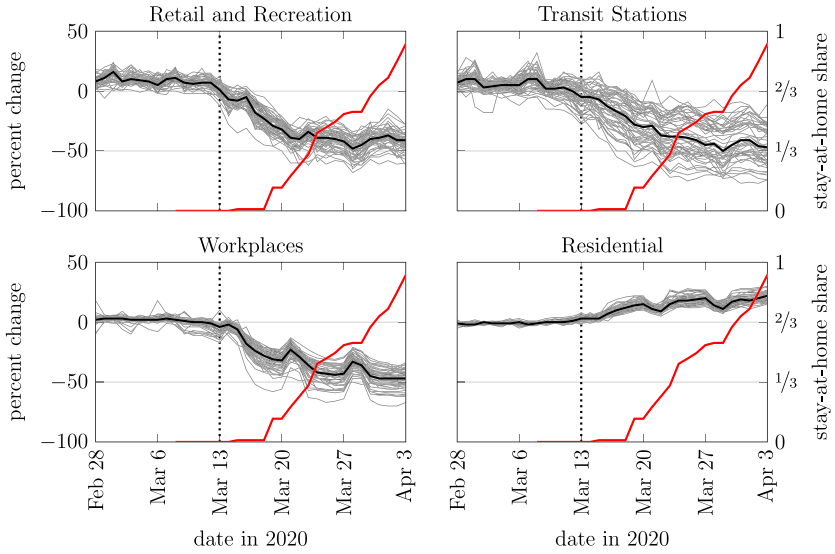
Google's COVID-19 Community Mobility Reports (https://www.google.com/covid19/mobility/) use Google devices' location history to track the number of visits and length of stay at places recognized as locations on Google maps. The activity metric is the percent change in activity relative to a baseline period of January 3 to February 6, 2020. We restrict attention to four categories of activity, each constructed by Google: "Retail and Recreation," "Transit stations," "Workplace," and "Residential."[2]

3.2. The crunch in March 2020

Fig. 1 shows the evolution of Google's measure of social activity across the US in March 2020. The thin black lines in each figure plot Google's activity metric across each of the fifty US states and the District of Columbia. The thick black line shows the cross-sectional median change in activity at any given date.

On March 13, 2020, indicated by a dotted line in Fig. 1, the US government issued a proclamation declaring COVID-19 to be a national emergency. While this proclamation had little formal

[2] We drop two categories, "Grocery and Pharmacy" and "Parks". Both categories have complicated relationships with social distancing. For example, visits to grocery stores surged as people stocked their cupboards in the middle of March, while visits to parks depended on whether they were open in a particular location.



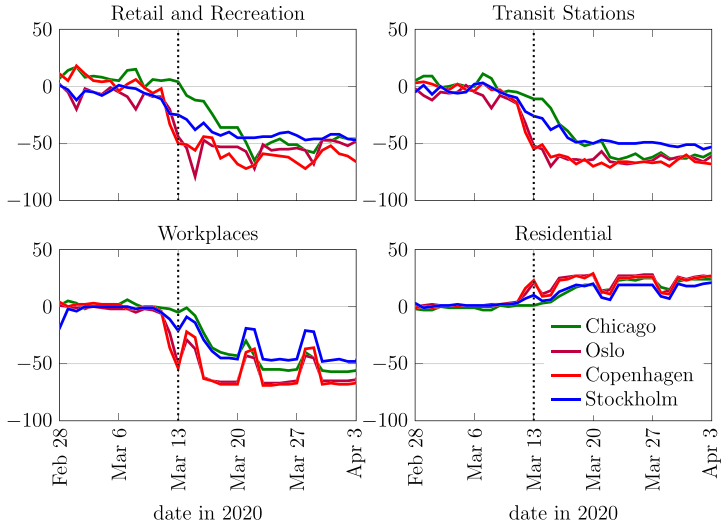
Notes: Each thin line corresponds to the Google Community Mobility Report for one of the 51 US states and the District of Columbia; see the text for a description of this measure. We retrieved this data from https://www.google.com/COVID19/mobility/ on February 19, 2021. The thick black line marks the median state for that day. These are shown on the left hand scale. The red line shows the fraction of the US population subject to stay-at-home or shelter-in-place orders and is shown on the right hand scale. The population subject to these orders is based on authors' own calculations using https://www.nytimes.com/interactive/2020/us/coronavirus-stay-at-home-order.html. The vertical dotted line indicates March 13, 2020, the date of the declaration of national emergency.

Fig. 1. Declining Activity, Early and Everywhere I.

content, it reflected a growing awareness of the dangers posed by SARS-CoV-2, the virus that causes COVID-19 in the US. In the following week, retail and recreation activity fell a median of 33 percentage points, transit station activity by 25 percentage points, and workplace activity by 28 percentage points. Over the same week, residential activity increased by 12 percent. All this occurred before any significant fraction of the US population was subject to stay-at-home orders, as indicated by the red lines in Fig. 1. In the next week, as those orders spread across half the country, the first three categories fell by a further 10, 15, and 11 percentage points, respectively, while residential increased by another 5 percentage points. The mobility index for workplaces fell by another 4 percentage points during the week after that, while the other three indexes stabilized.

Google's COVID-19 Community Mobility Reports offer information on international locations as well. We use that to compare the evolution of social activity in Chicago with three Scandinavian capitals, Oslo, Copenhagen, and Stockholm. These three cities are arguably similar in a lot of ways, but policies differed sharply. While Norway and Denmark locked down on March 12 and 13, respectively, Sweden is widely considered an outlier in that it instituted few restrictions on social activity in the spring of 2020.

Fig. 2 shows that social activity started falling in all four cities around March 13. The decline was sharpest in Oslo and Copenhagen, which were subject to formal stay-at-home orders, and more gradual in Stockholm and Chicago, which were not. Then on March 21, Illinois is-



Notes: Figure constructed equivalently to Fig. 1 but only these four cities. Activity metric is number of visits and length of stay at places recognized as locations on Google maps, relative to a baseline period (January 3–February 6, 2020). Categories constructed by data provider. Retrieved from https://www.google.com/COVID19/mobility/ on February 19, 2021. The vertical dotted line indicates March 13, 2020, the date of the U.S. declaration of national emergency.

Fig. 2. Declining Activity, Early and Everywhere II.

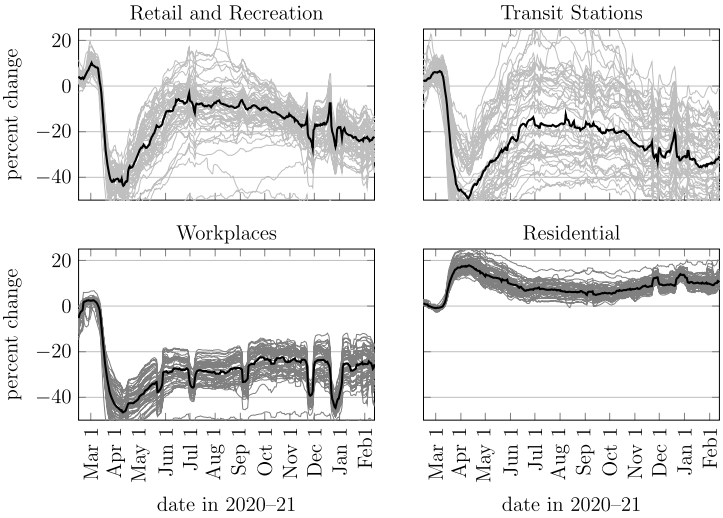
sued a shelter-in-place order. After that, the levels of retail, transit, and workplace activity in Chicago, Oslo, and Copenhagen were remarkably similar to each other and noticeably below Stockholm. Still, the significant decline in activity in Stockholm is consistent with our hypothesis that voluntary social distancing can substantially reduce activity even without formal stay-at-home orders.

3.3. Slow recovery and second wave of distancing

Fig. 3 plots the Google Data across US states for the full range of available data, until February 13, 2021. Social activity stopped declining and started to recover after Easter Sunday, April 12. As stay-at-home orders were eased in some states in late April and early May, there was no break in that trend. Instead, we saw a gradual ramp-up in activity that, despite a slowdown, continued through the summer. The recovery was, however, extremely slow and remained incomplete across all the different mobility metrics we study.

Starting in the fall, the trend reversed again and all mobility metrics have been continuously decreasing, recently stabilizing at low levels.

Our model-based analysis is primarily designed to speak to the period of March to October 2020, and the basic setup cannot capture the second wave of COVID-19 that occurred in the winter of 2020–21. However, in Appendix D we show how a simple extension of the model can naturally capture the emergence of a second wave of infections and increased social distancing, as we saw through the fall and winter of 2020–21.



Notes: Figure identical to Fig. 1 but with a wider date range and a 7 day centered moving average.

Fig. 3. Crunch, Recovery, and Second Wave of Social Distancing across Categories – Google Data.

Overall, this evidence indicates that there was and is widespread social distancing in the US, but that the majority of it was unrelated to mandatory stay-at-home orders.[3] It began before restrictions were put in place and it persisted after they were lifted and reflected the desire of individuals to reduce their exposure to the disease as it evolved over time.

We complement this empirical picture with additional evidence from SafeGraph micro-data in Appendix A. The bottom line is that we have direct and readily available measures of social activity. This is both the key endogenous variable in our model and the key driver of the pandemic. Since we model social activity rather than, say, consumption, our model can directly connect with this high-frequency data. In Section 6 we will confront the quantitative properties of our model with this evidence and argue that it offers a close account of the decline in social activity in March 2020 and its slow subsequent recovery.

4. Model

4.1. Setup

The basic epidemiological framework is a continuous time SIR model with a possibility of death, i.e. a SIRD model. We embed this into an economic model where individuals get utility from their social activity. Those who are susceptible to the disease can become infected through social interactions with infected individuals. The infected stochastically recover or die. Recovery confers immunity to the disease. Individuals do not know if they are susceptible or infected, but they can tell when they have recovered from the disease. Individuals are otherwise homogeneous.

[3] Brotherhood et al. (2021) use cross-sectional differences in terms of the cost associated with social distancing to show that people in crowded settings (slums) likewise distanced less, reflecting the same cost-benefit considerations we appeal to.

Preferences and health states Individuals discount the future at rate ρ and live forever unless they die from the disease. At any time $t \geq 0$, each individual is in one of the four states, susceptible (s), infected (i), recovered (r), or deceased (d). Let $N_\theta(t)$, $\theta \in \{s, i, r, d\}$, denote the measure of individuals in each state, with $\sum_\theta N_\theta(t) = 1$. To make the environment nontrivial, assume that $N_i(0) > 0$, so there is a seed of infection, and $N_s(0) > 0$, so there are people who can get infected.

Social activity and social interactions Each living individual gets utility $u(a(t))$ from a level of *social activity* $a(t) \geq 0$. We assume u is single-peaked and its maximum is attained at a finite value a^*. Without loss of generality, normalize $a^* = 1$ and $u(a^*) = 0$. The first normalization keeps the notation the same as the basic SIR model in the absence of a behavioral response to a disease outbreak. The latter allows us to use u as a measure of the utility cost of social distancing, i.e. of setting $a < a^*$.

We assume throughout that all individuals of a given type choose the same level of social activity, although when we study equilibrium, we consider the deviation of a single individual to a different level of activity. Let $A_\theta(t)$ denote the amount of social activity by each individual of type $\theta \in \{s, i, r\}$.

Disease transmission depends on *social interactions*, defined as the interaction between individuals' social activity. More precisely, the rate at which some individual of type θ has a social interaction with some other individual of type θ' is proportional to $A_\theta(t)N_\theta(t)A_{\theta'}(t)N_{\theta'}(t)$, the product of the total amount of social activity of the two types. In particular, the rate at which susceptible individuals get infected is $\beta A_s(t)N_s(t)A_i(t)N_i(t)$, where $\beta > 0$ captures the ease of transmitting the disease. Formally, this is a quadratic matching technology with random search.[4] We stress that there are not only social interactions between the susceptible and the infected, but also between all other groups. For instance, the recovered can optimally choose $A_r(t)$ without fear of infection; and susceptible individuals meet one another without consequences. Our model allows these interactions to happen, but implies that they do not affect the number of interactions where the disease gets transmitted.

We intend for social activity to include not only standard economic concepts like consumption and employment, but also all the other things that become risky when a novel and contagious respiratory disease is prevalent. Does a person use public transportation, go to the gym, or enjoy the amenities of urban life? Likewise, we intend to capture the margins of adjustment that policy makers have been primarily relying on in the battle against the disease. Do we close parks and beaches, do we ban gatherings among friends and family, indeed, do we lock people up at home?

None of these margins are captured by models of consumption and employment only, but each of them affects individual's utility and disease transmission, and each of them is an important part of policy. Indeed, in the calibrated model of Eichenbaum et al. (2020), eliminating all consumption and work would only reduce the transmission rate of COVID-19 by a third, surely

[4] An alternative would be a linear technology, in which case type θ individuals interact with type θ' individuals at rate $\frac{A_\theta(t)N_\theta(t)A_{\theta'}(t)N_{\theta'}(t)}{\sum_{\theta''} A_{\theta''}(t)N_{\theta''}(t)}$. This formulation makes sense if type θ individuals desire to interact with somebody at rate $A_\theta(t)$. With a linear technology, changing social activity by other people changes the distribution of social interactions without changing the level. Such an alternative might be the appropriate modeling choice for other epidemics such as HIV (Kremer and Morcom, 1998). See also Fenichel (2013). However, for COVID-19 it seems unlikely that additional social activity by non-infected individuals would reduce the infection risk of the susceptible, all else equal. We highlight, however, that this assumption is crucial for certain outcomes. For example, with a linear technology, it is optimal to boost the social activity of the recovered in order to moderate meetings between susceptible and infected people.

understating the extent to which individuals and policy can reduce their risk exposure. Our concept of social interactions is intended to capture all the ways that a contagious respiratory disease might spread, as well as the cost of curtailing that spread.

The assumptions that preferences u depend on social *activity* while disease transmission depends on social *interactions* are central to our view of social distancing. The former captures the idea that individuals value social activity, such as using public transportation, going to the gym, or dining at a restaurant.[5] In each of these examples, one can imagine reasons why some people's marginal utility of social activity is increasing or decreasing in the social activity of others. We assume it is unaffected. On the other hand, social interactions are critical for disease transmission, since if someone is the only passenger on the bus, is alone at the gym, and is the only diner in a restaurant, the risk of disease transmission is minimal.

Our disease transmission function captures the idea that if a particular group chooses little social activity, so $A_\theta(t)$ is small, then one is unlikely to make social contact with them. As a consequence, social activity displays a negative externality: it increases other people's social interactions, putting them at a higher risk of infection.

Recovery and death Infected individuals recover at rate $(1-\pi)\gamma$ and die at rate $\pi\gamma$, where $\pi \in [0,1]$ is the infection fatality rate,[6] the fraction of infected individuals who eventually die from the disease, and $1/\gamma$ is the expected duration of infection. We assume that at the moment an individual dies, he pays a utility cost v, the value of a statistical life. Equivalently, all individuals pay a utility cost $\kappa \equiv \pi v$ when they exit the infected state.

We assume that recovering from the disease confers lifetime immunity, so a recovered individual no longer transmits the disease and can no longer become sick. We recognize that the empirical evidence on the latter point is still unclear.

The end of the disease We assume a cure for the disease is found at rate δ. For simplicity we assume that a cure, once found, is perfect and immediately wipes out all costs of the disease.

Information Individuals do not know whether they are susceptible or infected, but they know when they have recovered from the disease. We capture this through the measurability restriction that $A_s(t) = A_i(t)$ and use $A(t)$ to denote this common level of social activity. In Appendix C.2, we show how to analyze the case where only some recovered individuals are aware that they have recovered, while others are subject to the same measurability restriction. Our qualitative analysis and quantitative results are similar in this case.

Reproduction number Our model has notions of two important epidemiological concepts, the basic and effective reproduction numbers. The *basic reproduction number* R_0 is defined as the expected number of susceptible people infected by a sick person in a world where almost everyone is susceptible and unaware of the disease. This was relevant, for example, when the disease

[5] For an environment where interactions are critical, see Diamond (1982).

[6] We assume that the infection fatality rate π is independent of the number of infected people, thus shutting down the possibility that the disease overwhelms the healthcare system. This is not meant to dispute that "healthcare capacity" (or "ICU") constraints are potentially an important feature of the COVID-19 outbreak. Instead, we do this to highlight that key properties—e.g. that optimal policy "flattens the curve"—arise even without such a constraint. We discuss this point further in Section 7.

first emerged. This is simply $R_0 = \beta/\gamma$. Throughout, we impose $\beta > \gamma$, since otherwise the disease never breaks out.

The *effective reproduction number* is the expected number of others infected by a sick person, given the current level of social activity $A(t)$ and the current fraction of susceptible people $N_s(t)$. This is $R(t) = R_0 A(t)^2 N_s(t)$. Social distancing drives the ratio of the effective reproduction number to the fraction of susceptible people, $R(t)/N_s(t)$, below the basic reproduction number R_0. As is well known, the number of infections grows (shrinks) if $R(t)$ is bigger (smaller) than 1.

4.2. Laissez-faire equilibrium

In this section, we consider the problem of an individual choosing his own social activity, taking the social activity of others and the number of others with each health status as given. We then impose the equilibrium restriction that individual outcomes must coincide with the aggregate.

An individual has rational beliefs about his own probabilities of being susceptible, infected, and recovered, which we denote by $n_s(t)$, $n_i(t)$, and $n_r(t)$, respectively. The individual knows when he is recovered but cannot distinguish between the susceptible and infected states. He thus chooses two time paths for social activity, $a(t)$ when he is susceptible or infected and $a_r(t)$ when he is recovered. Finally, the individual discounts future utility at rate ρ and recognizes that the problem ends at rate δ when a cure is found. Putting this together, the individual solves

$$\max_{\{a(t),a_r(t)\}} \int_0^\infty e^{-(\rho+\delta)t}\big((n_s(t)+n_i(t))u(a(t)) + n_r(t)u(a_r(t)) - \gamma n_i(t)\kappa\big)dt$$

subject to

$$n_s'(t) = -\beta a(t) n_s(t) A(t) N_i(t), \tag{1}$$

$$n_i'(t) = \beta a(t) n_s(t) A(t) N_i(t) - \gamma n_i(t), \tag{2}$$

$$n_r'(t) = (1-\pi)\gamma n_i(t),$$

$$n_d'(t) = \pi\gamma n_i(t),$$

with $n_s(0) = N_s(0)$, $n_i(0) = N_i(0)$, $n_r(0) = N_r(0)$, and $n_d(0) = N_d(0)$ given. Note that through the choice of $a(t)$, an individual affects his own transitions across different health statuses. An individual transitions from susceptible to infected at a rate proportional to the product of his own activity level $a(t)$, his own probability of being susceptible $n_s(t)$, and the total activity by infected people $A(t)N_i(t)$. He exits the infected state at rate γ, either recovering or dying, paying an expected cost κ. The individual assumes his own behavior does not affect the time path of $A(t)N_i(t)$.

To solve the individual's problem, first note that $a_r(t)$ affects the objective but none of the constraints. It is thus optimal to set $a_r(t) = a^* = 1$ and thus $u(a_r(t)) = 0$ for all t. Dropping this control variable and the unnecessary third and fourth constraints, write the current value Hamiltonian as

$$\begin{aligned} H(n_s(t), n_i(t), a(t), \lambda_s(t), \lambda_i(t)) &= (n_s(t)+n_i(t))u(a(t)) - \gamma n_i(t)\kappa \\ &\quad - \lambda_s(t)\beta a(t) n_s(t) A(t) N_i(t) + \lambda_i(t)\big(\beta a(t) n_s(t) A(t) N_i(t) - \gamma n_i(t)\big), \end{aligned}$$

where $\lambda_s(t)$ and $\lambda_i(t)$ are the co-states associated with the two remaining constraints.

There are three necessary first order conditions for optimal control. First, the derivative of the Hamiltonian with respect to the control variable $a(t)$ is zero:

$$(n_s(t)+n_i(t))u'(a(t)) = (\lambda_s(t)-\lambda_i(t))\beta n_s(t)A(t)N_i(t). \tag{3}$$

This states that the private return from social activity balances the risk of getting infected. Additionally, the derivatives with respect to the state variables $n_s(t)$ and $n_i(t)$ are equal to minus the time derivative of the costate, with a correction for discounting:

$$(\rho+\delta)\lambda_s(t) - \lambda_s'(t) = u(a(t)) + (\lambda_i(t)-\lambda_s(t))\beta a(t)A(t)N_i(t), \tag{4}$$

$$(\rho+\delta)\lambda_i(t) - \lambda_i'(t) = u(a(t)) - \gamma\big(\kappa+\lambda_i(t)\big). \tag{5}$$

There are two more necessary conditions for optimality, the transversality conditions

$$\lim_{t\to\infty} e^{-(\rho+\delta)t}\lambda_s(t)n_s(t) = \lim_{t\to\infty} e^{-(\rho+\delta)t}\lambda_i(t)n_i(t) = 0. \tag{6}$$

Equilibrium requires that individual and aggregate behaviors are consistent at every point in time, $n_s(t) = N_s(t)$, $n_i(t) = N_i(t)$, and $a(t) = A(t)$ for all $t > 0$. Imposing those restrictions on equations (1)–(6) gives us

$$N_s'(t) = -\beta A(t)^2 N_s(t)N_i(t), \tag{7}$$

$$N_i'(t) = \beta A(t)^2 N_s(t)N_i(t) - \gamma N_i(t), \tag{8}$$

$$(N_s(t)+N_i(t))u'(A(t)) = (\lambda_s(t)-\lambda_i(t))\beta A(t)N_s(t)N_i(t), \tag{9}$$

$$(\rho+\delta)\lambda_s(t) - \lambda_s'(t) = u(A(t)) + (\lambda_i(t)-\lambda_s(t))\beta A(t)^2 N_i(t), \tag{10}$$

$$(\rho+\delta)\lambda_i(t) - \lambda_i'(t) = u(A(t)) - \gamma\big(\kappa+\lambda_i(t)\big), \tag{11}$$

$$\lim_{t\to\infty} e^{-(\rho+\delta)t}\lambda_s(t)N_s(t) = \lim_{t\to\infty} e^{-(\rho+\delta)t}\lambda_i(t)N_i(t) = 0. \tag{12}$$

These equations and initial conditions for $N_s(0)$ and $N_i(0)$ fully summarize the model.

Note from equation (8) that the growth rate of infections, $N_i'(t)/N_i(t)$, is $\gamma(R(t)-1)$, where $R(t) \equiv R_0 A(t)^2 N_s(t)$ is the effective reproduction number. Thus infections grow if and only if the effective reproduction number exceeds 1. This is why a low effective reproduction number is critical for reducing the prevalence of the disease.

We solve this model through a backward shooting algorithm. We provide a detailed description of our numerical solution algorithm in Appendix B.

4.3. Social optimum

We now solve the problem faced by a benevolent social planner who dictates the time path of social activity $A(t)$ and $A_r(t)$. The planner, like the individual, recognizes that a reduction in contacts lowers utility directly, but she also recognizes the externalities associated with infections. The planner solves

$$\max_{\{A(t),A_r(t)\}} \int_0^\infty e^{-(\rho+\delta)t}\big((N_s(t)+N_i(t))u(A(t)) + N_r(t)u(A_r(t)) - \gamma N_i(t)\kappa\big)dt \tag{13}$$

subject to equations (7) and (8) and $N_r'(t) = (1-\pi)\gamma N_i(t)$. As in equilibrium, it is optimal to set $A_r(t) = a^* = 1$ and so $u(A_r(t)) = 0$ for all t, since this does not affect the evolution of the state variables. To characterize the path of $A(t)$, we again write down a Hamiltonian,

$$H(N_s(t), N_i(t), A(t), \mu_s(t), \mu_i(t)) = (N_s(t) + N_i(t))u(A(t)) - \gamma N_i(t)\kappa \\ - \mu_s(t)\beta A(t)^2 N_s(t)N_i(t) + \mu_i(t)\big(\beta A(t)^2 N_s(t)N_i(t) - \gamma N_i(t)\big).$$

The necessary first order condition with respect to the control A is

$$(N_s(t) + N_i(t))u'(A(t)) = 2(\mu_s(t) - \mu_i(t))\beta A(t)N_i(t)N_s(t), \tag{14}$$

while the necessary costate equations are

$$(\rho + \delta)\mu_s(t) - \mu_s'(t) = u(A(t)) + (\mu_i(t) - \mu_s(t))\beta A(t)^2 N_i(t), \tag{15}$$

$$(\rho + \delta)\mu_i(t) - \mu_i'(t) = u(A(t)) + (\mu_i(t) - \mu_s(t))\beta A(t)^2 N_s(t) - \gamma\big(\kappa + \mu_i(t)\big). \tag{16}$$

Finally, the planner also has necessary transversality conditions

$$\lim_{t\to\infty} e^{-(\rho+\delta)t}\mu_s(t)N_s(t) = \lim_{t\to\infty} e^{-(\rho+\delta)t}\mu_i(t)N_i(t) = 0. \tag{17}$$

Together with the state equations (7) and (8), we again have a system of four differential equations and one static equation. We again solve it using a backward shooting algorithm, as described in Appendix B.

There are two key differences between the equilibrium equations (9)–(11) and the corresponding optimal equations (14)–(16). First, the planner recognizes that raising everyone's meeting rate $A(t)$ increases meetings at rate proportional to $2A(t)$, while in equilibrium raising an individual's meeting rate $a(t)$ increases meetings at rate proportional to $A(t)$. This creates an extra factor of 2 in equation (14) compared to equation (9). Second, the planner recognizes that sick people get other people sick, while in equilibrium individuals do not care about this outcome. This shows up as an extra term on the right hand side of equation (16) compared to equation (11).

4.4. Perfect and imperfect altruism

In the laissez-faire setting, people only care about their own health. In reality, diseases are often transmitted to family and friends, and so it seems plausible that many people would like to reduce the risk of transmitting the disease, not just the chance of getting it. We show how our basic setup can be extended to a model of imperfect altruism, indexed by an altruism parameter $\alpha \in [0, 1]$. At one extreme, $\alpha = 0$, individuals would not socially distance if they knew they were sick. This is the laissez-faire equilibrium. At the other extreme, $\alpha = 1$, individuals fully internalize the cost of making others sick.

We modify the individual objective function to incorporate that individuals are concerned about making others sick:

$$\max_{\{a(t),a_r(t)\}} \int_0^\infty e^{-(\rho+\delta)t}\big((n_s(t) + n_i(t))u(a(t)) + n_r(t)u(a_r(t)) \\ - \gamma n_i(t)\kappa - \alpha\beta a(t)n_i(t)A(t)N_s(t)(\lambda_s(t) - \lambda_i(t))\big)dt.$$

The new piece is the last term. When an individual infects a susceptible person, at rate $\beta a(t)n_i(t)A(t)N_s(t)$, she suffers a utility loss equal to a fraction α of the difference $\lambda_s(t) - \lambda_i(t)$, where again $\lambda_\theta(t)$ is the costate variable on $n_\theta(t)$, $\theta \in \{s, i\}$. This difference represents the private cost of getting sick.

Table 1
Calibration to US data.

Parameters			
Parameter Description	Parameter	Value	Target
Conditional transmission prob.	β	$0.3+\gamma$	Initial doubling time
Rate that infectiousness ends	γ	$1/7$	Duration until symptomatic
Infection fatality rate (IFR)	π	0.0062	Hall et al. (2020)
Value of a statistical life (VSL)	v	31,755	Hall et al. (2020)
Arrival rate of cure	δ	$0.67/365$	Exp. time until vaccine/cure
Discounting	ρ	$0.05/365$	Annual discount rate
Fraction initially affected	$N_i(0)$	0.0000527	Deaths before March 13, 2020
Other			
Basic reproduction number	R_0	3.1	Implied by γ and β
Expected cost of infection	κ	197	Implied by π and v
Fraction initially susceptible	$N_s(0)$	0.9999223	no social distancing before $t=0$

Note: We calibrate the model at a daily frequency.

With this modification to the objective function, we can again write down the Hamiltonian and find the optimality and costate equations as well as the transversality conditions. Imposing $n_s(t)=N_s(t)$, $n_i(t)=N_i(t)$, and $a(t)=A(t)$ for all $t>0$ gives

$$(N_s(t)+N_i(t))u'(A(t)) = (1+\alpha)\beta A(t)(\lambda_s(t)-\lambda_i(t))N_s(t)N_i(t), \tag{18}$$

$$(\rho+\delta)\lambda_s(t)-\lambda_s'(t) = u(A(t))+\beta A(t)^2 N_i(t)(\lambda_i(t)-\lambda_s(t)), \tag{19}$$

$$(\rho+\delta)\lambda_i(t)-\lambda_i'(t) = u(A(t))+\alpha\beta A(t)^2 N_s(t)(\lambda_i(t)-\lambda_s(t))-\gamma(\kappa+\lambda_i(t)), \tag{20}$$

$$\lim_{t\to\infty} e^{-(\rho+\delta)t}\lambda_s(t)N_s(t) = \lim_{t\to\infty} e^{-(\rho+\delta)t}\lambda_i(t)N_i(t) = 0. \tag{21}$$

If $\alpha=0$, this returns the equilibrium equations (9)–(12). If $\alpha=1$, this returns the equations (14)–(17) describing the planner's solution in the setting where there is no altruism. Intermediate values of α capture some degree of imperfect altruism in the laissez-faire setting. We note that this logic does not imply that the equilibrium is efficient under perfect altruism, since the planning problem itself changes when individuals are altruistic.

5. Quantitative analysis

5.1. Calibration

We calibrate the model to US data on the COVID-19 outbreak after March 13, 2020, the date of the proclamation declaring COVID-19 to be a national emergency in the US. Table 1 summarizes our calibration.

We normalize the time unit to be a day. We then set $\rho=0.05/365$ to capture a 5 percent annual discount rate. In addition, we set $\delta=0.67/365$, which implies an expected time until cure of 1.5 years. We highlight that this jointly implies that individuals heavily discount future payoffs.

Next, we set $\gamma=1/7$, which implies that the average length of sickness is 1 week. We recognize that the disease lasts longer on average, particularly for people who develop serious symptoms. We choose this large value for γ to capture the length of the period when people

are infectious and either asymptomatic, pre-symptomatic, or have mild symptoms. Lauer et al. (2020) report a median incubation period for COVID-19 of five days and that 98 percent of people who develop symptoms after an exposure do so within 11.5 days. Appendix C.3 discusses an extension to the model where the infectious period may end in hospitalization, which in turn leads to either recovery or death.

We calibrate β to capture an initial 30 percent daily growth rate of the stock of infected, consistent with a doubling time of approximately three days. Specifically, if $N_s(t) = A(t) = 1$, equation (8) implies $N_i'(t)/N_i(t) = \beta - \gamma$, and so $\beta = 0.3 + \gamma = 0.443$ implies an initial daily growth rate of 0.3. These values in turn imply a basic reproduction number of $R_0 = \beta/\gamma = 3.1$.[7] Some recent estimates suggest an even higher value for R_0 (Sanche et al., 2020) and several authors work with a lower γ (e.g. Alvarez et al., 2020). We therefore offer a robustness exercise below where we use a lower value of γ and a correspondingly higher value of R_0, while still hitting the initial 30 percent daily growth rate.

We work with the following period utility function,

$$u(a) = \log a - a + 1. \tag{22}$$

We think of the first part as the gross returns from social activity, in particular consumption, and of the second part as the cost associated with it. This function has an interior maximum of 0 achieved at $a = a^* = 1$.

To measure the cost of disease κ, we build largely on Hall et al. (2020). Recall that κ is the product of the infection fatality rate (IFR) π and the value of a statistical life (VSL) v. The first term is relatively uncontroversial. Hall et al. (2020) work with two baseline IFRs of $\pi = 0.0081$ and 0.0044, and so we choose the intermediate value $\pi = 0.0062$.[8]

The VSL is a more controversial number, even though it is familiar in the economic analysis of environment regulation.[9] We follow Hall et al. (2020), who calculate a value of \$270,000 for each remaining year of life. They also offer a value of 14.5 years for the average remaining life expectancy of COVID-19 victims. This gives $v = \$3{,}915{,}000$. Roughly speaking, this means that the typical victim of COVID-19 would pay \$3,915 to avoid a 0.1 percent probability of death. With a discount rate of $\hat{\rho}$, this is equivalent to paying a constant stream of $\hat{\rho} \times \$3{,}915$ per day to avoid this death risk.[10] Compare this to current US per capita consumption of approximately \$45,000 per year (Hall et al. (2020)), and we reach the conclusion that people would permanently give up $\frac{\$3{,}915 \times 365}{\$45{,}000}\hat{\rho} = 31.755\hat{\rho}$ percent of their consumption to avoid a 0.1 percent death risk.

[7] There is still considerable disagreement on the uncurtailed doubling time or, alternatively, R_0 of COVID-19. Johns Hopkins University's Center for Systems Science and Engineering reports a doubling time of 2–5 days in the US in the early stages of the epidemic (Dong et al., 2020). Pan et al. (2020) report an effective reproduction number slightly above 3.0 in January 2020 in Wuhan.

[8] Robert Redfield, the director of the CDC, said during a call with reporters on June 25, 2020 that "Our best estimate right now is that for every case that was reported, there actually were 10 other infections" (https://n.pr/3dC55bu, retrieved June 26, 2020). With a case count of 2,697,326 and 129,007 deaths at that time, this would give $\pi = 0.0043$, close to the lower value in Hall et al. (2020) (https://coronavirus.1point3acres.com/en, retrieved July 1, 2020).

[9] See Greenstone and Nigam (2020) for a useful discussion of this value and the use of VSL in calculations such as ours.

[10] This notation allows for $\hat{\rho}$ to differ from ρ for expositional purposes, taking into account the short remaining life expectancy of the average COVID victim. It will turn out that the following calculations are (approximately) independent of the discount rate.

To see how to map this into our model, ask someone with preferences as in equation (22) what fraction x of her consumption she would be willing to give up to avoid a 0.1 percent probability of death. If the answer is $31.755\hat{\rho}$, then v solves

$$\frac{\log(1)}{\hat{\rho}} - 0.001v = \frac{\log\left(1 - 31.755\hat{\rho}\right)}{\hat{\rho}}$$

or $v = -1000\log\left(1 - 31.755\hat{\rho}\right)/\hat{\rho} \approx 31{,}755$ in our model units. The approximation is valid when $\hat{\rho}$ is small, as is the case when we use a daily discount rate. Note that this represents an "exchange rate" from model utils to US dollars of 31,755/3,915,000 $\approx 1/123$. We use that rate when reporting on the dollar cost of the disease below.

These values are similar to the ones chosen in several other recent paper on the outbreak. For instance, Alvarez et al. (2020) fixed the IFR at one percent and the VSL at \$1.33 million. Furthermore, summing the age-specific VSL from Greenstone and Nigam (2020) and weighting them by the CDC estimates of age-specific death rates leads to a VSL of \$4.5 million, very similar to the number we use. We also offer a robustness exercise below where we consider reducing κ by fifty percent to reflect a lower IFR or VSL.

Finally, we calibrate the initial infection rate, $N_i(0)$, to match the 51 people who died from COVID-19 in the US on or before March 13, 2020, the date 0 of our quantified model (https://coronavirus.1point3acres.com/en, retrieved July 1, 2020). Given a population of 328 million individuals and an IFR of $\pi = 0.0062$, this implies that $N_r(0) + N_d(0) = \frac{51}{328\times 10^6 \pi} = 0.0000251$. We assume that prior to date 0, people were unaware of the disease and so set $A(t) = a^* = 1$ for $t < 0$. Following the dynamics of the SIR model, this implies $N_i(0) = 1 - (N_r(0) + N_d(0)) - \exp(-R_0(N_r(0) + N_d(0))) = 0.0000527$, leaving the remaining share 0.9999223 of the people susceptible to the disease.

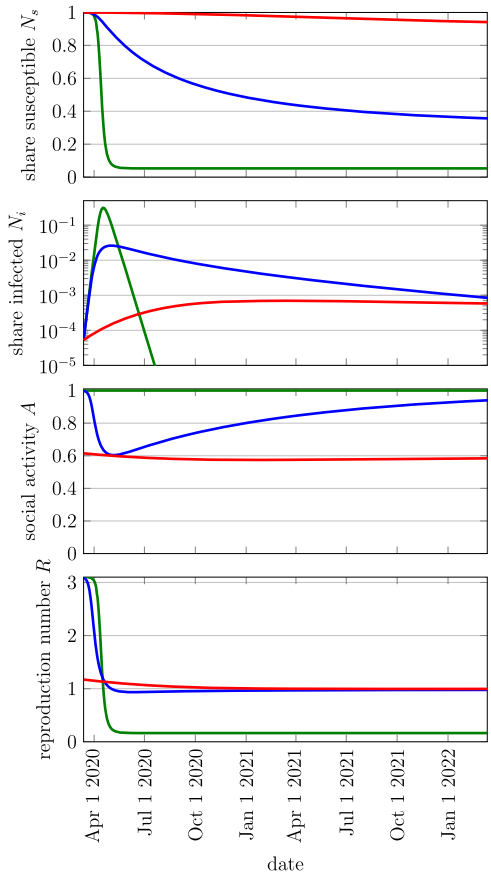
5.2. Basic SIRD model

We next turn to our quantitative results. As a simple benchmark, we begin with the basic SIRD model before turning to laissez-faire equilibrium and the social optimum.

The green line in Fig. 4 plots the dynamics of the pandemic in the SIRD model, that is without any behavioral response, $A(t) = a^* = 1$, under the assumption that a cure is not found. The top panel shows the share of people susceptible in levels. The second panel shows the share infected on a log scale. According to this model, the pandemic would have unfolded rapidly even though the infection rate started at half a basis point in mid March. The infection rate would have peaked in the second half of April at above 31 percent. As a consequence of the very high level of infections when reaching the herd immunity threshold, there would be large epidemic overshoot. As a consequence, by June 1st, only 5.5 percent of the population would have remained susceptible.

This model fails to capture the experience virtually anywhere in the world, even in places did not institute any restrictions on social activity. For instance, Sweden hits 1063 total cases, approximately 0.01 percent of the population, on March 15, 2020. One month later, this number rose to 12,443 confirmed cases, 0.12 percent of the population, despite the laissez-faire approach taken by the Swedish government.[11] In contrast, our calibrated SIRD model without any behavioral response predicts that it should take about eight days for the total case rate to rise from 0.01

[11] Retrieved from https://www.worldometers.info/coronavirus/country/sweden/.



Notes: See Table 1 for calibration. The second plot is drawn on a log scale.

Fig. 4. SIRD vs. Laissez-Faire vs. Optimal Policy.

percent to 0.12 percent of the population. Even if the total case rate is understated by an order of magnitude, approximately exponential growth implies that it still takes only nine days for the infection rate to increase from 0.1 percent to 1.2 percent of the population in the SIRD model.

Likewise, SafeGraph and Google data document a sharp decline in individual social activity even in laissez-faire settings. The basic SIRD model with constant parameters misses the fact that individual behavior responds to the risk of infection.

5.3. *Laissez-faire equilibrium*

We next turn to the disease dynamics in our laissez-faire equilibrium, which we depict as the blue line in Fig. 4. The difference between laissez-faire and the basic SIRD model is stark. Despite the government not intervening at all, the peak infection rate is less than one tenth of the level in the SIRD model, 2.6 percent at the beginning of May. In turn, the response of individual behavior substantially prolongs the epidemic, with the infection rate staying elevated for a much longer time. That is, equilibrium social distancing flattens the infection curve.

The third panel shows that individuals reduce their social activity by as much as 40 percent. However, the contraction in social activity does not occur immediately. Instead, it happens rapidly once the infection rate exceeds 0.1 percent. Conversely, activity starts to recover shortly after the infection rate peaks.

The fourth panel depicts the dynamics of the effective reproduction number $R(t)$. It is nearly equal to R_0 at the onset of the pandemic and then drops rapidly because of self-imposed social distancing. It falls below 1 in early May, exactly when the infection rate starts declining. Remarkably, it stays very close to 1 after this, never falling below 0.935. This is in line with the observation that N_i decays only slowly. We provide intuition for this finding in Section 5.5.

5.4. Optimal policy

The red line in Fig. 4 shows the optimal policy. Reflecting the external effects of social activity, one key property of the optimal policy is immediate social distancing in order to delay the spread of the infection and buy time to find a cure. While peak infection in the SIRD model occurs after 36 days and the equilibrium behavioral response delays the peak until 50 days have lapsed, the optimal policy delays it for almost a year. Because infections increase more slowly, the peak infection rate is almost forty times lower under the optimal policy than with laissez-faire, 0.07 percent.

On the other hand, once the infection rate peaks, it declines extraordinarily slowly along the optimal path. One way to quantify this is through the effective reproduction number $R(t)$, which never falls below 0.996. This means the infection rate never declines by more than 0.4 percent per week. We stress that this strong desire to "flatten the curve" is true even though there is no cost of peak-loading infection rates, i.e. even though κ is constant. A healthcare capacity constraint, where the cost κ is an increasing function of N_i, would make the case for flattening the curve even stronger. We have opted to omit this from the analysis to highlight that the optimal policy aggressively flattens the curve even without such a constraint.

We note that many places around the world have seen a "second wave" of infections. This is not what our model predicts either in equilibrium or in the optimum. We revisit this in Appendix D.

5.5. Key qualitative features

In Figures 14–18 in Appendix C.1, we change preferences, lower the cost of the disease κ, increase the duration of infectivity $1/\gamma$, reduce the cure probability δ, and start with a higher stock of initially infected $N_i(0)$. We find that there are some robust qualitative predictions of our model. Here we highlight four of them, the first two distinguish optimum from laissez-faire, and the last two are similar across these allocations.

Immediate social distancing The optimal policy delays the onset of the disease by imposing immediate social distancing, as depicted in the third panel of Fig. 4. In particular, the bottom panel shows that the optimal effective reproduction number is initially far below its uncurtailed counterpart $R_0 = 3.1$. Even if a full outbreak is eventually inevitable, the social gains from immediate social distancing are of first order because of discounting and because of the hope for a cure. The costs, however, are of second order, since $u'(a^*) = 0$.[12]

[12] In contrast, Alvarez et al. (2020) assume the population without disease is in a corner and so lockdowns have a first order cost.

On the other hand, there is no noticeable reduction in social activity under laissez-faire at the initial outbreak of the disease. As a result, the laissez-faire effective reproduction number is indistinguishable from R_0 in the early days of the disease. This reflects the fact that there is little private incentive to lower social activity when the risk of individual infection is negligible, and so individuals do not restrict activity until infections are more prevalent.

Persistent social distancing Under laissez-faire, social activity is almost back to its pre-pandemic level after 2 years. This is not the case under the optimal solution, which curtails activity for decades or until a cure is found. That is the flip-side of immediate social distancing. The optimal solution delays the onset of the disease in the hope of finding a cure. If no cure is found, herd immunity only builds very slowly and so restrictions on social activity must persist far longer than under laissez-faire.

Long run We now highlight three observations about the long run, common to both equilibrium and the social optimum.[13] Of course, those are only relevant in the case where no cure is found for a long time.

First, the disease vanishes in the long run. This reflects two assumptions. First, there are no newly-born individuals added to the population. Second, individuals attain permanent immunity after infection. Without either of these assumptions, an "interior steady state" with a strictly positive stock of infections would naturally emerge.[14] Third, the probability that an individual stays sick for t periods converges to 0 as t grows large, a natural assumption.

The next observation is that social distancing vanishes asymptotically, with social activity returning to its pre-pandemic level a^*. This is not obvious ex-ante because it is feasible to suppress social activity permanently in order to further improve health outcomes, for example by setting $A(t) = 1/\sqrt{R_0}$. To gain some intuition for this result, suppose the planner pursued a path of permanently suppressing social activity. Eventually, allowing a bit more social activity becomes arbitrarily cheap because the number of infected people is so small that any meaningful outbreak of the disease is pushed far off into the future. In contrast, there is a first order gain from the increase in social activity because there are always susceptible individuals suffering from the suppression of activity.

The key assumption here is that the disease cannot be eradicated through social distancing, which seems a plausible assumption. For example, even if the disease is temporarily pushed out of the human population, there is evidence that it can survive in animals and then reenter the human population through zoonotic transmission. Isolated countries like New Zealand have seen the disease reappear on its shores after prolonged absences.

An immediate consequence of vanishing social distancing is that the population eventually reaches herd immunity, with $N_s(t)$ falling below $1/R_0$. This is because when social distancing vanishes, $A(t) \to 1$, infections shrink if and only if $N_s(t) < 1/R_0$; see equation (8). Put differ-

[13] Since we only plot the time series for 2 years these are hard to see for the planner. These are, however, numerically accurate observations. We solve the model until the infection rate falls to $N_i(T) = 10^{-9}$, which takes 569,000 days, or sometime in the year 3578. The probability of not finding a cure by then is vanishingly small.

[14] This is in contrast to our setting where, in the long run, $N_i \to 0$ due to herd immunity. In such an interior steady state, there would be a constant flow of new infections—either infected for the first time or re-infected—which would offset the outflow due to death and recovery. The reason why optimal policy does not wipe out infections via social distancing is exactly the same as in our setting—permanent suppression is not optimal. We do not believe that optimal policy would be qualitatively different in such a setting.

ently, vanishing social distancing and a disappearing disease are only mutually consistent if there is herd immunity in the long-run.

The last two observations are closely intertwined. If a vaccine/cure had never been found, the disease would not have vanished due to permanent social distancing, but due to herd immunity. We stress that there are feasible paths where non-negligible levels of social distancing remain until there is a cure, and so herd immunity is never achieved. Such paths are not optimal. Of course, the optimal path towards herd immunity is extremely slow and keeps infections at a very low level. This also means that it is a scenario highly unlikely to fully play out, given that a cure puts an end to the pandemic.

The effective reproduction number At the onset of the pandemic, the effective reproduction number $R(t)$ differs significantly between the laissez-faire equilibrium and social optimum. Specifically, under the optimal path it drops immediately due to social distancing, while the decline is more continuous under laissez-faire. But under both policies, $R(t)$ eventually crosses 1 and after this point, the two paths are remarkably similar, with $R(t)$ never falling far below 1, effectively rolling over the stock of infections.

To understand why $R(t)$ never falls much below 1, it helps to consider the evolution of the state variables $N_s(t)$ and $N_i(t)$. Consider first the point when infections reach their peak, that is when $R(t) = 1$. Since $N_i'(t)/N_i(t) = \gamma(R(t) - 1)$, the number of infections is constant at this point, with one recovering person replaced by a newly infected person. As a consequence, the only reason this is not actually a steady state is that the pool of susceptibles drains, $N_s'(t) < 0$. But since under our calibration $N_i(t)$ is never very high, $N_s(t)$ moves slowly. The slow moving state variables imply optimal behavior is also slow moving, and so $R(t)$ must stay very close to 1 even after the peak in infections.

We add two observations to this. First, $N_i(t)$ is higher and hence $N_s(t)$ moves faster under laissez-faire than under the optimal policy. This is why $R(t)$ falls further under laissez-faire compared to the optimum. Second, suppose the peak level of infections was much higher, as would be the case if the cost of the disease were low. In this case, $N_s(t)$ would be fast moving and we would see a bigger decline in $R(t)$ after the peak of infections. We verify this in Figures 15–17 in Appendix C.1. In each figure, we change some parameter to increase the peak infection rate. $R(t)$ declines by more, but the decline is modest in each case.

Another perspective is that there is an important stabilizing force pushing $R(t)$ towards 1. When $R(t)$ is larger than 1, the number of infections grows, discouraging social activity. But a lower level of social activity mechanically reduces $R(t)$. The reverse happens when $R(t)$ is smaller than 1. Thus the recovery in social activity never pushes $R(t)$ too far from 1.

Finally, we offer another way to think about the finding that social distancing is never too extreme. In both the laissez-faire equilibrium and optimal policy, the level of social activity $A(t)$ never falls below $1/\sqrt{R_0}$, the critical value below which social distancing would wipe out the disease for any $N_s(t) \leq 1$. This result is robust to a wide range of parameters. The only exception is if the initial infection rate is large. Even in this case, we have $A(t) < 1/\sqrt{R_0}$ initially, but social activity later rises above this threshold and never falls below it again.

6. Model versus data

This section compares our model's predictions with evidence on social activity, on the effective reproduction number, and on mortality in order to evaluate our model's quantitative predictions.

As of the writing of this paper, we have witnessed an enormous wave of infections and deaths through the late fall and winter of 2020–21, in particular on the Northern hemisphere. To accommodate this, we offer an extension in Appendix D that allows for a simple form of time-varying parameters. Here, we restrict attention to the period of spring through fall of 2020.

6.1. Social activity

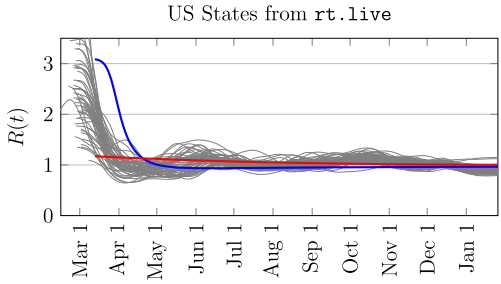
We first re-visit the quantitative evidence from Google from Section 3 and contrast it with our model's implications for social activity. Our hypothesis is that the model's laissez-faire equilibrium should capture individual behavior witnessed prior to implementation of any policy measures, even though we did not target the response of social activity to the COVID-19 outbreak in our calibration. Note that we do not claim that the hastily-implemented lockdowns and mobility restrictions were socially optimal.

Figs. 1 and 2 show an initial decline in activity of about 40 to 50 percent across the different metrics we study. In the US, much of this decline occurred before stay-at-home orders were issued and the decline was remarkably uniform across a variety of locations. Since the empirical metrics we have studied all have an inherent cardinality, their decline is quantitatively meaningful. Moreover, we can also compare these metrics to the decline in $A(t)$ in the laissez-faire equilibrium. As Fig. 4 shows, our model predicts that individual activity will decline by up to 40 percent in equilibrium. Thus, the model captures the magnitude of the decline in activity quite well.

On the other hand, while the speed of the decline in activity in the model is fast, the speed of decline in the data was even faster. In the model, $A(t)$ falls by 1.6 percent during the week of March 13–20 and takes three more weeks to fall by 31 percent, while in the data we observe a decline of that magnitude during the first week, before stay-at-home orders were implemented. This may reflect that individuals overestimated the risk of infection during this first week, a reasonable possibility given the scarcity of data and abundance of rumors. Alternatively, it may reflect altruistic behavior, which made the initial decline more in line with the red line in Fig. 4, i.e. an immediate 39 percent drop in activity as soon as people became aware of the disease.

The laissez-faire model also does a good job of capturing the speed and magnitude of the recovery in activity during the late spring, summer, and early fall of 2020. Fig. 3 displays a recover in retail and recreation activity which was nearly complete by July 1, before it started to reverse. The use of transit stations also peaked around the same time, although recovery was less complete. And the recovery in workplace activity was weaker and never really reversed. The SafeGraph data presented in Appendix A tell a similar story although there is certainly some variation across space and metrics. This is similar in our laissez-faire model as can be seen in Fig. 4. So if we think of the US as being close to laissez-faire since early June, our model also captures the evolution of activity during the summer in the US well. Of course, we caution that this comparison is somewhat heroic given the nature of earlier restrictions, the time-varying policies implemented, and the associated divergence of the state variables N_s and N_i from their paths in our benchmark model.

We could have directly targeted this type of data in our calibration, for example to select a value for the cost of the disease κ. That is, a natural approach would be to let the individual behavioral response "reveal" the perceived cost of infection instead of relying on direct measures of the infection mortality rate, π, and the value of a statistical life, v. Instead, we have opted with the direct measures which allows us to use the data on individual behavior, in particular prior to lockdown, as model validation.



Notes: We use results from rt.live who offer location-specific time series estimates for $R(t)$. Each thin grey line corresponds to one US state. We also plot equilibrium and optimum for the model which starts at March 13th.

Fig. 5. Effective reproduction number $R(t)$ throughout the pandemic.

6.2. *Effective reproduction number: roll-over after peak*

As we have noted, a robust feature of our analysis is that the effective reproduction number $R(t)$ remains close to 1 after crossing this threshold. This implies infections decline very slowly both under laissez-faire and under the optimal policy.

We now compare this feature with its direct empirical counterpart, a measure of $R(t)$ provided by the website `rt.live`. This uses time series on testing volume and positive test to estimate infection rates and hence the effective reproduction number. It offers separate measures for all 50 US states and the District of Columbia.[15]

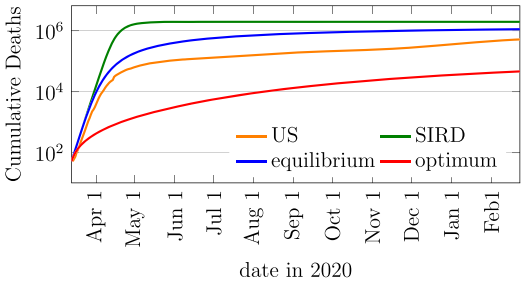
We plot the time series for each state separately in Fig. 5, along with equilibrium and optimal counterpart. We first note that the effective reproduction number in the data fell earlier and slightly faster than the laissez-faire model predicts. This is consistent with the overall picture that the evolution of the pandemic in the US falls somewhere "in between" laissez-faire and optimum.

More importantly, the data show an effective reproduction number of 1 across most US states for most of the pandemic, the same roll-over feature that arises in the model. In particular, $R(t)$ never fell much below 1 in all US states. The largest deviations of $R(t)$ from 1 occurred in the fall, when infection rates increased during a second wave of illness (see Appendix D for our take on later waves).

6.3. *Fatalities*

We now turn to the model's predictions for US COVID-19 fatalities. In Fig. 6, we plot the cumulative fatality rate in our model, which is just $\pi(1 - N_s(t) - N_i(t))$, multiplied by the US population of 328 million people. We plot this for the basic SIRD model along with the laissez-faire and optimal policy paths implied by our model. All four plots start with total COVID-19 fatalities for the US on March 13, 2020. According to the SIRD model, over 1.9 million Americans would have died by the mid-February 2021, almost all before May 1, 2020. Equilibrium behavior alone would have reduced that to 1,077,000. Actual deaths were lower, 484,700 on Feb 13, despite the second wave visible in the US data. This shows the importance of some combi-

[15] Data retrieved from download link on `rt.live`, https://d14wlfuexuxgcm.cloudfront.net/covid/rt.csv, on 2/11/2021.



Notes: Cumulative fatalities (without a cure) in the model and data. US data comes from https://coronavirus.1point3acres.com/en, retrieved on Feb 23, 2020. We calibrate the model to match data on cumulative deaths on March 13.

Fig. 6. Cumulative Fatalities.

nation of altruism and mandated social distancing. On the other hand, the optimal policy would have crushed this outcome, with only 41,046 dead.[16]

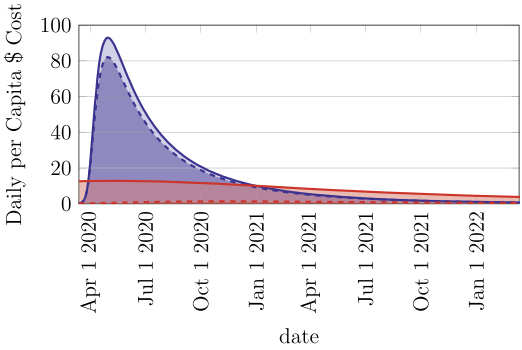
6.4. *Summing it up: the cost of the disease*

We close this section by using the model to understand the cost of the disease under different policies. In Fig. 7, we plot expected flow utility in the laissez-faire equilibrium and under optimal policy. We break each of these into two components, the total health cost at any point in time, $\kappa\gamma N_i(t)$, and the total utility cost created by social distancing, $(N_s(t)+N_i(t))u(A(t))$. We then discount these back to time 0, accounting both for impatience and for the chance of a cure. Finally, we convert utils into US dollars using the exchange rate discussed in Section 5.1.

In Table 2, we integrate these curves over the infinite future to find the expected present value of the disease and its components. We do this not only in the laissez-faire equilibrium and social optimum, but also under two extreme policies: the basic SIRD model, where $A(t)=1$ for all t by assumption; and a policy of permanently suppressing the disease, $A(t)=1/\sqrt{R_0}$ for all t, which ensures that $R(t)=N_s(t)<1$, so the disease always shrinks and herd immunity never emerges. The first column of the table reports the expected per capita health cost of the disease in dollars. The second column calculates the per capita dollar value of the expected reduction in social activity. Finally, the third column gives the total cost.

We first note that the composition of the cost differs sharply between the four policies. In the SIRD model, there is only a health cost by assumption. With a permanent lockdown at $A(t)=1/\sqrt{R_0}$, the health cost is negligible since the disease never takes off. We note that, in absolute terms the optimum comes fairly close to full suppression. At the same time, it does allow a small wave of infections and subsequent roll-over, resulting in health cost more than tenfold larger than

[16] It is worth noting that this mortality rate is higher than South Korea's or New Zealand's (both less than 0.001 percent). More broadly, while our model accounts for key features of the pandemic in the US and multiple other countries that shared a similar experience, there is vast heterogeneity across the globe in terms of how the pandemic played out over time. Some of this may be accounted for by different calibrations and our model might hence be able to explain it. Some of it might reflect geographical peculiarities that are not picked up in our model, such as the ability of an isolated island to close itself off from the world. And, of course, much of it might be driven by time-varying sub-optimal policies that reflect both political constraints and our still very limited understanding of the disease and its transmission.



Notes: Below dashed line: total daily per capita health cost $\kappa\gamma N_i(t)$ in period t. Between dashed and solid line: total daily per capita cost from social distancing, $(N_s(t) + N_i(t))u(A(t))$. We multiply each of those with $e^{-(\rho+\delta)t}$ to discount to time 0. We convert utils to US dollars using a factor of 123 as discussed in Section 5.1. Equilibrium in blue, optimum in red.

Fig. 7. Cost of Disease and its Components.

Table 2
Per Capita Welfare Cost of Disease.

	Health	Activity	Total
SIRD	\$21,223	0	\$21,223
Laissez-Faire	\$11,374	\$1,277	\$12,652
Optimum	\$860	\$7,227	\$8,087
$A(t) = \frac{1}{\sqrt{R_0}}$	\$78	\$8,334	\$8,412

Notes: First column: present value of per capita pure health cost of disease. Second column: present value of per capita cost of reduction in social activity. Third column: total cost. First row is the SIRD model, second row is the equilibrium, third row is the optimum, and the last row is a permanent lock down from the outset that never allows the disease to take off, $A = \frac{1}{\sqrt{R_0}}$. We convert utils to US dollars using a factor of 123, as discussed in Section 5.1.

those of a (non-optimal) full lockdown. The two more interesting policies fall in between. In the laissez-faire equilibrium, about 90 percent of the cost is health, while under the optimal policy, about 90 percent is the cost of social distancing. Fig. 7 shows that this decomposition does not vary dramatically over time.

Fig. 7 also shows that the time path of the cost differs sharply between laissez-faire and optimum. The planner flattens the curve through persistent social distancing, while the laissez-faire economy is effectively at herd immunity after 2 years. As a consequence, the expected flow costs are concentrated in the first year under laissez-faire, while a far larger fraction of the cost falls further into the future under the optimal path. At the same time, we see that the planner also incurs immediate costs from social distancing in excess of those that would occur in equilibrium.

Next, we note that the disease is very costly. The expected discounted flow cost per capita even under the optimal policy starts at $13 per person per day and is still at $4 per day two years later, with much of the decline reflecting the 74 percent possibility of a cure within that horizon. Under laissez-faire, the peak cost is $93 per day by early May, but then it declines rapidly. By the start of 2021, the expected daily flow costs under laissez-faire fall below those of the optimal path.

How do these numbers compare with outcomes we have seen in the real world? As we write this version of the paper, over half a million people have died in the US, a per capita health cost of $6004. At the same time, 2020 per capita GDP fell by $3250 below its 2010–2019 trend, while the unemployment rate remained nearly double the level of a year ago. These costs are again consistent with the hypothesis that the US has pursued a policy that lies somewhere in between laissez-faire and the optimum.

The benefits from pursuing the optimal policy are large. Comparing the aggregate costs in Table 2, the planner is able to cut expected per capita costs by 36 percent compared to laissez-faire. This reflects the large negative externalities of social activity during the COVID-19 outbreak, which in turn creates a strong case for mandatory social distancing policies. On the other hand, a policy of cutting social activity to $1/\sqrt{R_0}$ until a cure is found would reduce the mortality cost by another order of magnitude but would cause too much economic hardship, pushing total costs 4 percent above the optimal level. The observation that the gains from the optimal policy are small compared to what is achieved by this latter alternative reflects the fact that the optimal time path for infections keeps them at a very low level already (Fig. 4).

An important caveat is that these calculations correspond to the *expected* present value of costs, where uncertainty comes from the timing of the cure. If a cure comes very late, the optimal policy will create higher realized cost than the laissez-faire. This reflects the sharp difference in the time path of costs, with optimal policy delaying the disease in the hope of a cure. If there is no cure, then fatality rates are similar under the optimal policy and in laissez-faire since both regimes end with herd immunity, while the optimal policy imposes far greater costs of social distancing.

In Table 3 in Appendix C.1, we show how the level and composition of the cost of the pandemic vary with substantial changes to our calibration. We find that the laissez-faire results are quite robust, with health costs always accounting for more than 80 percent of the total. The gains from, and cost compositions of, optimal policy are more sensitive to the calibration. When the gains from social distancing are smaller because a cure is less likely or health costs are lower, we find that optimal policy allows an initial wave of infections. The same thing happens when the basic reproduction number is higher. In this case, effective social distancing requires a bigger reduction in activity, which the planner finds too costly. This means that in those cases, the laissez-faire policy is less inefficient and that the health costs of the two policies are more similar.

7. Discussion of key assumptions

Information structure We have assumed that individuals do not know they are sick until they recover or die. Most of the papers that we cite in Section 2 assume that individuals always know their health status, with Brotherhood et al. (2020) and Eichenbaum et al. (2021) being exceptions. In that case, individuals ramp up their social activity (or consumption and labor supply) at the moment they become sick because they have nothing left to lose. We do not believe that this is realistic, both because many individuals are pre-symptomatic or asymptomatic after they are infected, and because many people are altruistic enough to try not to get others sick.

In any case, if individuals know whether they are sick, eliminating the disease becomes both feasible and cheap with appropriate policy instruments. A planner who can directly gear policy towards the infected can isolate them and curb the spread of the disease at little cost. In practice, however, this only becomes relevant once widespread test, trace, and quarantine programs are available.

We have also assumed that the recovered know that they have gotten infected. This is inconsistent with the observation that many infected people never develop any symptoms. It is also inconsistent with the observation that the recovered were subject to the same restrictions on social activity as other people. Assuming the opposite has two opposing effects. On the one hand, social distancing is an even costlier tool of disease prevention because it applies to a larger population, especially in the later stages of the pandemic when many people have recovered. On the other hand, since social activity is still suppressed for recovered people, acquiring immunity through infection is somewhat less attractive. In Appendix C.2, we work through the math of the model and characterize both the laissez-faire policy and the social optimum. We find that changing this assumption does not have a big quantitative impact on either the laissez-faire policy or the optimal policy. Under laissez-faire, social activity already reverts back to normal fairly quickly after infections peak. It reverts back slightly faster when individuals do not realize they have recovered. Under the optimal policy, few people get sick and so the recovered population is of negligible size. The planner suppresses activity slightly more.

Available policy instruments The policy maker in our setting only has access to a very blunt policy tool: widespread social distancing. Perhaps most importantly, test, trace, and quarantine is unavailable to complement social distancing. This is a natural benchmark given the information structure we just discussed and given the realities of the pandemic thus far in the US and much of Europe.

The availability of a test, trace, and quarantine policy could alter our conclusions because eradicating the disease may become optimal (see Eichenbaum et al. (2021) and Berger et al. (2020) for related contributions that focus on test-trace-and-quarantine policies). The key observation is that when the share of infected $N_i(t)$ becomes small, widespread social distancing is a blunt tool because it primarily affects the large pool of susceptible individuals, $N_s(t)$. On the other hand, test, trace, and quarantine scales with the size of the infected population and so the cost vanishes as the disease vanishes. The important takeaway is that widespread social distancing is not a tool that should be used with the aim of eradicating the disease.

Healthcare capacity constraint We have assumed that the probability of death π and hence the cost of the disease κ are independent of the number of infected people $N_i(t)$. This ignores the possibility that a sudden surge in infections may overwhelm the healthcare system (see Alvarez et al. (2020) for a closely related paper where fatality risk depends on the stock of current infections). This is not because we want to deny the relevance of healthcare capacity constraints. They were clearly an important consideration during the early weeks of the COVID-19 outbreak; see https://twitter.com/drewaharris/status/1233267475036372992. Appendix C.3 shows how to extend our baseline model to allow for a hospitalization state without any additional computational complexity. It is then straightforward to also allow for an "ICU" constraint. However, our point is that economic logic would push us to flatten the curve and largely suppress the outbreak via social distancing even if there were no healthcare capacity constraints.

Heterogeneity The fatality risk of COVID-19 is far higher for older individuals and those who have certain pre-existing conditions. Our setup abstracts from such heterogeneity.[17] We believe that this picks up the realities of policy during the pandemic, which was applied across the board in a non-discriminatory fashion. Of course, for this reason our laissez-faire setting might look particularly bad since it does not take into account that those at a high risk would naturally reduce activity more and those with little risk less so. Nonetheless, we do not think that our key takeaways with respect to voluntary distancing, optimal policy, and the evolution of $R(t)$ after peak infections would be substantially altered in a setting with heterogeneous individuals.

Another source of heterogeneity lies in the propensity to spread the disease. In particular, epidemiologists have commented on the importance of super-spreaders for COVID-19. For example, Endo et al. (2020) estimates that about ten percent of infected people account for 80 percent of new infections. If doing so were feasible, it would be optimal to focus social distancing on those people with a high propensity to be super-spreaders. Since it is unclear whether such people can be identified before the super-spreading events, the alternative is to focus social distancing policies on the activities where super-spreading is most likely, such as prisons, elder care facilities, and bars.

Deterministic arrival of a cure We have assumed that a cure for the disease arrives stochastically at rate δ. Alternatively, we can solve the model when the cure arrives deterministically at date T. This has little qualitative impact on our results in the initial phase of the disease. In particular, A, N_s, N_i, and R all behave very similarly to our benchmark, both in equilibrium and optimum. As T gets closer, however, social activity rises since the cost of infection falls to zero at the terminal date. This leads to a large increase in infections both in equilibrium and optimum, an implication that strikes us as problematic given the uncertainty inherent in medical progress.

Vaccine versus cure Finally, we have experimented with both a stochastic and a deterministic arrival rate of a vaccine, rather than a cure (see, e.g. Makris and Toxvaerd (2020) for how the (eventual) arrival of a vaccine versus an effective treatment affects equilibrium behavior). A vaccine effectively shifts people from susceptible to recovered but does not protect infected people. This alternative has no interesting quantitative impact on our results.

8. Conclusion

This paper uses a standard dynamic economic model to integrate privately optimal behavior and policy analysis into an epidemiological model of COVID-19. The model speaks to data on social activity and does a good job of describing observed behavioral patterns after the US woke up to the dangers posed by the disease in mid-March 2020.

Our quantitative exercises reveal several robust patterns. Even in laissez-faire, individuals significantly lower social activity in order to reduce the risk of getting infected. This behavior does not internalize the risk of getting other people sick, however. An optimal policy that internalizes this cost, immediately curtails social activity to buy time. Still, social distancing does not try to eradicate the disease. Instead, if there is no effective treatment, optimal social distancing only vanishes in the very long run, once the population achieves herd immunity. Along the way, it is

[17] See Brotherhood et al. (2020) and Acemoglu et al. (2020) for two related contributions that focus on age-specific risk and policies.

optimal to largely suppress the outbreak via social distancing yet to roll over a small stock of infections, which means that the effective reproduction number stays very close to 1.

Our framework is general and tractable. Equilibrium behavior and optimal policy are encoded in a set of two differential equations that can jointly be solved with the epidemiological block of the model. We therefore view the tools offered here as a natural building block to explore the role of alternative policies. The model can also be extended to analyze issues like heterogeneity and to explore the role of geography in determining disease transmission.

Appendix. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jet.2021.105293.

References

Abel, Andrew B., Panageas, Stavros, 2020. Optimal management of a pandemic in the short run and the long run. Technical Report. National Bureau of Economic Research.

Acemoglu, Daron, Chernozhukov, Victor, Werning, Iván, Whinston, Michael D., 2020. Optimal targeted lockdowns in a multi-group SIR model.

Alvarez, Fernando, Argente, David, Lippi, Francesco, 2020. A Simple Planning Problem for COVID-19 Lockdown.

Atkeson, Andrew, 2020. What will be the economic impact of COVID-19 in the US? Rough estimates of disease scenarios. Technical Report. National Bureau of Economic Research.

Barro, Robert J., Ursua, José F., Weng, Joanna, 2020. The Coronavirus and the Great Influenza Epidemic-Lessons from the coronavirus potential effects on mortality and economic activity.

Berger, David W., Herkenhoff, Kyle F., Mongey, Simon, 2020. An SEIR Infectious Disease Model with Testing and Conditional Quarantine. Technical Report. National Bureau of Economic Research.

Bethune, Zachary, Korinek, Anton, 2020. COVID-19 Infection Externalities: Pursuing Herd Immunity or Containment? Technical Report, Working Paper.

Bognanni, Mark, Hanley, Doug, Kolliner, Daniel, Mitman, Kurt, 2020. Economic Activity and COVID-19 Transmission: Evidence from an Estimated Economic-Epidemiological Model.

Brotherhood, Luiz, Kircher, Philipp, Santos, Cezar, Tertilt, Michèle, 2020. An economic model of the Covid-19 epidemic: the importance of testing and age-specific policies.

Brotherhood, Luiz, Cavalcanti, Tiago, Da Mata, Daniel, Santos, Cezar, 2021. Slums and Pandemics.

Budish, Eric, 2020. $R < 1$ as an Economic Constraint: Can We "Expand the Frontier" in the Fight Against Covid-19?

Dewatripont, Mathias, Goldman, Michel, Muraille, Eric, Platteau, Jean-Philippe, 2020. Rapid identification of workers immune to COVID-19 and virus-free: a priority to restart the economy. Technical Report, Discussion paper. Universit Libre de Bruxelles.

Diamond, Peter A., 1982. Aggregate demand management in search equilibrium. J. Polit. Econ. 90 (5), 881–894.

Diamond, Peter A., Maskin, Eric, 1979. An equilibrium analysis of search and breach of contract, I: steady states. Bell J. Econ. 10 (1), 282–316.

Dong, Ensheng, Du, Hongru, Gardner, Lauren, 2020. An interactive web-based dashboard to track COVID-19 in real time. In: The Lancet Infectious Diseases.

Eichenbaum, Martin S., Rebelo, Sergio, Trabandt, Mathias, 2020. The Macroeconomics of Epidemics.

Eichenbaum, Martin S., Rebelo, Sergio, Trabandt, Mathias, 2021. The Macroeconomics of Testing and Quarantining.

Endo, Akira, Abbott, Sam, Kucharski, Adam J., Funk, Sebastian, et al., 2020. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China.

Fenichel, Eli P., 2013. Economic considerations for social distancing and behavioral based policies during an epidemic. J. Health Econ. 32 (2), 440–451.

Garibaldi, Pietro, Moen, Espen R., Pissarides, Christopher A., 2020. Modelling contacts and transitions in the SIR epidemics model. Technical Report, CEPR Covid Economics Working Paper, Issue 5.

Glover, Andrew, Heathcote, Jonathan, Krueger, Dirk, Rios-Rull, Jose Victor Health versus Wealth: On the Distributional Effects of Controlling a Pandemic. Technical Report, Working Paper.

Greenstone, Michael, Nigam, Vishan, 2020. Does Social Distancing Matter.

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

Hall, Robert E., Jones, Charles I., Klenow, Peter J., 2020. Trading off Consumption and COVID-19 Deaths.

Jones, Callum, Philippon, Thomas, Venkateswaran, Venky, 2020. Optimal Mitigation Policies in a Pandemic: Social Distancing and Working from Home.

Kaplan, Greg, Moll, Benjamin, Violante, Gianluca Pandemics According to HANK. Technical Report, Working Paper.

Keppo, Jussi, Kudlyak, Marianna, Quercioli, Elena, Smith, Lones, Wilson, Andrea For Whom the Bell Tolls: Avoidance Behavior at Breakout in COVID19. Technical Report, Working Paper.

Kermack, William Ogilvy, McKendrick, Anderson G., 1927. A contribution to the mathematical theory of epidemics. Proc. R. Soc. Lond. Ser. A, Contain. Pap. Math. Phys. Character 115 (772), 700–721.

Kremer, Michael, 1996. Integrating behavioral choice into epidemiological models of AIDS. Q. J. Econ. 111 (2), 549–573.

Kremer, Michael, Morcom, Charles, 1998. The effect of changing sexual activity on HIV prevalence. Math. Biosci. 151 (1), 99–122.

Krueger, Dirk, Uhlig, Harald, Xie, Taojun, 2020. Macroeconomic Dynamics and Reallocation in an Epidemic. Technical Report, CEPR Covid Economics Working Paper, Issue 5.

Lauer, Stephen A., Grantz, Kyra H., Bi, Qifang, Jones, Forrest K., Zheng, Qulu, Meredith, Hannah R., Azman, Andrew S., Reich, Nicholas G., Lessler, Justin, 2020. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. Ann. Intern. Med.

Makris, Miltiadis, Toxvaerd, Flavio, 2020. Great Expectations: Social Distancing in Anticipation of Pharmaceutical Innovations.

Pan, An, Liu, Li, Wang, Chaolong, Guo, Huan, Hao, Xingjie, Wang, Qi, Huang, Jiao, He, Na, Yu, Hongjie, Lin, Xihong, Wei, Sheng, Wu, Tangchun, 2020. Association of public health interventions with the epidemiology of the COVID-19 outbreak in Wuhan, China. JAMA 232 (19), 1915–1923.

Philipson, Tomas J., 2000. Economic epidemiology and infectious diseases. In: Handbook of Health Economics, vol. 1B, pp. 1761–1799.

Philipson, Tomas J., Posner, Richard A., 1993. Private Choices and Public Health: The AIDS Epidemic in an Economic Perspective. Harvard University Press.

Philipson, Tomas J., Posner, Richard A., 1995. The microeconomics of the AIDS epidemic in Africa. Popul. Dev. Rev., 835–848.

Piguillem, Facundo, Shi, Liyan, 2020. The optimal Covid-19 quarantine and testing policies. Technical Report. Einaudi Institute for Economics and Finance (EIEF).

Rachel, Lukasz, 2020. An Analytical Model of Covid-19 Lockdowns. Technical Report. LSE.

Rowthorn, Bob R.E., Toxvaerd, Flavio, 2020. The optimal control of infectious diseases via prevention and treatment. Technical Report.

Sanche, Steven, Lin, Yen Ting, Xu, Chonggang, Romero-Severson, Ethan, Hengartner, Nick, Ke, Ruian, 2020. Early Release-High Contagiousness and Rapid Spread of Severe Acute Respiratory Syndrome Coronavirus 2.

Toxvaerd, Flavio, 2019. Rational disinhibition and externalities in prevention. Int. Econ. Rev. 60 (4), 1737–1755.

Toxvaerd, Flavio, 2020. Equilibrium Social Distancing. Technical Report.