Journal Pre-proof

Health versus Wealth: On the Distributional Effects of Controlling a Pandemic

Andrew Glover, Jonathan Heathcote, Dirk Krueger, José-Víctor Ríos-Rull

PII: S0304-3932(23)00086-7
DOI: https://doi.org/10.1016/j.jmoneco.2023.07.003
Reference: MONEC 3507

To appear in: *Journal of Monetary Economics*

Received date: 21 December 2021
Revised date: 7 July 2023
Accepted date: 7 July 2023

Please cite this article as: Andrew Glover, Jonathan Heathcote, Dirk Krueger, José-Víctor Ríos-Rull, Health versus Wealth: On the Distributional Effects of Controlling a Pandemic, *Journal of Monetary Economics* (2023), doi: https://doi.org/10.1016/j.jmoneco.2023.07.003

Health versus Wealth: On the Distributional Effects of Controlling a Pandemic*

Andrew Glover† Jonathan Heathcote‡ Dirk Krueger§

José-Víctor Ríos-Rull¶

Abstract

To slow the COVID-19 virus, many countries shut down parts of the economy. Older individuals have the most to gain from slowing virus diffusion. Younger workers in shuttered sectors have most to lose. We build a model in which economic activity and disease progression are jointly determined. Individuals differ by age, sector and health status. Disease transmission occurs at work, at home, through consumption, and in hospitals. Optimal economic shutdowns in 2020 are milder when taxes are distortionary, and when the government does not have access to debt. A harder, shorter shutdown is preferred when vaccines become available in early 2021.

*We thank George Alessandria, Andrew Atkeson, Tim Lee, Simon Mongey, Andreas Peichl, Josep Pijoan-Mas, Cezar Santos, Stefanie Stantcheva, Marc Suchard as well as several anonymous referees for very helpful comments, and attendees of virtual seminars at many institutions for valuable feedback. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Banks of Kansas City and Minneapolis or the Federal Reserve System. Krueger thanks the NSF (grant SES-175708.) for financial support.

†Federal Reserve Bank of Kansas City; `Andrew.glover@kc.frb.org`

‡**Corresponding Author:** Federal Reserve Bank of Minneapolis and CEPR; Research Department, Federal Reserve Bank of Minneapolis, 90 Hennepin Avenue, Minneapolis, MN 55480 (612) 204 6385, `heathcote@minneapolisfed.org`

§University of Pennsylvania, NBER, and CEPR; `dkrueger@econ.upenn.edu`

¶University of Pennsylvania, UCL, CAERP, CEPR and NBER; `vr0j@upenn.edu`

Keywords: COVID-19; Shutdowns; Redistribution; Economic Policy

J.E.L. classification codes: E1, I14, H12

Research Highlights

1. We explore the optimal shutdown policy in response to the COVID-19 pandemic.
2. Individuals in our quantitative model differ by age and by sector.
3. Redistribution through taxes, transfers, and debt is modeled explicitly.
4. Heterogeneity and costly redistribution significantly impact the optimal shutdown policy.
5. With vaccines a harder but short shutdown is preferred.

1 Introduction

A pandemic such as the COVID-19 crisis constitutes a large shock to global welfare, with adverse impacts on societal health and economic wealth. What is the optimal policy response to such shocks when social contact is central to both disease transmission and economic activity? Debate in 2020 centered on the question of how aggressively to restrict economic activity in order to slow the spread of a pandemic and how quickly to lift restrictions as it shows signs of subsiding, either naturally or in response to a vaccination campaign.

There is substantial disagreement about the answer to this question and about the factors, both in terms of the medical nature of the disease as well as the structure of the economy, that determine this answer. In this paper, we argue one source of disagreement is the fact that the benefits and costs of "lock-down" policies are large and very unequally distributed among different groups of the population. The young and the old and workers in sectors differentially impacted by lockdowns have vastly diverging preferences concerning these policies.

Standard epidemiological models miss this disagreement because they assume a representative agent setting in which all households face the same trade-off between restrictions on social interactions that slow disease transmission but also depress economic activity. In reality, for a pandemic such as COVID-19, the benefits of slower viral transmission accrue disproportionately to older households, which face a much higher risk of serious illness or death from infection. In contrast, the costs of reduced economic activity are disproportionately borne by younger households facing the brunt of lower employment. For these younger households, the costs of mitigation policies depend on their sector of work. Sensible lock-down policies designed to reduce viral spread focus on reducing activity in sectors in which there is a social aspect to consumption and sectors that produce goods or services perceived to be non-essential. We will call this part of the economy the "luxury sector" henceforth. During the COVID-19 pandemic, for example, restaurants, bars and other establishments in the broader hospitality sector were closed first. The fact that workers cannot easily reallocate across sectors implies that lock-downs have very disparate impacts on young households specialized in different sectors. Thus, different groups in the economy (old versus young, workers in different sectors, healthy versus sick) likely have very different views about the optimal mitigation strategy.

One way to try to build a coalition in favor of mitigation efforts is to use redistributive tax and transfer policies to mitigate the increase in economic inequality that shutdowns entail. However, redistribution is costly in practice. And the more costly it is, the larger and more unequal will be the economic costs of mitigation measures. It is therefore important to study optimal lockdown and redistribution policies jointly. This is what we set out to do in this paper.

To do so, we build a novel macro-epidemiological model of a health pandemic that incorporates the interaction between macro-mitigation and micro-redistribution policies. We then apply the model to study the optimal policy response to the COVID-19 crisis, both for the first phase of the pandemic in 2020, in which we assume no vaccines were on the horizon, and then for the second phase, starting in 2021, when a gradual roll out of effective vaccines took place.[1]

Our model has three key elements: (i) a household sector with heterogeneous individuals, (ii) an epidemiological block where consumption, production, and purely social interactions determine health transitions, and (iii) a government that can use distortionary taxes, transfers and debt to spread the economic costs of shutdowns across individuals and over time.

We distinguish between three types of people: young workers in a basic sector, young workers in a luxury sector, and old retired people. The output of workers in the two sectors is combined to produce a single final consumption good. Workers are immobile across sectors. Consumption of basic sector output does not contribute to virus transmission, and workers in this sector are not subject to shutdowns. In contrast, the policy maker can choose to reduce employment and output in the social contact-intensive luxury sector in order to reduce virus transmission during production and social consumption.

The epidemiological model builds on a standard Susceptible-Infectious-Recovered (SIR) diffusion framework but permits a richer set of health states that are quantitatively important for our analysis. We label our variant of the SIR model the *SAFER* model, reflecting the progression of individuals through a sequence of possible health states. Individuals start out as susceptible, S (i.e., healthy, but vulnerable to infection), and can then become infected but asymptomatic, A; infected with fever-like symptoms, F; infected and needing emergency hospital care, E; recovered, R (healthy and immune); or dead. The transition rates between these states vary with age:

[1] Our analysis of these two phases assume that policymakers do not foresee the emergence of new more contagious variants such as "Delta," which surged in the United States in late summer 2021, or "Omicron," which became the predominant strain in early 2022, which we think of as the third phase of the pandemic. Relatedly, they do not anticipate that in this third phase individuals who have experienced infection or been vaccinated can be subsequently reinfected. In this paper we focus on the first two phases of the pandemic. See Glover et al. (2022) for a partial analysis of the third phase with a focus on the Delta variant.

the old are much more likely to experience adverse health outcomes conditional on becoming infected.

At the heart of the model is the two-way interaction between the distribution of health and economic activity. We model virus transmission from co-workers in the workplace, from fellow consumers in the marketplace, from friends and family at home, and from the sick in hospitals. Because they do not work, the old do not face direct exposure at work, but virus transmission in the workplace indirectly increases infection rates in other settings. The three infected subgroups spread the virus in very different ways: the asymptomatic are unlikely to realize they are contagious and will continue to work and to consume; those with a fever will stay at home and infect only family members, while those in hospital care may pass the virus to health care workers.

The government maximizes a utilitarian social welfare function and has at its disposal two sets of policy levers. First, at each date it can choose what fraction of activity in the luxury sector to shut down, which we call the "extent of mitigation". Mitigation slows the spread of the virus by reducing the rate at which susceptible workers and consumers become infected, but it also reduces to zero the market income of mitigated workers. Second, the government chooses how much income to transfer to those not working, either because they are old, because they are unwell, or because their workplaces have been closed. Transfers must be ultimately financed out of taxes on workers, but in our baseline policy specification the government can use debt to smooth tax rates over time. A utilitarian government wants to redistribute, but internalizes that higher tax rates depress labor supply and output. When mitigation increases, the government optimally trades off equity versus efficiency by both increasing tax rates – implying larger efficiency costs – and by tolerating more inequality between workers and non-workers – implying less equity. We show that the distortions induced by redistribution reduce the dynamic incentives of the government for mitigation. In particular, we prove theoretically that the marginal welfare costs of mitigation are larger when redistribution is costly than when the government has access to lump-sum taxes.

We then use this model to characterize quantitatively the optimal path of mitigation, both for the first phase of the pandemic in 2020 in the absence of a vaccine and, separately, for the second phase (the first half of 2021) when effective vaccinations that protect individuals both from contracting and from spreading the disease are gradually administered. We calibrate the model to U.S. data and show that under the actual mitigation path the model captures the dynamics of COVID-19 related deaths well. We then ask what level and time path of economic lockdowns a utilitarian government would choose and how these contrast with the preferred policies of the three different groups of the population.

We highlight four findings. First, in the absence of a vaccine (and absent the expectation of one arriving in the near future) utilitarian optimal policy locks down about 30 percent of the nonessential sector in early 2020, with a temporary relaxation during the summer months, when infections and deaths are low. This is a compromise between vastly different policy preferences of different groups. To start with, one would expect disagreement between workers in the two different sectors, since only the luxury workers are subject to lockdown risk. However, this disagreement can effectively be addressed through a redistributive tax-transfer policy in which workers share the cost of lockdowns with the unemployed through higher taxes and transfers. Disagreement across age groups is much harder to deal with, as the old receive most of the health benefits of lockdowns and pay none of the costs in terms of higher taxes. As a result the old prefer much stricter and longer lockdowns than the young. By the same token, there is much more at stake for the old than the young, in that the welfare gains from switching from the benchmark mitigation path to the utilitarian-optimal one are about 40 times larger for the old than the young, while welfare differences across the two young groups are relatively minor.[2]

Second, the optimal mitigation path depends on the set of fiscal instruments to which the government has access. Suppose the government can impose type-specific lump-sum taxes, and that redistribution between workers and non-workers is therefore costless, as it implicitly is in representative agent models. Shutdowns are then less costly, and as a result optimal shutdowns are more extensive. We also consider a fiscal constitution under which the government must run a balanced budget date by date to quantify the importance of access to government debt. In that scenario, shutdowns necessitate immediate increases in tax rates, which tempers the planner's appetite for mitigation. The broader policy implication is that the optimal mitigation policy in response to a pandemic is sensitive to the details of the social insurance system and to the amount of fiscal space that a country enjoys.

Third, the optimal mitigation path is also strongly influenced by the details of how the threat to health from COVID-19 varies with age. Our utility specification incorporates an additive flow utility of being alive, whose value we calibrate to replicate a standard "value of a statistical life".

[2]The young account for 85 percent of the population in both the model and the data. Thus, if policy is determined by majority rule rather than by a utilitarian government, then one might expect shutdowns to reflect the preferences of the young and to therefore be too modest relative to the utilitarian optimum.

Given this specification, our planner is more concerned about total years of life lost due to COVID than about mortality *per se*. Because the risk of dying from COVID is much higher for older individuals with relatively low normal residual life expectancy, the planner in our economy chooses to mitigate less than it would if mortality rates did not vary by age, as in models based on the representative agent paradigm.

Fourth, expectations about vaccine arrival and distribution are critical determinants of the optimal profile for shutdown policies. If no vaccines are on the horizon (say in Spring 2020, the first phase of the pandemic), then the government cannot strongly affect the share of the population that will eventually fall ill. In such a scenario, optimal lockdowns are relatively modest and geared mostly toward avoiding excess demand for emergency hospital beds and associated excess mortality, which we model explicitly. In contrast, with knowledge that a vaccine is coming soon (the beginning of the second phase of the pandemic in early 2021), shutdowns can eliminate rather than merely delay infections, and much harsher lockdowns are optimal. We show that a utilitarian government chooses to lock down the entire luxury sector at the start of 2021, when it knows effective vaccines will be distributed, whereas the lockdown would be only half as extensive if there were no vaccines on the horizon. But as an increasing share of the population subsequently obtains immunity via vaccination, the government re-opens the economy much more rapidly than it would without a vaccine.

1.1 Related Literature

Our paper contributes to a by now substantial literature on the interaction between pandemics and economic activity, with a focus on the current COVID-19 crisis. Important early references include Atkeson (2020), Eichenbaum et al. (2021) and Argente et al. (2021).[3]

We wish to stress the following contributions of our paper relative to the literature. First, the paper is one of the first COVID-19 studies to explicitly incorporate multiple age groups. We emphasize not only the enormous age-related differences in the disease burden of COVID but also the stark policy disagreements across these different groups. We share the focus on the age dimension of heterogeneity with Acemoglu et al. (2021), Boppart et al. (2022) and Brotherhood et al. (2020).

[3]Other notable papers in the rapidly growing economics literature on COVID-19 include Fernandez-Villaverde and Jones (2022), Greenstone and Nigam (2020), Krueger et al. (2022), Toxvaerd (2020), Farboodi et al. (2021), Guerrieri et al. (2022), Bayer and Kuhn (2020), Berger et al. (2022), Chari et al. (2021), Hall et al. (2020) and Toda (2020).

A second distinctive feature of our framework is that the economic side of our environment is modeled in an explicit structural way. Each of the key household constituencies solves a maximization problem subject to a budget constraint, and there is no fictitious representative household that pools the economic costs of shutdowns. Therefore, the optimal policy cannot be reduced to a simple trade-off between lost output versus lives saved: the distribution of consumption and hours worked, as well as the distribution of mortality, are central policy considerations. We explore optimal policy from the perspective of a Ramsey government that uses realistically blunt policy instruments to affect household behavior. In contrast, most of the extant literature has focused on the dichotomous extremes of laissez-faire equilibrium or socially optimal allocations. The model of Kaplan et al. (2020) features richer heterogeneity than ours, but their paper does not study optimal lockdown policies or optimal redistribution.

A third novel feature of our model is that it allows us to characterize the optimal fiscal response to the COVID pandemic in closed form. In particular, we derive expressions for optimal time-varying transfers to non-workers and for optimal tax rates that depend on the expected present values of the measures of workers and non-workers in the economy. These in turn depend on the share of the population that is either mitigated or sick with COVID. We also show how the optimal fiscal response changes depending on whether or not the government has access to debt finance (see also Arellano et al. 2023). Our closed-form model of optimal redistribution could be applied in other contexts. One advantage of our model is that it accommodates transfers to individuals with zero market income, in contrast to the specification used by Benabou (2002) and Heathcote et al. (2017).

Finally, our paper is one of the few that explicitly model the interaction between the deployment of vaccines and optimal mitigation (see also Gonzalez-Eiras and Niepelt, 2020, Bognanni et al., 2020, Garriga et al., 2022). Gollier (2021) also explores the positive and normative effects of the extent and timing of vaccine deployment in a model with multiple age groups. But he considers neither the policy conflicts among age groups nor the jointly optimal redistribution and lockdown policies.

In Section 2, we describe how we model the joint evolution of the economy and the population. In Section 3, we then explain how we model mitigation and redistribution policies and the optimal policy problem. The calibration strategy is discussed in Section 4. The findings are in Section 5, and Section 6 discusses optimal policy in the presence of a vaccination campaign. Section 7 concludes.

2 The Model

We model an open economy in continuous time in which the government can borrow and lend freely at a fixed exogenous interest rate. We first describe the individual state space, spelling out the nature of heterogeneity by age and health status. In Section 2.2, we describe the multi-sector production technology and explain how mitigation shapes the pattern of production. Section 2.3 explains the details of our *SAFER* extension of the standard *SIR* epidemiological model and the channels of disease transmission. In Section 2.4 and Section 2.5 we set out the lifetime and period utility function and labor supply decisions, respectively, and Section 2.6 discusses the aggregation of the household and government sector. Section 2.7 discusses the key model assumptions we have made.

2.1 Household Heterogeneity

Time starts at $t = 0$ and evolves continuously. All economic variables, represented by Roman letters, are understood to be functions of time, but we suppress that dependence whenever there is no scope for confusion. Parameters are denoted with Greek letters. Generically, we use the letter x to denote population measures, with superscripts specifying subsets of the population.

Agents can be young or old, denoted by y and o. We think of the young as individuals below the age of 65 and their measure is given by x^y. For simplicity, and given the short time horizon of interest, we abstract from population growth and from aging and death unrelated to COVID-19 during the period of analysis.[4] Within each age group, agents are differentiated by health status, i, which can take six different values: susceptible s, asymptomatic a, miserable with a fever f, requiring emergency care e, recovered r, or dead d. Individuals in the first group have no immunity and are susceptible to infection. The a, f, and e groups all carry the virus – they are subsets of the infected I group in the standard *SIR* model – and can pass it onto others. However, they differ in their symptoms. The asymptomatic have no symptoms or only mild ones and thus unknowingly spread the virus. We model this state explicitly (in contrast to the prototypical *SIR* model), because a significant percentage of individuals infected with COVID-19 experience no or only very mild symptoms. Those with a fever are sufficiently sick to know they

[4]Thus, there are no individuals who enter the economy during the pandemic; for an analysis of the differential welfare effects of aggregate shocks between newborn and older individuals, see Glover et al. (2020b).

are likely contagious, and they stay at home and avoid the workplace and market consumption. Those requiring emergency care are hospitalized. The recovered are again healthy, no longer contagious, and immune from future infection. A worst-case virus progression is from susceptible (s) to asymptomatic (a) to fever (f) to emergency care (e) to dead (d).[5] However, recovery (r) is possible from any of the a, f and e states.

2.2 Economic Activity: Technology and Mitigation

Young agents in the model are further differentiated by the sector in which they can work. A mass x^b of the young work in an essential, basic b sector, while the rest of the young of mass x^ℓ work in a luxury, non-essential sector, denoted ℓ. Shutdowns only apply to the ℓ sector, and require some or all of the ℓ sector workers to stay at home in order to reduce virus transmission in the workplace and through luxury consumption activities. We call such a policy a mitigation policy, and use m_t to denote the fraction of luxury workers who are instructed to not work at time t. We assume that individuals cannot change the sector they work in. In terms of notation, superscripts denote the dimensions of household heterogeneity, that is, age, sector, and health status, in that order. For example, x^{ybs} is the measure of young individuals working in the basic sector who are in the susceptible health state.

The production technology is linear in labor input in both sectors. Thus, output in the basic sector is given by the measure $x^{bw} = x^{ybs} + x^{yba} + x^{ybr}$ of young workers employed there, times the number of hours h^b they work:

$$Y^b = \left[x^{ybs} + x^{yba} + x^{ybr}\right] h^b = x^{bw} h^b. \tag{1}$$

Note that this specification assumes that asymptomatic individuals carrying the virus continue to work while those with fever stay at home.[6] In contrast to the basic sector, output in the luxury sector depends on mitigation policy and is given by

$$Y^\ell(m_t) = (1 - m_t)\left[x^{y\ell s} + x^{y\ell a} + x^{y\ell r}\right] h^\ell = (1 - m_t)x^{\ell w} h^\ell. \tag{2}$$

[5] Note that in the standard SEIR model, agents in the exposed state E have been subjected to the virus and may fall ill, but until they enter the infected state I, they cannot pass the virus on. Our asymptomatic state is a hybrid of the E and the I states in the SEIR model: asymptomatic agents have no symptoms (as in the SEIR E state) but can pass the virus on (as in the SEIR I state). Berger et al. (2022) make a similar modeling choice.

[6] One could instead imagine a policy of tracing contacts of infected people, which would allow the government to keep some portion of asymptomatic workers at home.

Since production is linear in labor in both sectors with identical productivity, the relative price of luxury goods in terms of the basic good is 1, and GDP is given by $Y = Y^b + Y^\ell$.

2.3 Health Transitions: The *SAFER* Model

We now describe the dynamics of individuals across health states. At date $t = 0$, the total mass of living individuals is one, μ^y denotes the share that is young, and μ^b is the share of the young who work in the basic sector. At each point in time, we denote populations by age and sector by $x^{yb} = \sum_{i\in\{s,a,f,e,r\}} x^{ybi}$, $x^{y\ell} = \sum_{i\in\{s,a,f,e,r\}} x^{y\ell i}$, and $x^o = \sum_{i\in\{s,a,f,e,r\}} x^{oi}$. Thus, at $t = 0$, $x^{yb} = \mu^y\mu^b$, $x^{y\ell} = \mu^y(1 - \mu^b)$, and $x^o = (1 - \mu^y)$. At any point in time we will let $x^i = x^{ybi} + x^{y\ell i} + x^{oi}$ for $i \in \{s, a, f, e, r\}$ denote the total number of individuals in health state i. Finally, let $x = \sum_{i\in\{s,a,f,e,r\}} x^i = x^{yb} + x^{y\ell} + x^o$ denote the entire living population.

In our model, the crucial health transitions that can be affected by mitigation policies are from the susceptible to the asymptomatic state. The number of such workers who catch the virus is their mass (x^{ybs} for young basic sector workers, for example) times the number of virus-transmitting interactions they have. We model four sources of possible virus contagion: people can catch the virus from colleagues at work, from market consumption activities, from family or friends outside work, and from taking care of the sick in hospitals, which we index w, c, h, and e, respectively. For a given type of individual, the flow of new infections from each of these activities is the product of the number of contagious people they can expect to meet, denoted by $x_j(m_t)$ for $j \in \{w, c, h, e\}$, and the likelihood that such meetings result in infection, which we label infection-generating rates $\beta_j(m_t)$.

The numbers of contagious people in each activity are given by the following population measures

$$x_w(m_t) = x^{yba} + (1 - m_t)x^{y\ell a}, \tag{3}$$

$$x_c = x^a, \tag{4}$$

$$x_h = x^a + x^f, \tag{5}$$

$$x_e = x^e, \tag{6}$$

where these measures reflect the assumptions that symptomatic (s or e) individuals neither work nor shop and that basic and luxury sector workers can meet in the workplace. Note that the number of contagious workers depends on the mitigation choice m_t.

In modeling the infection-generating rates, we recognize that different sectors of the economy are heterogeneous with respect to the extent to which production and consumption generate risky social interaction. For example, some types of work and market consumption can easily be done at home, while for others, avoiding interaction is much harder. A sensible shutdown policy will first shutter those sub-sectors of the luxury sector that generate the most interaction. Absent detailed micro data on social interaction by sector, we model this in the following simple way. Assume workers are assigned to a unit interval of sub-sectors $i \in [0, 1]$ where sub-sectors are ranked from those generating the least social interaction to those generating the most. Also assume the sub-sector-specific infection-generating rates are $\beta_w^i = 2\alpha_w i$ and $\beta_c^i = 2\alpha_c i$, where (α_w, α_c) are parameters governing the intensity by which meetings generate infections. When the government asks a fraction m_t of luxury workers to stay at home, we assume it targets the sub-sectors generating the most interactions; that is, $i \in [1 - m_t, 1]$. The average infection-generating rates of the sub-sectors that remain are then $\alpha_w(1 - m_t)$ and $\alpha_c(1 - m_t)$, respectively.[7] By assumption, the government does not mitigate any workers in the basic sector, so the average workplace infection-generating rate in that sector is α_w. The economy-wide infection-generating rate for work-related infections is the following employment-share-weighted average across the two sectors:

$$\beta_w(m_t) = \frac{x^{bw}}{x^{bw} + (1 - m_t)x^{\ell w}} \times \alpha_w + \frac{(1 - m_t)x^{\ell w}}{x^{bw} + (1 - m_t)x^{\ell w}} \times \alpha_w(1 - m_t). \tag{7}$$

The infection-generating rate for consumption, $\beta_c(m_t)$, is similar, except that (i) only luxury consumption is associated with infection risk, and (ii) we assume that the infection-generating rate is proportional to the number of luxury sector workers working, which we think of as a proxy for the number of stores that are open. Thus,

$$\beta_c(m_t) = \frac{(1 - m_t)x^{\ell w}}{(1 - \mu^b)\mu^y} \times \alpha_c(1 - m_t), \tag{8}$$

where the denominator is the pre-mitigation number of luxury workers. The key property of these expressions is that as mitigation increases, the average social-interaction-generating rate falls. When all workers are healthy and there is no mitigation, $\beta_w(0) = \alpha_w$ and $\beta_c(0) = \alpha_c$.

[7] $E\left[2\alpha_w i | i \leq (1 - m_t)\right] = \frac{2\alpha_w}{1-m_t}\int_0^{1-m_t} i di = \frac{2\alpha_w}{1-m_t}\frac{(1-m_t)^2}{2} = \alpha_w(1 - m_t)$.

Equations (9)-(11) below capture the flows of basic sector workers, luxury sector workers, and older individuals out of the susceptible state and into the asymptomatic state:

$$\dot{x}^{ybs} = -\left[\beta_c(m_t)x_c + \beta_h x_h\right]\ x^{ybs} - \beta_w(m_t)x_w(m_t)\ x^{ybs} - \beta_e x_e\ x^{ybs}, \tag{9}$$

$$\dot{x}^{y\ell s} = -\left[\beta_c(m_t)x_c + \beta_h x_h\right]\ x^{y\ell s} - \beta_w(m_t)x_w(m_t)(1-m_t)x^{y\ell s}, \tag{10}$$

$$\dot{x}^{os} = -\left[\beta_c(m_t)x_c + \beta_h x_h\right]\ x^{os}. \tag{11}$$

Consider the first outflow rate in equation (9). The flow of young basic sector workers getting infected through consumption is the number of such workers who are susceptible, x^{ybs}, times the number of contagious shoppers, x_c, times the infection-generating rate, $\beta_c(m_t)$. The flow of young basic sector workers getting infected from co-workers is similarly constructed.

The rate at which young basic workers contract the virus at home, $\beta_h x_h$, depends on the number of contagious workers in the household, x_h, defined in equation (5). Note that both asymptomatic and fever-suffering individuals are at home. We assume that caring for those requiring emergency care is a task that falls entirely on basic workers. The risk of contracting the virus from this activity is proportional to the number of hospitalized people, $x_e = x^e$, with infection-generating rate β_e, which reflects the strength of precautions taken in hospitals.

Parallel to equation (9), equation (10) describes infections for the susceptible population working in the luxury sector. For this group, the risks of infection from market consumption and at home are identical to those for basic sector workers. However, individuals in this sector work reduced hours when $m_t > 0$ and thus have fewer work interactions in which they could get infected. Furthermore, luxury sector workers do not take care of hospital patients, and thus the last term in equation (9) is absent in equation (10). Equation (11) displays infections among the old who get infected only from market consumption and from interactions at home.

The remainder of the epidemiological block simply describes the transition of individuals through the health states (asymptomatic, fever-suffering, hospitalized, and recovered) once they have been infected. These transitions are described in equations (26) to (37) in Appendix A, with parameters that are allowed to vary by age. Transition into death occurs from the emergency care state at age-dependent rates $\sigma^{yed} + \varphi(x^e)$ and $\sigma^{oed} + \varphi(x^e)$, where φ is the excess mortality rate when hospital capacity is overused and is given by

$$\varphi(x^e) = \lambda_o \max\{x^e - \Theta, 0\}. \tag{12}$$

In (12) the term in the max operator defines the extent of hospital overuse given capacity Θ, treated as fixed in the time horizon under consideration. The parameter λ_o controls how much the death rate of the hospitalized rises (and the recovery rate falls) once capacity is exceeded.

2.4 Lifetime Utility Function

Preferences are defined over consumption and hours worked and also incorporate utility from being alive and being in a specific health state. Lifetime utility for the young is given by

$$E_0 \left\{ \int_0^{T^y} e^{-\rho t} \, S_t^y \left[u(c_t^y, h_t^y) + \overline{u} + \widehat{u}_t^i \right] dt \right\}, \tag{13}$$

where ρ denotes the discount rate, T^y is remaining life expectancy (absent premature death from COVID), and S_t^y denotes the probability of surviving to date t. Flow utility at date t is the sum of a term involving consumption and labor supply, $u(c_t^y, h_t^y)$, a flow utility from simply being alive, $\overline{u}$, and a flow value that varies with health state i, $\widehat{u}_t^i$. We assume that $\widehat{u}_t^s = \widehat{u}_t^a = \widehat{u}_t^r = 0$ and that $\widehat{u}_t^e < \widehat{u}_t^f < 0$. Thus, having a fever is bad, and being treated in the hospital is worse. If an individual dies of COVID, all utility terms are zero thereafter. Preferences for the old are similar, except that they do not work, so $h_t^o = 0$. In addition, the old have a shorter normal residual life expectancy, T^o, and face greater COVID mortality risk, reflected in lower survival probabilities, S_t^o.

In equilibrium, expected utility of a young individual will depend on the sector in which she works, for two reasons. First, sectors differ in the share of economic activity being shut down (and thus, for the individual worker, in the probability of being able to work when healthy). Second, a worker's sector will affect her distribution of health outcomes.

2.5 The Period Utility Function, Household Consumption and Labor Supply

Households value consumption and labor supply (if they work) according to the following Greenwood, Hercowitz and Huffman style utility function:

$$u(c,h) = \log\left(c - \frac{h^{1+\frac{1}{\chi}}}{1+\frac{1}{\chi}}\right),$$

where utility from household consumption $c = c^b + c^\ell$ is a linear aggregate of consumption in the two sectors. Since one unit of labor produces one unit of output in both sectors, the wage of in both sectors is $w^b = w^\ell = 1$. Recall that the sector an individual can work in is part of her type and cannot be adjusted during the time horizon under consideration.

The government taxes labor income at a flat rate τ and provides everybody not working with a transfer T, which is simply consumed. Healthy, non-mitigated households solve

$$\begin{aligned} \max_{c,h} U &= \log\left(c - \frac{h^{1+\frac{1}{\chi}}}{1+\frac{1}{\chi}}\right), \\ \text{s.t. } c &= (1-\tau)h, \end{aligned}$$

with solution

$$h = (1-\tau)^\chi, \tag{14}$$

$$c = (1-\tau)h = (1-\tau)^{1+\chi}, \tag{15}$$

$$U = -\log(1+\chi) + (1+\chi)\log(1-\tau). \tag{16}$$

For non-working households, the budget constraint and period utility are

$$c^n = T, \tag{17}$$

$$U^n = \log(c^n) = \log(T). \tag{18}$$

2.6 The Government and Aggregation

The government purchases goods of both sectors of the economy and raises taxes at rate τ to pay for these expenditures and for transfers T. We assume that government purchases are a constant share g of domestically produced output in both sectors: $G^i = gY^i$ for $i \in \{b, \ell\}$. For future reference, we write the measures of working, nonworking and total individuals as

$$\begin{aligned} x^w(m) &= x^{bs} + x^{ba} + x^{br} + (1-m)(x^{\ell s} + x^{\ell a} + x^{\ell r}) = x^{bw} + (1-m)x^{\ell w}, \\ x^n(m) &= x^o + x^{bf} + x^{be} + x^{\ell f} + x^{\ell e} + m(x^{\ell s} + x^{\ell a} + x^{\ell r}). \\ x(m) &= x^w(m) + x^n(m). \end{aligned}$$

2.7 Discussion of Model Assumptions

Our economy has the feature that the economic side of the model is analytically tractable. As we will shortly see, this makes it possible to jointly characterize optimal mitigation and redistribution policies. However, our simple model abstracts from several ingredients that may be important for modeling the COVID-19 pandemic.

First, and perhaps most importantly, individuals in our model cannot take any choices that directly reduce their infection risk.[8] Eichenbaum et al. (2021), Farboodi et al. (2021) and Engle et al. (2021) are important examples of models that endogenize behavioral responses to mitigate infection risk. If agents in our model could reduce their risk of infection by reducing hours of work or consumption, economic activity would endogenously contract when the pandemic hits, and there would be less need for government intervention to further restrict economic activity.

However, even if we were to give individuals scope to reduce their infection risk, government restrictions on activity would still be optimal, for two reasons. The first reason is the standard externality logic: when trading off the costs of reduced economic activity against the benefits of reduced infection risk, selfish individuals internalize the payoffs from reducing own infection risk, but not the benefits of reducing the risk of infecting others. The second reason we would not expect adequate endogenous private mitigation is that in our heterogeneous agent model, most virus transmission occurs among young individuals. For the young, the risk of dying from COVID is very small, and thus young individuals in the model will not choose to take very costly actions to reduce their infection risk. Thus, if we were to introduce it, endogenous private mitigation would be much more limited in our heterogeneous agent model than in a representative agent

[8] Our calibration captures these choices in a simple reduced-form way by imposing a one-time discrete decline in infection-generating rates when COVID first emerges, and a gradual increase in those rates in 2021 that we link to the pace of vaccinations.

setting.

In addition, with or without endogenous private mitigation, optimal policy trades off the marginal costs of additional economic restrictions against the marginal benefits of better health outcomes. Thus, given optimal policy, we would expect an extended model with private mitigation to exhibit very similar or identical dynamics for health and economic outcomes to the ones we compute. Only the decomposition between the portion of the economic slowdown that is endogenous versus policy-directed would change.

We have considered one natural extension of the model to introduce a behavioral margin by allowing individuals to choose to not work –i.e., to volunteer to be mitigated– and to thereby eliminate the risk of infection in the workplace. In this extension, workers who choose to not work receive transfers. We find that, given our utilitarian optimal path for mitigation, neither basic nor luxury sector workers would ever choose not to work.[9] The logic is simply that the private cost of being mitigated – in terms of lost current income – outweigh the private gains, in terms of reduced infection risk and extended longevity.[10]

Note that in this extended model, public and private mitigation measures are equally efficient, so there is no motive for the government to moderate government shutdowns in order to encourage more workers to voluntarily stay at home.[11] Thus, our utilitarian optimal mitigation path would

[9] The specific thought experiment is as follows. Given perfect foresight over the aggregate path of the pandemic, we evaluate expected lifetime utility at the start of the pandemic for young individuals in our baseline economy and for hypothetical young individuals who know they will not work for the duration of one future calendar month – for example, December 2020.

[10] To understand this, consider the following calculation that gives a sense of willingness-to-pay. Through December 1 2021, the fraction of Americans below the age of 65 who experienced COVID-related deaths was 0.072 percent (the CDC reports 195, 195 such deaths). Given our calibrated values for ρ, T^y and $\bar{u}$ we can ask how much consumption a young individual would be willing to forgo, for one year, to completely eliminate the risk of dying from COVID. Given consumption and hours equal to the per capita average values $\bar{c}$ and $\bar{h}$, the answer is the value for x that solves

$$\int_0^{365} \exp(-\rho t)\left(\log\left((1-x)\bar{c} - \frac{\bar{h}^{1+\frac{1}{\chi}}}{1+\frac{1}{\chi}}\right) + \bar{u}\right)dt + \int_{365}^{T^y} \exp(-\rho t)\left(\log\left(\bar{c} - \frac{\bar{h}^{1+\frac{1}{\chi}}}{1+\frac{1}{\chi}}\right) + \bar{u}\right)dt$$

$$= (1 - 0.01 * 0.072) \int_0^{T^y} \exp(-\rho t)\left(\log\left(\bar{c} - \frac{\bar{h}^{1+\frac{1}{\chi}}}{1+\frac{1}{\chi}}\right) + \bar{u}\right)dt$$

which is 10.6 percent of average annual consumption, or \$4, 675. This number is small relative to the income loss from not working, even when the utility benefit from the reduction in hours worked is taken into account.

[11] In this version of the model the government uses distortionary taxes to fund transfers to all non-workers, irrespective of whether transfer recipients have been told to not work by the government or have chosen to not work to reduce infection risk.

be unchanged if people could choose not to work.[12]

Two other model assumptions that are important for retaining tractability on the economic side of the model are that households cannot save, and that basic and luxury goods enter separably in preferences. See Kaplan et al. (2020) and Guerrieri et al. (2022) for examples of papers that allow for savings and that introduce richer sectoral variation and cross-sector spillovers.

On the health side, we refine the textbook SIR model by splitting the infected state into three sub-states, the asymptomatic A, fever F, and emergency room E states. Our motivation for doing so is two-fold. First, we think a quantitative analysis of optimal lockdowns needs to distinguish between infected individuals without and with symptoms. We find it plausible to assume that individuals with COVID symptoms (those in the F state) will stay home, and not work or shop. However, individuals infected with COVID are typically contagious for several days before they develop symptoms, and it is the presence of these individuals in the A state that creates a rationale for lockdowns. In particular, the value of lockdowns depends crucially on the relative shares of A versus F individuals in the population: lockdowns would be pointless in our model if every infected person was symptomatic (F state), and would be most valuable if the infected were all asymptomatic (A state). Thus, getting these shares approximately correct (which the *SAFER* model allows) seems crucial for quantifying optimal policy.

Second, the presence of the E state allows us to model the potential for hospital overload (especially in the first phase of the pandemic). Section 5.1 demonstrates that including this overload mechanism in the model is quantitatively very important for understanding why the optimal mitigation path seeks to "flatten the curve."

3 Fiscal Policy: Taxes, Transfers and Government Debt

As described above, in addition to financing exogenous public spending parameterized by g, the government is responsible for three choices: a path of mitigation (shutdowns) m_t, redistribution through proportional taxation on workers at rate τ_t and lump-sum transfers T_t to individuals who do not or cannot work (which include those unemployed because of shutdowns, those with fever or who are hospitalized, and those who have retired). These fiscal choices imply a path of

[12] One could imagine alternative models in which private mitigation is more efficient than public mitigation. For example, suppose COVID mortality risk varies across workers, and that individual workers are better informed about their idiosyncratic risk than the government. In such an environment a better policy than mitigating workers at random might be to let individuals decide whether or not to work, while providing transfers to those who choose not to.

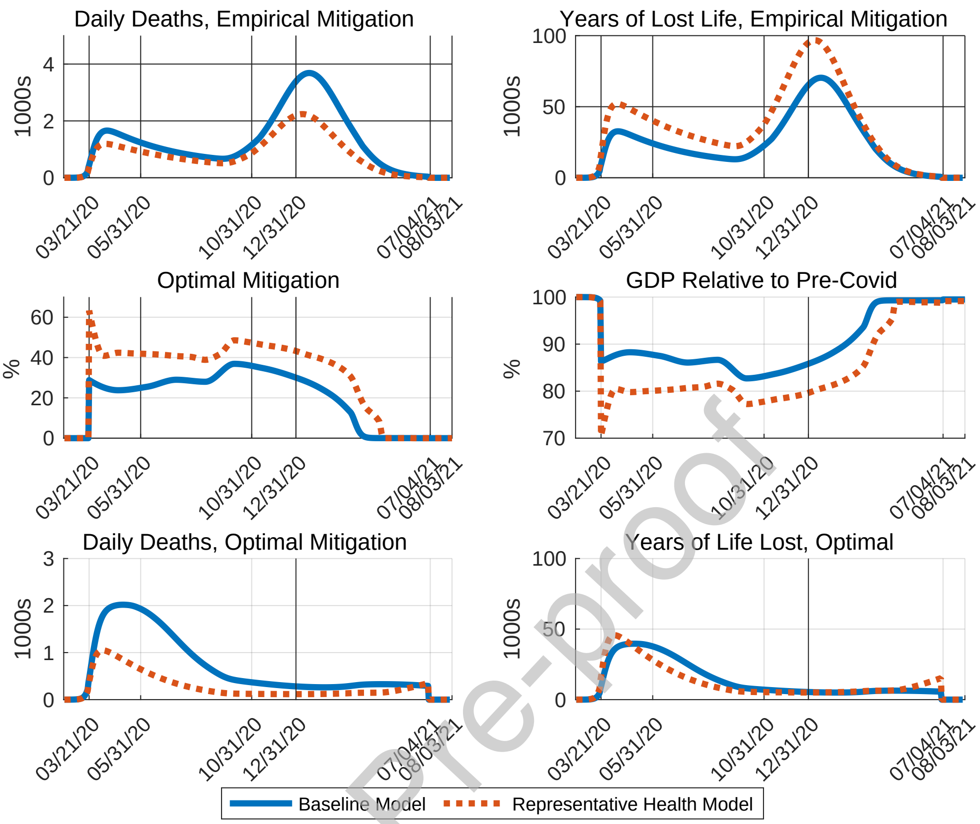

Figure 5: Role of Variation by Age in Health Threat from COVID-19: Baseline Model and Representative Health Model.

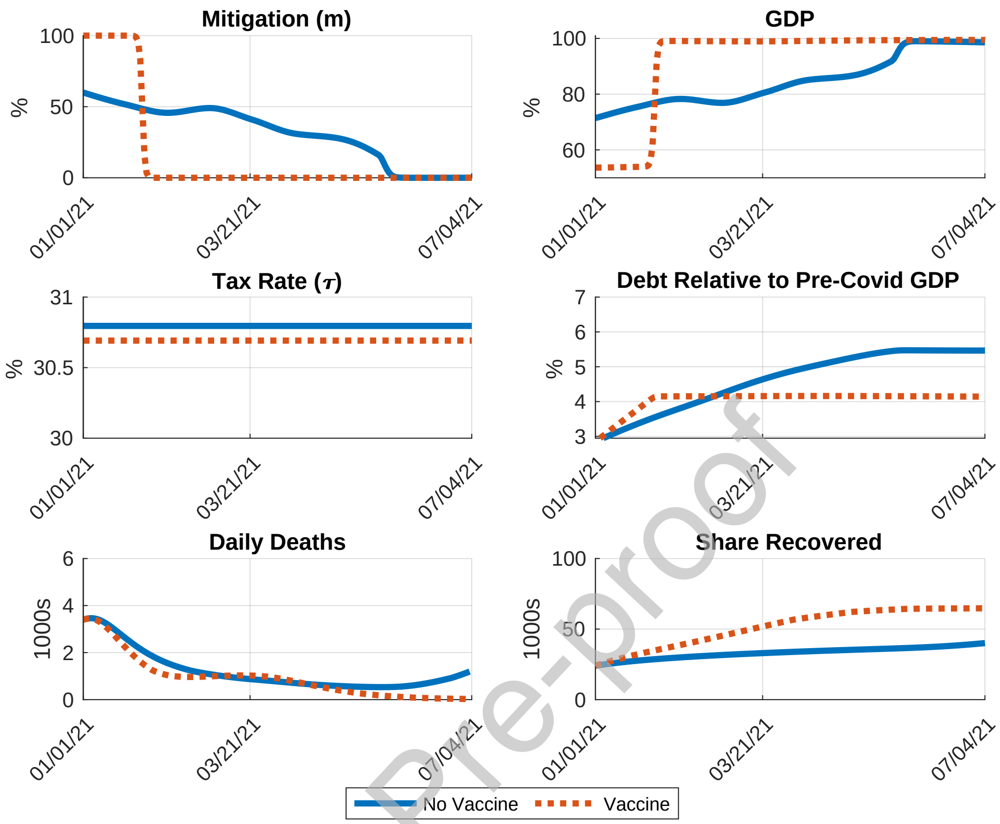

Figure 6: Utilitarian-Optimal Mitigation Paths with and without Vaccines