Health and Economic Inequality During Pandemics*

Aditya Goenka† Lin Liu‡ Haokun Pang §

August 29, 2024

Abstract

This paper studies the evolution of health and economic inequality during pandemics. It focuses on optimal preventive and treatment actions by agents who differ in their productivity, wealth and health status. Unlike the canonical Aiyagari-Bewley-Huggett framework where the income risk is exogenous, in our framework it is affected by individual choices and the endogenously determined economic and epidemiological variables. We have a heterogeneous agent continuous time (HACT) model where even if there are no exogenous productivity shocks the infection process introduces individual level heterogeneity and macroeconomic level fluctuations due to the interaction of the infection process, the control decisions, and effects on economic variables such as labour productivity and capital. Consistent with the evidence from the Covid-19, income, wealth, and health (proxied by infection rates) inequality increase during pandemics. The increase in income inequality is temporary but that of wealth is persistent. The increase in inequalities is driven by the increasing elasticities of health expenditures with wealth. We characterise the policy functions, the stationary equilibrium distributions and simulate the transitional dynamics to stationary equilibrium. We also evaluate the effect of government's income support scheme.

Keywords: Covid-19, SIRS model, Wealth and Health Heterogeneity, Aiyagari Model, HACT, Optimal Health Expenditure

JEL Classification: *E13, E22, D15, D50, D63, I10, I15, I18, O41, C61.*

*We thank seminar participants in the 2024 AMES Econometric Society; 2024 MMF Conference, Surrey; 2023 Conference on Advances in the economics and operational control of Covid, Rennes; CEMAP Conference 2023, Durham; SAET 2023, Paris; Rice and RSB Conference on Frontiers in Modelling in Epidemiological Economics 2023, Paris; 2023 EWMES Econometric Society, Manchester; 2024 AMSCW Econometric Society, Delhi; University of Birmingham, Delhi School of Economics, Presidency University, University of Sheffield, and Zhongnan University of Economics and Law, especially Raouf Boucekkine, Shankha Chabraborty, and Franck Portier for their helpful comments. The usual disclaimer applies. Aditya Goenka acknowledges research support from the British Academy.

†Department of Economics, University of Birmingham, Email: a.goenka@bham.ac.uk

‡Management School, University of Liverpool, Email: lin.liu@liverpool.ac.uk

§Department of Economics, University of Birmingham, Email: h.pang.1@bham.ac.uk

I. Introduction

The Covid-19 pandemic has had asymmetric effects in terms of health outcomes (infections and mortality)[1] as well as having increased income inequality with many countries (summarized in Stantcheva (2022)). The mechanisms driving these changes are not fully understood. Are the asymmetric outcomes driven only by the underlying heterogeneities such as age, household characteristics, occupations and ability to Work-From-Home, or are they also driven by endogenous optimal responses of individuals in response to the infections in society? Abstracting from all heterogeneities other than wealth inequality, the paper studies the role of optimal health and economic decisions in determining the joint evolution of inequality in health and economic outcomes. A parsimonious heterogeneous agent continuous time framework (HACT) is used to model the joint determination of health inequalities measured in terms of infections and economic inequality in terms of income and wealth inequality.

The paper incorporates a disease dynamics into a heterogeneous agent model, à la Aiyagari (1994). In Aiyagari (1994), income shocks are exogenously given. In our paper, income are directly affected by the health status of the individuals. Individuals, who are healthy either susceptible or recovered from the diseases, receive full income, while individuals who are infected receive no income. Moreover, the transmission probabilities between health states are endogenous, determined by health expenditures.

There are two types of health expenditures, incurred in response to a pandemic: (1) Preventive health expend expenditures that reduce consumption, for example, reducing mobility and isolating at home, cost incurred by wearing a mask and other preventive actions that reduce utility of consumption, etc. These reduce the chance of catching infectious diseases and thus, affects the transmission probability from being healthy and susceptible to being infected. (2) Treatment or recuperative expenditures incurred when being infected and requiring treatment. This is any expenditure increases recuperation rate and which reduces consumption. The change in health state acts like a productivity shock and as the health expenditures affect evolution of the health status it makes the individuals states endogenous. This is different from the classical Aiyagari model where the individual productivity shocks are exogenous. As individual productivity is determined by the health status, which

[1] See Blundell et al. (2020), ONS (2020), and PHE (2020) that document differences in health outcomes in UK across several dimensions — income deciles, OSA areas measured in terms of deprivation, rural-urban and regional differentials, age and gender differentials, attitudes towards risk, etc. Also see Belot et al. (2020), Borgonovi, Andrieu and Subramanian (2020), Brown and Ravallion (2020), Coven and Gupta (2020), Fan, Orhun and Turjeman (2020), Galasso et al. (2020), Lewandowski, Lipowska and Magda (2021), Mongey, Pilossoph and Weinberg (2021), Papageorge et al. (2021), Weill et al. (2020).

is controlled by health expenditures, the individual productivity shock becomes endogenous in our framework. In the Krusell-Smith extension of the classical model (Krusell and Smith, 1998; Fernández-Villaverde, Hurtado and Nuno, 2023), there can be an aggregate shock which is also exogenous. In our framework, as the individual decisions affect the evolution of infections which in turn affects wages and interest rates, the aggregate shock also becomes endogenous. Thus, our framework is an extension of the HACT model as in Achdou et al. (2022).

The epidemiological component of the model is a *SIRS* model which matches the dynamics of Covid-19. As we use a heterogeneous agent framework, the individuals take the decisions of other as given when choosing their optimal actions. As they are infinitesimal, they ignore the effect of their actions on the aggregate variables. Thus, we have the classical *disease externality* which has been modelled in different ways (see Gersovitz and Hammer (2004), and Goenka and Liu (2020)). In this mean field, we want to know how rational optimizing decisions by agents who differ in wealth and health status affect the equilibrium evolution of the disease and the economy. Thus, we shut down all other forms of heterogeneity such as education, ability to work from home, gender, age, location, access to health services, ethnicity, etc. which are important elements in matching the data.[2] This enables us to characterize the policy functions of the different agents and see to how these differ depending on wealth. If being ill is not just a productivity shock, but also brings a welfare loss, then the policy functions for preventive and treatment expenditures are increasing in wealth. In equilibrium only the susceptible individuals spend resources on prevention and only the infective on treatment. We also show that the elasticity of both types health expenditures are increasing in wealth. The endogenous optimal response to the pandemic, thus ends up increasing health inequalities as in equilibrium, the wealthier are less likely to get infected and recover faster. The two types of economic inequalities are affected by different mechanisms. Income inequality is largely affected by productivity (the health shock) and wages. The income inequality as documented in Stantcheva (2022) increases but its effects are not long lasting if there is an unanticipated shock to infection dynamics (which acts like a MIT shock). Wealth inequality is driven flow of income, returns on assets, and accumulation of assets. As capital adjusts slower than infections in the model (there are no adjustment costs) and health status of the wealthier is better, the wealth inequality also increases. This is much longer lasting in the event of an unanticipated change to disease dynamics.

The economic epidemiology literature has largely studied models where individ-

[2]See Adams-Prassl et al. (2020), Alipour, Falck and Schüller (2023), Bartik et al. (2020), Dingel and Neiman (2020), Gottlieb et al. (2021), Hensvik, Le Barbanchon and Rathelot (2020), and Lekfuangfu et al. (2020).

ual are homogeneous (Alvarez, Argente and Lippi, 2021; Atkeson, 2021) for early partial equilibrium models, Goenka, Liu and Nguyen (2021, 2022) for general equilibrium models with capital. Acemoglu et al. (2021) have age heterogeneity in a partial equilibrium framework. Kaplan, Moll and Violante (2020) have a HACT model with *SIRS* dynamics and wealth heterogeneity where the disease dynamics are exogenous. Angelopoulos et al. (2021) model wealth heterogeneity in a non-compartmental model so the disease dynamics are not explicitly modelled. Glover et al. (2023) model age hetergoeneity in an 3-period overlapping generations framework with SIR dynamics. In our framework, the model is a HACT model with capital and the *SIRS* dynamics are affected by optimal endogenous decisions. The model becomes a mean-field game. The usual approach in the earlier literature was to study an optimal control problem using Hamiltonians. There is a difficulty as one is not able to derive policy functions but only trajectories, and the sufficiency conditions can fail due to the non-convexity of the disease dynamics (see Goenka, Liu and Nguyen (2014, 2021, 2022)). As we use a dynamic programming approach, the second problem is avoided (see Calvia et al. (2023)). However, policy functions become difficult to compute. We adopt a finite-differencing method (Achdou et al. (2022)) to solve the computational issues.

In the model there are a continuum of individuals who can be in one of three health states, S healthy and Susceptible to the disease, I infected with the disease and Infective, or R, Recovered from the disease and immune from infection. This immunity is not long-lasting and individuals can become Susceptible as the virus mutates. They also have different wealth levels due to past saving decisions, so that that individual state is two-dimensional. They make optimal decisions on preventive and treatment actions, and how much to consume and save. Consistent with the later part of the pandemic we concentrate on morbidity from the disease which reduces labour productivity. The infection process acts like a productivity shock but this is endogenous and the health decisions act like a private partial self-insurance mechanism. The individual takes aggregate disease dynamics as given when making the health decisions and ignores the effect of own preventive and treatment actions on the evolution of the disease so that there is a disease externality. We characterize the optimal decisions and show how these depend on the individual states. In the baseline case where morbidity decreases utility, only the Susceptible make preventive expenditures and only the Infective on treatment. As mentioned above both the preventive and treatment expenditures are increasing in wealth and their elasticity is increasing as well. If morbidity acts only as a productivity shock then the optimal health expenditures are zero.[3] In a stationary equilibrium the wealthier individuals

[3]This is consistent with the results on the welfare loss from mortality in Goenka, Liu and Nguyen (2022).

have lower infection rates consistent with the higher health expenditures. One of the questions is if the wealthier have better lower infection rates to what extent do their decision increase inequalities. We look at different specifications and show that both preventive social distancing and treatment increase capital and income, reduce the disease incidence, but also increase the wealth inequality.

We also study the transitional dynamics to stationary equilibrium given a specific initial distribution of health status. The initial rate of recovered group is used to interpret the vaccination rate, as both vaccinated and recovered people carry immunity to infection. The simulated dynamics match the aggregate infection motion of Omicron B.A.1 outbreak in UK. As expected, infections increases that reduces labour supply and this drop in labour supply will increase wages. This increases the income inequality as the effect off falling ill is accentuated when wages are increasing. There is also a drop in investment leading to a decrease in interest rates. Other than the decrease in capital, the other changes are temporary. There is an increase in wealth inequality which is driven by the slow recovery of capital. The wealthier increase their health expenditures, which has leads to better health outcomes which also increase the differences is wages earned on average (note that in the model, the wage is the same for everyone who have the same productivity). The decrease in aggregate earned income and increased health expenditures, decreases aggregate consumption.[4]

We apply our model to evaluate the effect of government income support plan. We assume there is a government compensates the infected group by lump-sum transfer during the first 3 months after pandemic outbreak. Government budget constraint is imposed exogenously that the finance of budget is abstracted from the model. We find the income support generates trade-off between health and inequality at the aggregate level. The support scheme firstly reduces both income and wealth inequality and the effect is smaller when government has tighter constraint. Consistent with the empirical evidence, when the support is large enough, the rising wealth inequality could be turned around. However, the support scheme also mitigates the value loss of being infected and thus discourages health expenditure. Therefore, the aggregate infection rate is higher and the equilibrium labour and production are lower with income support.

The plan of the paper is as follows: section II. present new empirical facts on wealth distribution during Covid-19 pandemic, section III. presents the model, sec-

[4]Chetty et al. (2023) show that there are differences in consumption and savings across income groups. The income groups in their will be highly correlated with the wealth groups in our model. Their evidence indicates the high income groups sharply reduced their consumption and increased savings. While this also happens in our model, we will be able to match their results better if we distinguished between consumption of goods and services. The decrease in consumption documented in the paper by higher income groups was largely driven by decrease in consumption of services.

tion IV. addresses the computation and calibration, section V. characterises the stationary distribution and policy functions, section VI. studies transitional dynamics under various MIT shocks, and section VII. concludes.

II. Empirical Evidence

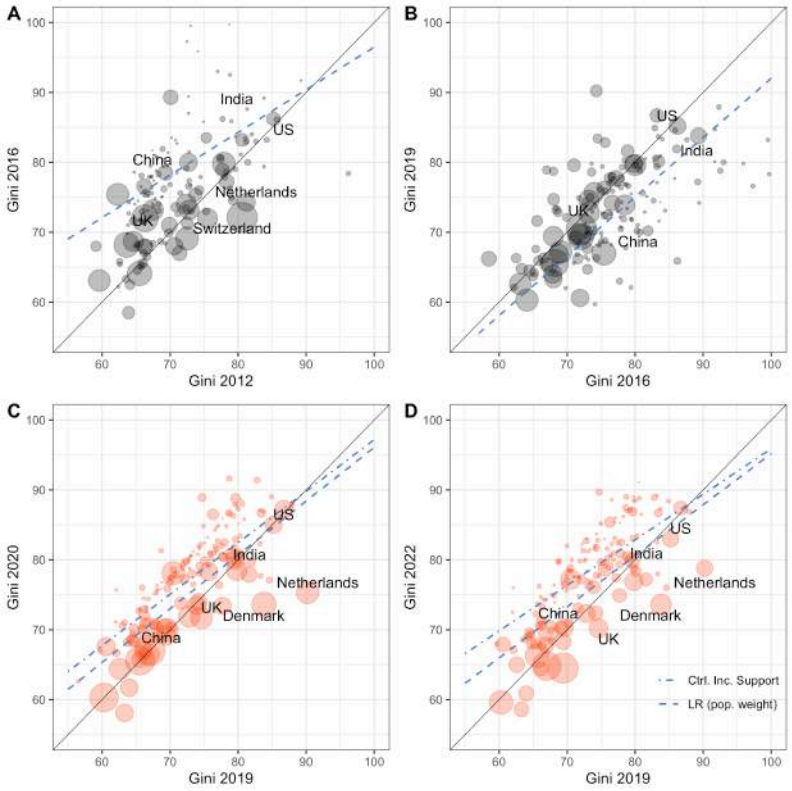
In this part of the paper, we present some empirical evidence. We firstly focus on the change of wealth inequality from 2012 to 2022 using the data in Global Wealth Report (GWR) by Credit Suisse.[5] The reports are released annually, providing information on wealth inequality of 165 countries from 2010. Figure 1 plots the wealth Gini index of each country at 2016, 2019, 2020, 2022 against the index at the base year.[6] The size of the scatter reflects the wealth per capita for each country. In Panel A, we find the wealth inequality worsened during the period of 2012-2016, with most countries lying above the 45 degree line. For the subsequent phase from 2016 to 2019, as shown in Panel B, a large number of countries, especially for those poor economies, shift below the 45 degree line. This implies a trend of improving wealth equality. When it goes to the post-Covid era in Panel C and D, we observe different trajectories that economies with higher wealth per capita, indicated by larger scatter, locate below the 45 degree line. The poorer economies with smaller scatters, in contrast, lie above the line. This evidence shows that the wealth equality is further progressive after 2019 for some richer economies like UK and Netherlands, but is worsened for economies with less wealth. The divergence in wealth Gini is even more significant when it comes to a relatively longer run towards 2022 in Panel D.

We suspect the different motion of wealth inequality between richer and poorer countries is due to the government income support scheme in Covid-19 relief. The Covid-19 Government Response Tracker (Hale et al., 2022) shows that there are 168 out of 185 countries ever covered part of the income loss during the pandemic. The short-run policy response could turn the direction of income/wealth equality (Stantcheva, 2022). Hence, we merge the GWR data with the tracker data. The dash-dot line in Panel C and D of Figure 1 are the fitted value of linear model,

[5] See https://www.credit-suisse.com/about-us/en/reports-research/global-wealth-report.html

[6] In the time series plot at Appendix Figure A1, we could identify that the global wealth Gini index increased from 2012 to 2016 and then decreased. We hence split the sample period into 3 phases: before 2016; 2016-2019 and the post Covid era from 2020

Figure 1 – Change of Gini Index



Notes: (a) This figure shows the change of Gini index from 2012 to 2019. The horizontal axis is the Gini index at the base year (2012,2016,2019). The horizontal axis is the Gini index evaluation year (2016, 2019, 2020, 2022). (b) The size of the scatters measures the medium of wealth per capita. (c) The dash line is the regression fitted value; the dash dot line is the fitted value removing the effect from government support scheme. All regressions are weighted by number of adults.

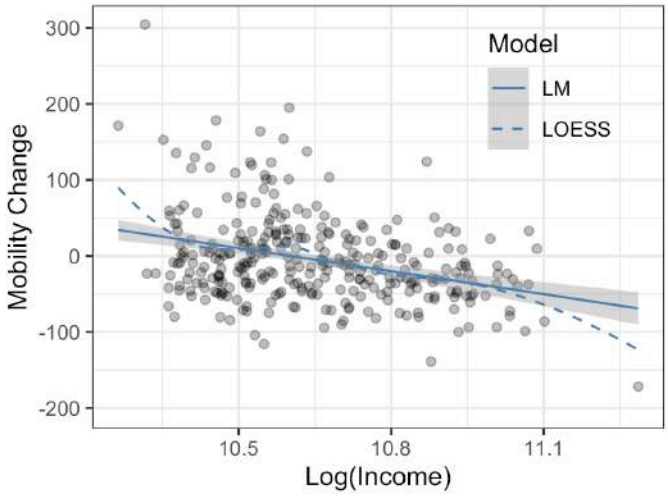
controlling the level of income support.[7] Compared with the effect without income support (dash line), we find that the fitted lines lie well above. It indicates that the wealth equality would be further worsened without government support scheme.

Our model will link the change of wealth inequality to individual optimal choice of health expenditure. However, it could be hard to provide supportive empirical evidence using the country data of the GWR. We thus turn to the subregional data of UK lower tier local authorities. We use the mobility index from Google Community Mobility Report (GCMR) data[8] as a proxy to general preventive health expenditure.

[7] We regress $Gini_i = \beta_0 + \beta_1 BaseGini_i + \beta_2 GovSupp_i + \varepsilon_i$, where $GovSupp_i$ is the average government income support index for country i from Feb 2020 to Dec 2023. The the dash-dot line removes the effect of support index that it plots the fitted value $\hat{Gini}_i = \hat{\beta}_0 + \hat{\beta}_1 BaseGini_i$. The dash line is the fitted value for the regression without controlling government support. Both regressions are weighted by number of adults in each country. The estimation results are shown in Appendix Table B1

[8] https://www.google.com/covid19/mobility/. The GCMR data provides the change of mobility index compared to Jan 2020. The data is of daily frequency, we collapse the data to monthly frequency by simple average.

Figure 2 – Income and Mobility on Oct 2022



Notes: (a) The horizontal axis is the log income for UK lower tier local authorities. The vertical axis is the change of mobility index related to Jan 2020. (b) The solid and dash lines are the fitted value of linear and Loess model.

Smaller value of the index implies higher preventive response to the disease. Then, we merge the mobility data with 2018 UK household annual income.[9]

We firstly consider cross-sectional evidence on Oct. of 2022 when all the social restrictions were removed to get rid of the impact of lockdown. Figure 2 plots the mobility change against the average log household income for UK local authorities. The solid line with gray shadow indicates the linear fit with 0.99 confidential interval to the data. We can observe a significant negative correlation between the change of community mobility with household income.[10] Furthermore, we notice that the fitted value is positive for the poor area while negative for the rich area. This implies individuals living in richer area have reduced mobility more, consistent with higher preventive actions after the pandemic while those in the poorer area do not. We also find that the linear model fits the correlation well, with small difference to the non-parametric LOESS fit (dash line) for most range of log income.

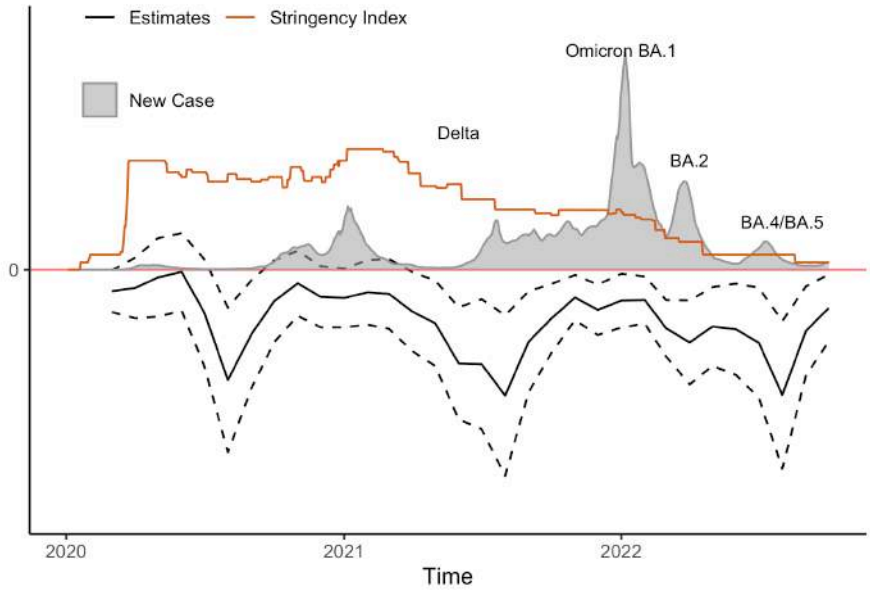
We further estimate the evolution of the marginal effect of income over time using Equation 1. In the specification, exploit the area-by-time variation. $M_{i,t}$ denotes the mobility index for local authority i at time t. The authority fixed effect is θ_i[11] and the time fixed effect η_t. Our main regressors are the interactive term between

[9] We average the MSOA level data to local authority level for merging.

[10] Weill et al. (2020) also find the same pattern in USA where the response to mandates for social distancing is much higher in richer counties than poorer counties. In our model there are no mandates and all response is endogenous. The results we present are also consistent with Yan et al. (2021) who find disentangle effects of mandates and endogenous response and find that the former increase the latter.

[11] This term absorbs all endogeneity generated by cross-sectional variety, e.g. sector composition that leads to different work-from-home rate.

Figure 3 – Income and Mobility



Notes: (a) The horizontal axis is the observation time. The black solid and dash lines are the estimates with 99% confidential interval for $\beta^{(\tau)}$ in regression Equation 1. (b) The gray area denotes the rate of new case of Covid infection in UK. (c) The orange line is the government stringency index. All components in this Figure are properly scaled.

log income $\log(I_i)$ and time indicator $T_t^{(\tau)}$ which equals 1 if time $t = \tau$, 0 for the rest. We set t = Feb2020 as the based time and exclude it from the summation part. By these settings, $\beta^{(\tau)}$ estimates the marginal effect on mobility $M_{i,t}$ from income, related to the base time. In the estimation, we also weight the samples by the population and cluster the standard error at the level of local authority.

$$M_{i,t} = \theta_i + \eta_t + \sum_{\tau \neq \text{Feb2020}} \beta^{(\tau)} \log(I_i) \times T_t^{(\tau)} + \varepsilon_{i,t}. \tag{1}$$

The black solid line in Figure 3 plots the corresponding estimates for $\beta^{(\tau)}$, with the 0.99 confidential interval on the dash line. We also plot the national infection rate in the gray shadowed area and the stringency index in the orange line.[12] The estimators are significantly negative for periods after the March of 2021, when the stringency index began to decline from its peak level. This implies individuals in the richer area reduced mobility more than those in the poorer area. We also detect a negative correlation between marginal effect $\beta^{(\tau)}$ with the infection rate for the periods after mid 2021. Therefore, the rich areas have higher prevention response to the disease when the infection rate is high.

In the next section we will build a heterogeneous agent model to model the

[12]The stringency is provided by the Government Response Tracker. It measures the level of social restriction. All the components in Figure 3 are properly scaled for purpose of visualization.

empirical observations. The model links the growing wealth inequality with individuals' optimal health policy. Our predictions are consistent with the empirical findings that the wealthier individuals take higher preventive actions (and treatment response) to the disease and thus subjected to smaller risk of income loss by infection during the pandemic. The heterogeneous response to aggregate infection generates income and subsequent wealth gap between the rich and poor.

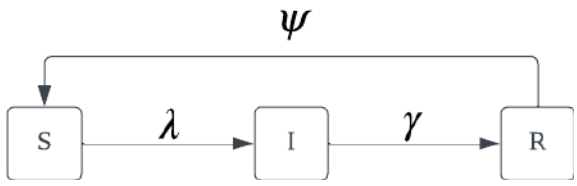
III. The Model

The model is a marriage of the *SIRS* epidemiological model and Aiyagari-Bewley-Huggett model as in Kaplan, Moll and Violante (2020). In the Aiyagari-Bewley-Huggett model, individuals face uninsured idiosyncratic labour market risk, which is exogenous. In our model, individuals face a health risk, where the probability of transiting between health status is endogenous and affected by two types of health expenditure - preventive and treatment health expenditure. This is partial self-insurance against the health shock. Thus, the model is different from Kaplan, Moll and Violante (2020) which does not model the response of individuals to the pandemic and thus, the partial self-insurance against the health shock.

A. The Model Setup

Individuals. We assume that the transmission of infectious diseases follow the SIRS epidemiology model. There is a continuum of individuals who have different health status $h \in [\mathcal{S}, \mathcal{I}, \mathcal{R}]$, denoting susceptible, infected and recovered, respectively. Individuals could be healthy and susceptible to infectious diseases, that is, $\mathcal{S}$. They can get infected at the rate of λ, with their health status becoming $\mathcal{I}$. The infected individual could recover from diseases at the rate of γ and gain immunity. The corresponding health status is $\mathcal{R}$. But, the recovered individuals could lose immunity with rate ψ and back to the health status $\mathcal{S}$. Figure 4 shows the flow chart for the health status.

Figure 4 – Flow of Epidemiological Compartments



Notes: This Figure shows the motion of health status. The numbers on the arrows are the transition probabilities.

In the model, individuals infected with infectious diseases are unable to work in the full capacity. Individual productivity z depends on his health status and we

assume $z : h \to [0, 1]$ with $0 \leq z(\mathcal{I}) < z(\mathcal{S}) = z(\mathcal{R}) = 1$.[13] Therefore, the idiosyncratic health shocks for individuals are essentially productivity shocks, which affect their income level. This model setup in fact is the same as the Aiyagari-Bewley-Huggett model where individuals face idiosyncratic income shocks. However, different from the Aiyagari-Bewley-Huggett model, where the income process is given exogenously, here the transition probability between health status is modelled endogenously. There are two types of health expenditure. Investment in preventive health expenditures $m_{\mathcal{P}}$ can reduce probability of getting infected λ, and investment in treatment health expenditures $m_{\mathcal{T}}$ can shorten the during of being infected and increase probability of getting recovered γ. With slight abuse, we will introduce health expenditures in detail later.

In the model, individuals also differ in their wealth level a. Denote w as wage and r as interest rate. They receive labour income $wz(h)$ and capital income ra, and accumulate wealth according to:

$$\dot{a} = wz(h) + ra - c - m_{\mathcal{P}} - m_{\mathcal{T}}, \tag{2}$$

where c is the consumption, $m_{\mathcal{P}}$ and $m_{\mathcal{T}}$ are the health expenditure. Individuals also face a borrowing constraint $a > \underline{a}$, where $-\infty < \underline{a} \leq 0$. The idiosyncratic component $z(h)$ generates heterogeneity in optimal policies and eventually in individual states. We denote the joint distribution on individual wealth and health as $g(a, h)$ with the probability measure μ.

Individuals derive utility from streams of current and future consumption c and health status h, discounted at rate $\rho \geq 0$. The consumption utility function is CES that $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$. Individuals also suffer disutility χ if infected. It captures the direct health lost of being infected beside productivity lost.[14] Hence, individuals maximize his utility by choosing streams of consumption c and health expenditures - $m_{\mathcal{P}}$ and $m_{\mathcal{T}}$:

$$\max_{\{c, m_{\mathcal{P}}, m_{\mathcal{T}}\}} \quad \mathbb{E}_0 \int_0^\infty e^{-\rho t}[u(c) - \chi \mathbb{1}_{(h=\mathcal{I})}]dt. \tag{3}$$

Infection and Recovery. We endogenize the flow of epidemiological compartments using health expenditure in a similar way as Goenka, Liu and Nguyen (2014). Specifically, for infection process, we assume individuals contact with other at a rate α, which is a decreasing function of preventive expenditure, i.e. $\alpha(m_{\mathcal{P}})$ in

[13] Patient clinical reported outcome finds infection brings productivity loss to workers (Di Fusco et al., 2022). We also model Long Covid by letting $z(\mathcal{R}) < 1$ later.

[14] See e.g. Acemoglu et al. (2021); Alvarez, Argente and Lippi (2021), the non-pecuniary welfare loss is usually associated with mortality rate. In our paper, because mortality is not modelled to be in lined with the later evidence of Omicron variant, we assume morbidity could also induce non-pecuniary loss as people fall ill.

Assumption 1-(1). The preventive expenditure could be interpreted as any economic cost that reduces the chance of infection e.g. wearing mask, conducting regular lateral flow test, social distancing etc.. Now, to give a bird's eye view to the contact mechanism, lets' simply assume two individuals i, j with different preventive expenditures, say, $m^i_{\mathcal{P}}$ and $m^j_{\mathcal{P}}$. They would contact with each other with at a rate of the product $\alpha(m^i_{\mathcal{P}})\alpha(m^j_{\mathcal{P}})$.

The infection takes place once susceptible individual contacts with any infective individuals in the economy. Therefore, in the case of two individuals, susceptible i is infected with probability $\alpha(m^i_{\mathcal{P}})\alpha(m^j_{\mathcal{P}})$ if j carries the disease, with probability 0 if j doesn't. That is, the infection probability for i would be $\alpha(m^i_{\mathcal{P}})\alpha(m^j_{\mathcal{P}})\mathbb{1}_{(h^j=\mathcal{I})}$. We denote the latter two terms $\alpha(m_{\mathcal{P}})\mathbb{1}_{(h=\mathcal{I})}$ as the infectious contact rate. Thus, if individuals interact with all members in the economy, we can express the infected probability for a susceptible individual with $m_{\mathcal{P}}$ unit of preventive expenditure as

$$\lambda = \alpha(m_{\mathcal{P}})\zeta. \tag{4}$$

The first component $\alpha(m_{\mathcal{P}})$ is a susceptible individual's own contact rate. The second part ζ is the average infectious contact rate in the economy.[15] We further assume the population is large enough that individual's choice on own preventive expenditure has negligible effect on the social average variable. That is, ζ is perceived and taken as given in the individual maximization problem.

For recovery process, we assume recovery rate for the infective group is increasing with treatment expenditure, i.e. $\gamma(m_{\mathcal{T}})$ in 1-(2).

Assumption 1. *(1) The contact rate for individual is a decreasing function of preventive expenditure that*

$$\alpha(m_{\mathcal{P}}) : \mathbb{R}_+ \to \mathbb{R}_+ \quad with \quad \alpha' < 0; \alpha'' > 0, \alpha(0) = \bar{\alpha}; \alpha(\infty) = \underline{\alpha}. \tag{5}$$

(2) The recovery rate is an increasing function of treatment expenditure that

$$\gamma(m_{\mathcal{T}}) : \mathbb{R}_+ \to \mathbb{R}_+ \quad with \quad \gamma' > 0; \gamma'' < 0; \gamma(0) = \underline{\gamma}; \gamma(\infty) = \bar{\gamma}. \tag{6}$$

With these assumptions, we can represent the motion of health status in Figure 4 by a hypothetical transitional matrix as Equation 7. We mark that the transition

[15] This assumption is inspired by heterogeneous age group model in Hethcote (2000). The expression is a matching process between the susceptible and infected group. To help understand the interpretation of this expression, we assume discrete uniform distribution of N population. If individual i meets everyone in the economy, the probability of contacting with an infected people would be $1/N \sum_j \alpha(m^i_{\mathcal{P}})\alpha(m^j_{\mathcal{P}})\mathbb{1}_{(h^j=\mathcal{I})} = \alpha(m^i_{\mathcal{P}})[1/N \sum_j \alpha(m^j_{\mathcal{P}})\mathbb{1}_{(h^j=\mathcal{I})}]$. The second term would be the average infective contact rate. The continuous distribution case is an analogy integrating by density distribution function.

matrix is just for illustration. In our continuous-time setup, the transition will be a Poisson process, with the intensity in the matrix.

$$\pi(h'|h) = \begin{bmatrix} 1-\alpha(m_{\mathcal{P}})\zeta & \alpha(m_{\mathcal{P}})\zeta & 0 \\ 0 & 1-\gamma(m_{\mathcal{T}}) & \gamma(m_{\mathcal{T}}) \\ \psi & 0 & 1-\psi \end{bmatrix}_{3\times 3}. \tag{7}$$

Firms. The production landscape is a standard perfect competitive market. The representative firm has Cobb-Douglas production functions, giving $Y = AK^{\beta}L^{1-\beta}$, where Y is output, K is aggregate capital depreciated at rate δ, L is labour, A is technology and β is capital income share with $0 < \beta < 1$. The price of output is normalized to 1, and the profit maximization problem is given as:

$$\sup_{\{K,L\}} \Pi = AK^{\beta}L^{1-\beta} - rK - wL - \delta K. \tag{8}$$

The equilibrium interest rate and wage are given by the F.O.C. as Equation 9.

$$\begin{aligned} w &= A(1-\beta)\left(\frac{K}{L}\right)^{\beta}, \\ r &= A\beta\left(\frac{K}{L}\right)^{\beta-1} - \delta. \end{aligned} \tag{9}$$

B. Equilibrium

There are two types of equilibrium conditions we need to consider. One type of equilibrium conditions is the standard market clearing conditions. The demand for capital from firms equals to the supply of asset from individuals. Similarly, in labour market, the demand for labour from firms equals to the supply of labour from individuals. That is, aggregate capital and aggregate labour satisfy

$$\begin{aligned} K &= \int ag(a,h)d\mu, \\ L &= \int z(h)g(a,h)d\mu. \end{aligned} \tag{10}$$

The other type of equilibrium conditions is the epidemiological equilibrium condition. As noted in Equation 4, the infection probability depends on both individual's own contact function and his perception about the average infectious contact rate ζ. In equilibrium, the individuals' perception about the average infectious contact rate is in fact the true value.

$$\zeta = \int \alpha(m_{\mathcal{P}})\mathbb{1}_{(h=\mathcal{I})}g(a,h)d\mu. \tag{11}$$

The formulation of our heterogeneous agent model is recognized to be a Mean-Field-Game (MFG) in Mathematics. Mean-field theory studies the strategic decision making by small interacting agents in very large populations. The paradigm of MFG is initiated by Lasry and Lions (2007); Huang, Malhamé and Caines (2006). When the number of players goes to infinity, the interaction between individuals is mean-field-type that he or she observes only averages of functions of the private states of the other players. Individuals' effect on distribution or aggregate variables is trivial. In the Nash equilibrium of the game, given a distribution, individual would not deviate from their optimal policy. Meanwhile, the distribution is the probability behaviour under individuals' optimal behaviour.

In our model, the joint distribution of individual states $g(a, h)$ formulates the aggregate variables and price. Individuals take the price as given when optimizing their life-time utility. In equilibrium, the price is exactly the aggregate outcome from individuals' optimal policy. In summary, we are looking for the equilibrium defined as follows.

Definition 1. *(Nash Equilibrium in MFG) Choice variables* $\{c^*, m_{\mathcal{P}}^*, m_{\mathcal{T}}^*\}$ *and distribution* $g^*(a, h)$ *satisfying*

(1) $\{c^*, m_{\mathcal{P}}^*, m_{\mathcal{T}}^*\}$ *solves the optimization problem Equation 3 subjected to the motion of wealth and health (Equation 2 and Equation 7), given the price* (w^*, r^*) *and social infective contact rate* ζ^*.

(2) Distribution $g^*(a, h)$ *is the outcome of optimal policy* $\{c^*, m_{\mathcal{P}}^*, m_{\mathcal{T}}^*\}$.

(3) Distribution $g^*(a, h)$ *generates aggregate variables* $\{w^*, r^*, \zeta^*\}$ *using Equation 9, Equation 10 and Equation 11.*

IV. Computation and Calibration

A. HJB-KFP PDEs

We follow the Partial Differential Equation (PDE) viewpoint[16] in searching the equilibrium (see e.g. Achdou et al. (2013, 2020); Lauriere (2021); Achdou et al. (2022) etc.). The dynamic programming representation of the problem consists of two PDEs: (1) Hamiltonian-Jacobian-Bellman Equation (HJB); (2) Kolmogorov Forward Equation (or Fokker-Planck Equation, KFP).

Denote the value of maximization problem Equation 3 as $v(a, h)$. The value function satisfies the HJB as

[16] Another approach is the Stochastic Differential Equation (SDE) view point (see e.g. Carmona, Delarue et al. (2018); Cardaliaguet et al. (2019) etc.)

$$\begin{aligned}\rho v(a,h) = \sup_{\{c,m_{\mathcal{P}},m_{\mathcal{T}}\}} & u(c) - \chi \mathbb{1}_{(h=\mathcal{I})} + \partial_a v(a,h)[wz(h) + ra - c - m_{\mathcal{P}} - m_{\mathcal{T}}] \\ & + \Lambda^{h'}(m_{\mathcal{P}}, m_{\mathcal{T}}, h)[v(a,h') - v(a,h)] \\ & + \nu_{\mathcal{P}} m_{\mathcal{P}} + \nu_{\mathcal{T}} m_{\mathcal{T}} + \partial_t v(a,h).\end{aligned}$$

In this expression, the discounted value equals the objective function plus the marginal value change brought by the motion of wealth and health. $\Lambda^{h'}(m_{\mathcal{P}}, m_{\mathcal{T}}, h)$ is the Poisson rate of switching from status h to h' (entries in the transitional matrix Equation 7) under the continuous time setup. This term departs from Aiyagari (1994) and Achdou et al. (2022) that the motion of the discrete state h is endogenized by health expenditure. $\nu_{\mathcal{P}}$ and $\nu_{\mathcal{T}}$ are the slack variables for the constraints $m_{\mathcal{P}}, m_{\mathcal{T}} \geq 0$. The FOCs for the maximization problem are listed as follows.

$$\begin{aligned} c: \quad & u'(c) - \partial_a v = 0, \\ m_{\mathcal{P}}: \quad & -\partial_a v + \frac{\partial \Lambda^{h'}(m_{\mathcal{P}}, m_{\mathcal{T}}, h)}{\partial m_{\mathcal{P}}}[v(a,h') - v(a,h)] + \nu_{\mathcal{P}} = 0, \\ m_{\mathcal{T}}: \quad & -\partial_a v + \frac{\partial \Lambda^{h'}(m_{\mathcal{P}}, m_{\mathcal{T}}, h)}{\partial m_{\mathcal{T}}}[v(a,h') - v(a,h)] + \nu_{\mathcal{T}} = 0, \\ & \nu_{\mathcal{P}} m_{\mathcal{P}} = \nu_{\mathcal{T}} m_{\mathcal{T}} = 0. \end{aligned} \tag{12}$$

The first FOC derives the optimal consumption policy that

$$c^* = u'^{-1}(\partial_a v). \tag{13}$$

For the second FOC, notice from Equation 7 that the infective probability is a function of preventive expenditure only for the susceptible group $\mathcal{S}$. For the rest of the health groups, preventive expenditure will not have impact on their transition probability. That is, for group $h = \{\mathcal{I}, \mathcal{R}\}$, we have $\partial_{m_{\mathcal{P}}} \Lambda^{h'}(m_{\mathcal{P}}, m_{\mathcal{T}}, h) = 0$. This implies a non-zero slack variable that $\nu_{\mathcal{P}} = \partial_a v > 0$. Therefore, the optimal preventive health policy for the infective and recovered groups is spending nothing.

$$m^*_{\mathcal{P}}(a, \mathcal{I}) = m^*_{\mathcal{P}}(a, \mathcal{R}) = 0. \tag{14}$$

For the susceptible group, we have $\Lambda^{\mathcal{I}}(m_{\mathcal{P}}, m_{\mathcal{T}}, \mathcal{S}) = \alpha(m_{\mathcal{P}})\zeta$. Thus, the optimal preventive expenditure satisfies the F.O.C. $-\partial_a v + \alpha'(m_{\mathcal{P}})\zeta[v(a,h') - v(a,h)] = 0$. Similar procedure could be gone through for the treatment expenditure $m_{\mathcal{T}}$. To round up, the optimal health policy can be derived as Equation 15

$$
\begin{aligned}
m^*_{\mathcal{P}} &= \begin{cases} 0; & h = \{\mathcal{I}, \mathcal{R}\} \\ \alpha'^{-1}\left(\dfrac{\partial_a v(a,\mathcal{S})}{\zeta[v(a,\mathcal{I}) - v(a,\mathcal{S})]}\right); & h = \mathcal{S}. \end{cases} \\
m^*_{\mathcal{T}} &= \begin{cases} 0; & h = \{\mathcal{S}, \mathcal{R}\} \\ \gamma'^{-1}\left(\dfrac{\partial_a v(a,\mathcal{I})}{v(a,\mathcal{R}) - v(a,\mathcal{I})}\right); & h = \mathcal{I}. \end{cases}
\end{aligned} \tag{15}
$$

The first expression implies that infected people would spend nothing on preventive expenditure to reduce their infectious contact rate, although it could lower the risk of transmitting disease to others. Such an externality is the nature of decentralized economy, where we abstract from altruism that individuals derive utility only from their own consumption and health status. Hence, individuals in group $\mathcal{I}$ would substitute all their preventive expenditure for treatment to shorten the infected period. Similarly, recovered people would not spend on prevention either, as they know they carry immunity.

For the susceptible group, the preventive expenditure is an increasing function of the value loss of being infected $v(a,\mathcal{I}) - v(a,\mathcal{S})$. Susceptible individuals would spend more on prevention if infection brings larger value loss for them. Therefore, one could expect that the equilibrium preventive expenditure would be larger with greater disease-induced productivity loss or direct punishment χ. The treatment expenditure is an increasing function of the value gain of recovery $v(a,\mathcal{R}) - v(a,\mathcal{I})$. Thus, people spend less on treatment if the symptoms get milder or there are no long-run health effect from infection.

The second part for the PDE system for the MFG is the Kolmogolov Forward Equation (KFP).

$$
\begin{aligned}
\frac{\partial g(a,h)}{\partial t} = &- \frac{\partial}{\partial a}[s(a,h,m_{\mathcal{P}},m_{\mathcal{T}})g(a,h)] \\
&- \Lambda^{h'}(m_{\mathcal{P}},m_{\mathcal{T}},h)g(a,h) + \Lambda^{h}(m_{\mathcal{P}},m_{\mathcal{T}},h'')g(a,h'').
\end{aligned} \tag{16}
$$

where $s(a,h,m_{\mathcal{P}},m_{\mathcal{T}})$ is the savings defined as $\dot{a}$ in Equation 2. The KFP evaluates the motion of the joint distribution of wealth and health. The change of the joint distribution over time $\partial_t g(a,h)$ could be broken into to the marginal change of a and h. Specifically, the first component is the flow of population at margin a. The last two components can be interpreted as the population flow-out of health state h to the next state, and flow-in from the previous state.

Under the Finite-Differencing-Scheme (Achdou et al., 2013), we can rewrite the HJB-KF system using matrix notation as Equation 17 and Equation 18, where $\mathbf{V}$ and $\mathbf{g}$ is the matrix for value and distribution. $\mathcal{A}$ is a sparse matrix for HJB equation with its adjoint $\mathcal{A}^*$ when the optimal conditions Equation 13, Equation 15 holds.

We show the construction for these matrices in Appendix D.

$$\text{(HJB)} \quad \rho \mathbf{V} = u(\mathbf{V}) + \mathcal{A}\mathbf{V} + \partial_t \mathbf{V}, \tag{17}$$

$$\text{(KFP)} \quad \partial_t \mathbf{g} = \mathcal{A}^* \mathbf{g}. \tag{18}$$

We can further represent the equilibrium conditions Equation 9 to Equation 11 by function $\mathcal{F}(\mathbf{g}_t) = \mathbf{0}$. Therefore, to find the Nash Equilibrium defined in Definition 1 is to solve the dynamic programming problem defined as follow.

Definition 2. *(Dynamic Programming for Nash Equilibrium in MFG) Find sequence of distribution* $\{\mathbf{g}^*\}_t$ *and value function* $\{\mathbf{V}^*\}_t$ *such that*

(1) $\{\mathbf{V}^*\}_t$ *solves the HJB Equation 17.*

(2) Motion of distribution $\{\mathbf{g}^*\}_t$ *satisfies KFP Equation 18.*

(3) $\{\mathbf{g}^*\}_t$ *clears the market that* $\mathcal{F}(\mathbf{g}^*) = \mathbf{0}$.

B. Parameterization

We calibrate the model in seasonal frequency. The parameterization for economy side of the model is standard. We let $\sigma = 2$ in the CES utility function and subjective discount rate $\rho = 0.0138$. The production function is Cobb-Douglas production with productivity $A = 1$, capital income share $\beta = 0.36$ and depreciation rate $\delta = 0.05$. For the direct disutility of being infected, we calibrate $\chi = 0.3$ in the baseline and vary this parameter in the comparative study.[17] We choose this value is to match the empirical observation that wealthier are more preventive to the disease with higher health expenditure. As we will show later, the health expenditure elasticity of income lies between 5% to 40% under this calibration. However, if we assume no non-pecuniary punishment with $\chi = 0$, health expenditure is inelastic to income.

The main goal for our parameterization is to calibrate the epidemiological part of the model to match the latter evidence of Covid-19. We firstly handle the key functional form of contract rate and recovery rate. We follow Goenka, Liu and Nguyen (2021) that the contact function and recovery function are assumed to be

$$\begin{aligned} \alpha(m_{\mathcal{P}}) &= \epsilon_0 (m_{\mathcal{P}} + \epsilon_2)^{\epsilon_1}, \\ \gamma(m_{\mathcal{T}}) &= \bar{\gamma} - \eta_0 (m_{\mathcal{T}} + \eta_2)^{\eta_1}. \end{aligned} \tag{19}$$

[17] In models with disease-induced mortality, e.g. Acemoglu et al. (2021); Alvarez, Argente and Lippi (2021), the weight χ is associated with the Value of Statistical Life (VoSL). The magnitude of χ is calibrated to 30-70 times output. Our paper is abstracted from mortality. We therefore calibrate this parameter such that the health expenditure elasticity of income be in a reasonable range.

In this expression of $\alpha(m_{\mathcal{P}})$, ϵ_1 is the maximum contact rate elasticity of preventive expenditure. We assume unit elasticity that $\epsilon_1 = -1$. We let $\epsilon_0 = 0.18, \epsilon_2 = 0.005$ so that $\alpha(0) = 36$. This implies individuals would meet 14 people on average per day if the prevention expenditure is 0 for everyone.[18] We notice that it is hard to find empirical evidence to calibrate the contact function. However, as will summarize later, the outcome of the parameters we choose match the basic reproduction number R_0 of the Omicron variant.

For the recovery function $\gamma(m_{\mathcal{T}})$, early evidence on alpha and pre-alpha variants shows the symptoms clearance duration is around 6 to 20 days [19](Beigel et al., 2020; Lechien et al., 2020). For the Delta variant, Hakki et al. (2022) finds 93% of samples shed viral RNA for over 7 days after symptom onset. And it takes around 14.6 days to clear the viral. Latter evidence on Omicron variants shows a shorter infective duration. The literature review by UK Health Security Agency (2023) shows that most studies report a clearance time around 7 to 15 days. Hence, we assume $\eta_1 = -1, \eta_2 = 0.005$ in analogy to the contact function. We calibrate $\bar{\gamma} = 12.85, \eta_0 = 0.034$ such that the recovery duration is bounded between 7 and 15 days.

For the rate of losing immunity ψ, Cagigi et al. (2021) shows the IgG levels had significant waning from 3 to 8 months[20] in patients who recovered from moderate and severe disease. Similar evidence of waning immunity could be found on antibody by vaccination. Gilboa et al. (2022) finds the IgG level decays at 2.26% (1.32%) per day for second dose (third dose) of BNT162b2 vaccine. The microneutralization assays begin to decay within 3 months after the third dose of vaccine. However, the data shows the most reinfection happens among different variants of virus (ONS, 2023a). The antibody could be even less effective against different variants. Therefore, considering that disease mutation is abstracted in our model, we assume a slightly higher rate of reinfection. We calibrate $\psi = 3/5$ such that the immunity only lasts for 5 months.

To further link our calibration with the empirical data of Covid-19, Table I summarizes the epidemiological characteristics of the model in stationary equilibrium and the empirical data. We roughly calculate the basic reproduction number R_0. Appendix C shows how R_0 is obtained in our heterogeneous agent model. We find

[18]The Poisson intensity of contacting others, regardless of the health status, would be $\alpha\bar{\alpha}$, where $\bar{\alpha}$ is the average contact rate in the economy. If all individuals spend 0 preventive expenditure, the contact rate is homogeneous that $\alpha(0)\bar{\alpha}(0) = \alpha(0)^2 = 1296$. This implies 1296 contacts on average per period. As our model is of seasonal frequency, the calibration implies 14.4 contacts per day.

[19]The double blind trial of Remdesivir (Beigel et al., 2020) finds the median time to recovery for patients not receiving oxygen, receiving low-flow oxygen, high-flow oxygen are around 6, 9 and 20. Patients with severe symptoms receiving ECMO or mechanical ventilation recover in around 1 month.

[20]Although systemic IgG levels were durable for up to 8 months, airway IgG and IgA declined significantly within 3 months.

our model produces $R_0^{(ss)} = 9.23$ at the baseline stationary equilibrium. We find this value matches the empirical studies on Omicron variant. Liu and Rocklöv (2022) summarize estimated R_0 of Omicron variants in the recent studies. The Omicron variant has an average basic reproduction number of 9.5 and a range from 5.5 to 24. For population composition, our calibration yields a 4.1% of social infection rate at stationary equilibrium. The value matches the UK observation after 2023 when the pandemic is stabilized. However, it is hard to find statistics on recovery rate. We find the recovered group in our calibration is close to proportion of adults with antibodies over 800 ng/ml.[21] We will vary the parameters in the coming section of comparative studies.

Table I – Calibration

	Model		Data
	Mean	Median	
Basic Rep. Num. R_0	9.236		9.5 ave., range 5.5-24
Days to Infection	19.183	18.833	-
Days to Recover	7.241	7.244	around 7 to 15
Days to Lose Immunity	150		around 90 to 240
Fraction $\mathcal{S}$	10.8%		-
Fraction $\mathcal{I}$	4.1%		2%-5% after 2023 (UK)
Fraction $\mathcal{R}$	85.1%		77%-80% Feb 2023 (UK)

Notes: (a) Data source: R_0 Liu and Rocklöv (2022) etc.; Days to recover UK Health Security Agency (2023); Days to lose immunity Cagigi et al. (2021); Gilboa et al. (2022) etc.; UK data ONS (2023a). (b) The data of recovery population is proxied by fraction of population with antibody more than 800 ng/ml

V. Stationary Equilibrium

A. Baseline Results

In the stationary equilibrium, the value function and distribution are not changing overtime. That is, $\partial_t \mathbf{V} = 0, \partial_t \bar{\mathbf{g}} = 0$, where $\bar{\mathbf{g}}$ is the stationary distribution. Hence, we can re-write conditions in Definition 2. A fixed-point iteration algorithm is applied to solve the system. We present the detail in Appendix E.

$$\begin{aligned} \rho \mathbf{V} &= u(\mathbf{V}) + \mathcal{A}\mathbf{V}, \\ \mathbf{0} &= \mathcal{A}^* \bar{\mathbf{g}}, \\ \mathcal{F}(\bar{\mathbf{g}}) &= 0. \end{aligned} \tag{20}$$

[21] ONS (2023b) estimates that 800 ng/ml is the highest level which can produce a historic back-series and enables enhanced monitoring of antibody levels and waning.

In the baseline model, we choose $\chi = 0.3$. Figure 5 presents the stationary equilibrium of the baseline model. In each panel, the horizontal is wealth and curves of different colours represent different health status. The health expenditure is increasing with individual wealth.[22] While we find the main part of the health policy functions being close to linear, the extreme poor individuals near the borrowing constraint behave differently. We observe that they substitute treatment health expenditure and consumption for prevention, with an upward turn in the preventive policy function. This is because their assets cannot go further down when getting close to the borrowing constraint. Individuals would keep their preventive expenditure to be high enough to prevent further income loss from future infection. When boundary individuals accumulate assets, that is, a moving marginally away from the borrowing constraint, they would substitute prevention expenditure for treatment expenditure.

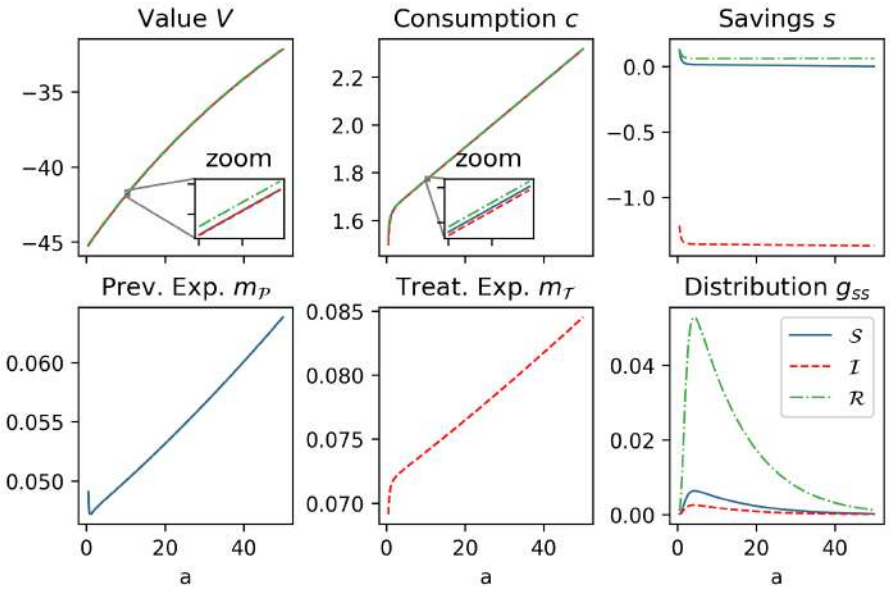
The first two Panels in the first row of Figure 5 plots the value and consumption policy functions for each health status. Qualitatively similar to the health policies, the value and consumption also show momentum on wealth. When it comes the comparison among different health status, we observe the value and consumption for recovered individuals are higher than the other two types in level. This is because they gain temporal immunity after went through infection. The recovered people hence don't need to spend on either prevention or treatment as implied by the F.O.C. we derived in the previous section. Infected individuals have the lowest value and consumption policy among health groups because they are incapacitated with the lowest labour income, and they also spend extra money on treatment.

The disparity in optimal policies generates heterogeneity in savings and wealth. For savings, as shown in the last Panel in the first row of Figure 5, we observe that the infected individuals consume their assets with a negative saving rate, while the other two types of individuals have higher savings. The last panel in the second row of Figure 5 plots the stationary wealth distribution for each epidemiological compartments. We find the wealth distribution in stationary equilibrium is positive skewed for all health statuses. We also plot the wealth share in each percentile (Lorenz Curve) in Appendix Figure A2. The Gini Index for wealth in our model is around 0.412, which is much lower than the empirical observations by survey data (around 0.7 for UK; around 0.85 for US.) (CreditSuisse, 2023). This is due to the nature of Aiyagari model that heterogeneity from idiosyncratic shock is not enough to characterize large skewness of wealth distribution.[23]

[22] The wealth effect of health expenditure our model captures is also found in the model by Hall and Jones (2007) and specially, for preventive expenditure, Ozkan (2014). Life expectancy and survival probability are the key mechanism in these literatures. Our model generates similar observations via incapacitation of infection.

[23] see e.g. Krueger, Mitman and Perri (2016); Stachurski and Toda (2019)

Figure 5 – Baseline Model ($\chi = 0.3$)



Notes: (a) This Figure plots the baseline results ($\chi = 0.3$) at the stationary equilibrium. The subfigures plot the value function; policy function for consumption, savings and health expenditure; distribution. (b) The horizontal axis is the individual wealth. The blue, orange and green lines denote health status susceptible $\mathcal{S}$,infective $\mathcal{I}$ and recovered $\mathcal{R}$ respectively.

However, we notice that the wealth Gini index in our model is still higher than the basic Aiyagari model with uninsured shock, which is roughly 0.32.[24] We believe the insurance mechanism of health expenditure worsens the wealth equality. Too see this, we compare the outcomes of different models in Table II. In the second row, we shut down both preventive and treatment expenditure that the idiosyncratic risk is now uninsured. We find the disease is more transmissive with higher basic reproduction number R_0. However, the model with uninsured risk brings lower wealth Gini index (0.365). When we only enable one of the health expenditure, as in the third and fourth rows, the wealth Gini slightly increases, but is still lower than the baseline case. This is because the rich people are unable or less able to migrate the negative impact using more health expenditure. The risk of income loss is identical regardless of individuals' wealth, while being lower for the wealthy in the baseline.

To provide some more insights to the wealth and income inequality, we tabulate the income distribution against wealth and health in Table III. We split the total population by the 1^{st} and 3^{rd} quantile of wealth and income distribution. The entries in the table denote the share of population in each income group. In Panel (a), we can see that the largest component in the low income group is the people below

[24]The last row of Table II shows the Gini index of a standard Aiyagari model, where possible productivity are $\{0.2, 1\}$ in lined with our baseline model. The Poisson intensities are $\{0.5, 0.5\}$.

Table II – Model Comparison

Model	R_0	agg.Capital	agg.Income	Wealth Gini
Baseline	9.236	14.447	1.838	0.412
Exog. Disease	216.0	13.869	1.763	0.365
$m_{\mathcal{P}}$ only	10.427	13.966	1.775	0.37
$m_{\mathcal{T}}$ only	104.282	14.393	1.831	0.407
Aiyagari	-	43.981	2.376	0.319

Notes: This table compares the stationary equilibria of different models. The first row is the baseline model in our paper. The second row we assume the exogenous disease dynamics. The third and forth rows assume that individuals optimize only one health expenditure. The final row shows the results for a standard Aiyagari model of two types of productivity $\{0.2, 1\}$ with the same Poisson intensity.

Table III – Income Distribution in the Baseline Model

(a) Low Income Group

	$\mathcal{I}$	$\mathcal{S}$	$\mathcal{R}$
Low a	0.042	0.089	0.748
Mid a	0.081	0.0	0.0
High a	0.04	0.0	0.0

(b) Middle Income Group

	$\mathcal{I}$	$\mathcal{S}$	$\mathcal{R}$
Low a	0.0	0.007	0.056
Mid a	0.0	0.104	0.833
High a	b	0.0	0.0

(c) High Income Group

	$\mathcal{I}$	$\mathcal{S}$	$\mathcal{R}$
Low a	0.0	0.0	0.0
Mid a	0.0	0.005	0.039
High a	b	0.117	0.839

Notes: 0<b<1e-5

the first quantile of the wealth distribution, which takes up more than 80%. Some people of the wealthiest 25% could also fall into low income when they are infected and lose labour income. They only share around 4% in the group. Panel (b) and (c) show the distribution for middle and high income group respectively. The middle and high income group mainly consist of those healthy individuals (susceptible and recovered). Only few infected people of the wealthiest 25% could remain middle or high income. This is due to their low labour income share.[25] The impact from disease-induced labour loss is insignificant when capital income is high enough.

B. Income Elasticity of Health Expenditure

Figure 5 shows the policy function for health expenditure. To further disentangle the implication on individual's health policy from wealth heterogeneity, we solve for the income elasticity of health expenditure. Instead of using the instantaneous policy functions directly, we consider the expenditure over a certain period (0 to τ) in the

[25] In Appendix Figure A3, we plot the share of labour income in total income for the infected and uninfected group. The labour income share is decreasing with wealth. When individuals get infected, the labour income share reduce that people rely more on capital income.

stationary equilibrium (Achdou et al., 2022; Laibson, Maxted and Moll, 2022) as follows

$$\begin{aligned} M_{\mathcal{P}}(a_0, h_0) &= \mathbb{E}\left[\int_0^{\tau} m_{\mathcal{P}}(a_t, h_t)dt \middle| a_0, h_0\right], \\ M_{\mathcal{T}}(a_0, h_0) &= \mathbb{E}\left[\int_0^{\tau} m_{\mathcal{T}}(a_t, h_t)dt \middle| a_0, h_0\right]. \end{aligned} \tag{21}$$

$M_{\mathcal{P}}$ and $M_{\mathcal{T}}$ sum up the health expenditure from time 0 to τ for individuals starts with state (a_0, h_0). These expressions are important complementations to the instantaneous policy functions as they additionally capture the information at the dimension of time. In our model, health expenditure are restricted to individuals who are infected or susceptible. However, the health status is stochastic. The duration for each health status varies along the wealth space, and thus differs the cumulative health spending in a certain period. This means that whether the riches spend more on health over time in our model is not intuitively explicit. For example, although wealthier individual spend more on prevention and treatment, they consequently experience fewer shots and shorter duration of infection. The second advantage of using expenditure over time is that for most survey data, we mainly observe individuals' time-aggregated variables at some specific frequency.

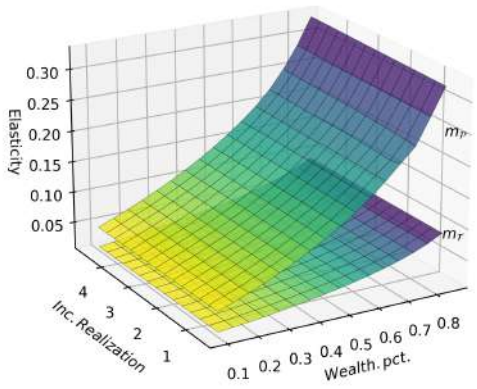
We apply the new measurement to solve for the Income Elasticity of Health Expenditure as Equation 22, where we assume there is a windfall income Δ on individuals' asset. These expressions measure the fractional change on health expenditure over τ periods, if the wealth endowment is increased by $\Delta/a_0 \times 100\%$. In Appendix D, we show how Equation 21 is computed by the Feynman-Kac Formula.

$$\begin{aligned} \varepsilon_{M_h, y} &= \frac{\partial M_h(a_0, h_0)}{\partial y}\frac{y}{M_h} \\ &= \frac{M_h(a_0 + \Delta, h_0) - M_h(a_0, h_0)}{\Delta}\frac{a_0}{M_h(a_0, h_0)} \\ h &= \{\mathcal{P}, \mathcal{T}\} \end{aligned} \tag{22}$$

We set the time period τ to one-year duration and presents the computational results at Figure 6. The surface represents the estimated elasticities corresponding to income realization Δ and wealth percentile. We focus on the baseline case where $\chi = 0.3$. We find that the income elasticity is positive but lower than unity[26] for both types of health expenditure. Preventive expenditure is more elastic than the treatment expenditure, with its surface lying well above. We also find the elasticity is increasing with wealth level. Individuals at the top wealth decile exhibit elasticity around 5 times more than those at the lowest decile. However, we find the estimated

[26]This is consistent with the empirical evidence that health expenditure is a necessary good rather than luxury good (Di Matteo, 2003; Freeman, 2003; Moscone and Tosetti, 2010).

Figure 6 – Health Expenditure Elasticity of Income



Notes: (a) This Figure plots the health expenditure elasticity of income when the health punishment is 0.3. (b) The x-axis denotes the percentiles in the wealth distribution. The y-axis denotes the realization of windfall income Δ.

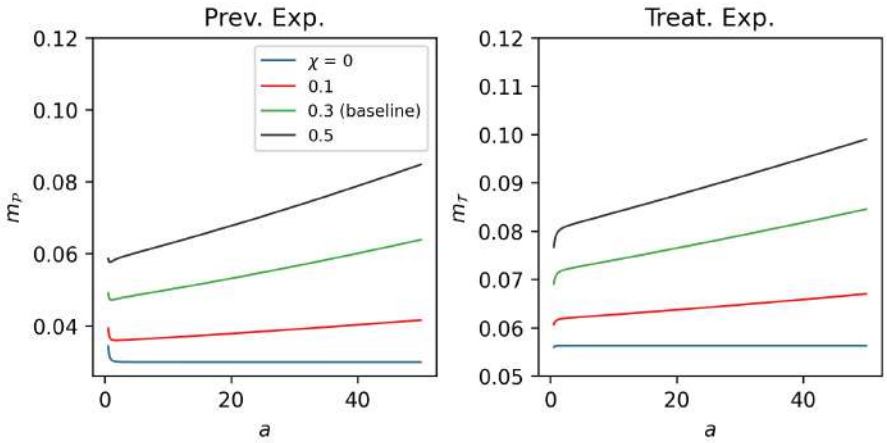
elasticity is not sensitive to the rate of windfall income Δ.

C. Comparative statics

To the study long-run implication of the model, we solve the stationary equilibria under different parameters for comparative statics. Specifically, we conduct the experiment by varying disutility χ. Figure 7 presents health policy function with varying degrees of disutility from being infected. When lowering the level of disutility, health expenditure become less sensitive to wealth. This is because when χ decreases, the extra punishment of being infected is lowered. Therefore, the value of infection $v(a, \mathcal{I})$ increases and changes the denominator in F.O.C. Equation 15. For the special case of $\chi = 0$, the mechanism via direct loss of being ill is shut down. Infection is therefore a pure income shock. However, temporary income drop of being infection is trivial for the wealthier individuals. Hence, we observe very small heterogeneity in health expenditure along the wealth distribution as the blue line. The corresponding income elasticity of health expenditure also gets smaller as shown in Appendix Figure A4.

Table V provides more details on how variables change in response to change in disutility. Panel (a) compares the aggregate variables. Overall, the infection rate for the rich (top 25%) is smaller. When the disutility increase, the gap in infection rate is widened. The richer are getting less likely to be infected, related to the poor. The disparity of infection outcomes among wealth groups matches our the empirical

Figure 7 – Health Policy Function



Notes: The first and second Panel of this Figure plot the health policy functions varying health punishment χ. The slope is higher when the punishment increases.

observation in UK. ONS (2022a) documents a statistically significant 1.8% gap on infection risk[27], 2.8% gap on self-reported hospital admission rate between people living in the most and least deprived area.

The infection rate gap generated in the model is due to the optimal policy that the rich spend more on prevention, as shown in Panel (b) and in lined with our observation in Figure 7. Price (w, r) do not changed too much. This is because aggregate capital and labour move at the same direction. Capital is increased. Labour is also increased because of less infection. The influence on pricing is then cancelled out in the long run. However, as we will show in the later session, the prices would change in the short run when converging to the stationary equilibrium or being subjected to shocks.

The next few blocks in the Panel (a) compare the wealth and income inequality. We observe the Gini index for wealth and inequality moves in opposite directions as punishment χ increases. On the one hand, income equality is improved with higher χ. This is due to the property that infection harms labour productivity such that the majority of infected individuals has the instantaneous income under the 25% percentile of the income distribution (see Table III). Therefore, the reduction of overall infection rate implies smaller size of low income group, and thus a more equal income distribution. On the other hand, the wealth inequality becomes more profound with the wealth Gini index and the wealth shared by the wealthiest 25% increase slightly.

We provide a possible explanation in Table VII, where we use Feynman-Kac

[27]For Delta variant, the difference was 1.9pp (most deprived: 7.3%, 95% CI: 6.9 to 7.7; least deprived: 5.4%, 95% CI: 5.2 to 5.6). While for Omicron the difference was 1.8pp (most deprived: 6.7%, 95% CI: 6.2 to 7.1; least deprived: 4.9%, 95% CI: 4.6 to 5.1)

Table V – Comparative Study

(a) Aggregate Variables

χ	0	**0.3**	0.5
Infection Rate			
Aggregate	4.344	4.107	3.97
Bottom 25%	4.444	4.235	4.115
Top 25%	4.307	4.023	3.862
diff.	-0.137	-0.212	-0.253
Capital	14.418	14.447	14.463
Prices			
Wage Rate	1.694	1.694	1.694
Interest Rate	0.014	0.014	0.014
Inequality			
Wealth Gini	0.41	0.412	0.423
Income Gini	0.072	0.071	0.07
Wealth Share			
Bottom 25%	6.72	6.67	6.49
Top 25%	52.89	53.03	54.08
diff.	46.18	46.36	47.59

(b) Contol Variables

χ	0	**0.3**	0.5
Consumption			
Aggregate	1.83	1.829	1.827
Bottom 25%	1.682	1.681	1.678
Top 25%	2.052	2.053	2.059
diff.	0.37	0.372	0.38
Preventive Exp.			
Aggregate	0.03	0.052	0.066
Bottom 25%	0.03	0.048	0.06
Top 25%	0.03	0.057	0.074
diff.	-0.0	0.009	0.015
Treatment Exp.			
Aggregate	0.056	0.075	0.085
Bottom 25%	0.056	0.072	0.081
Top 25%	0.056	0.079	0.092
diff.	0.0	0.007	0.01

Notes: (a) This table compares the stationary results of models with different health punishment $\chi \in \{0, 0.3, 0.5\}$, where $\chi = 0.3$ is used as the baseline. For the special case $\chi = 0$, infectious disease is a pure income shock. (b) Panel (a) compares the aggregate outcome of price, inequality measures, and infection rate. Panel (b) compares the optimal controls at bottom and top wealth quantile.

Formula[28] to calculate the expected income and savings over a 3-years window. In the first block, we find that expected income increases for both poor and rich, while their income gap is widened when χ goes up. This is because the richer spend more on prevention. Hence, they are less likely to be infected and thus experience smaller productivity loss in a given period. When breaking the income into categories, we find both labour and capital income gap expand.

The change of income stream affects the savings choice. When it comes to the expected savings in the second block of Table VII, we find the savings for the poor is generally higher. As χ increases, savings is reduced for the poor, but increased for the rich. The saving gap between the poor and rich gets smaller. Therefore, the richer accumulate more capital and implies a more unequal wealth distribution.

We also implement the same exercise by varying the infectiousness of the disease. Appendix B Table B2 presents the corresponding results where we change the coefficient ϵ_0 in the contact rate function. In the Appendix B Table B5 and Table B8, we conduct another two experiments by raising the effectivity of health expenditure. Specifically, we increase ϵ_1 and η_1 to allow more curvature on $\alpha(m_{\mathcal{P}}), \gamma(m_{\mathcal{T}})$.

[28] We would like to calculate the expected variables (income, expenditure, savings) over a period τ for each individual with state (a_0, h_0), i.e. $\mathbb{E}[\int_0^\tau y(a_t, h_t)dt | a_0, h_0]$.

Table VII – Comparative Study (cont.)

χ	0	**0.3**	0.5
Expected Income in 3-yr duration			
Bottom 25%	20.5	20.532	20.534
Top 25%	25.0	25.07	25.204
diff.	4.5	4.539	4.669
Labour Income diff.	0.0	0.013	0.02
Capital Income diff.	4.461	4.488	4.612
Expected Savings in 3-yr duration			
Bottom 25%	2.793	2.701	2.755
Top 25%	-1.029	-0.991	-0.916
diff.	-3.822	-3.692	-3.672

Notes: This table uses the Feynman-Kac formula to calculate the expected value within a 3-years duration. We compare the expected income and savings for models with different punishment χ.

VI. Transitional Dynamics

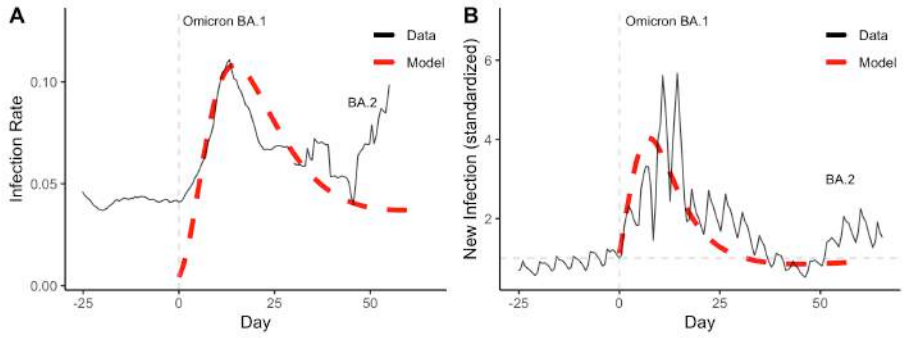
In this section, we move on from the discussion on the long-run stationary equilibrium to the short-run dynamics. Specifically, we consider two practices on the transitional dynamics of the model: (1) Given the arbitrary initial distribution of individual states, solve for the convergence path towards the stationary equilibrium; (2) Given that the economy is initially at the stationary equilibrium, impose a temporary unanticipated shock (MIT shock) to the model.

A. Transition Path

We would like to find a full dynamic sequence of distribution $\{g(a,h)\}_t$ and value function $\{v(a,h)\}_t$ satisfying the Nash equilibria in Definition 1 for $t \in [0,\infty)$. To give a bird's eye view, we solve the HJB (Equation 17) backwards knowing the stationary solution $v^*(a,h)$; solve the KFP (Equation 18) forward given a specific initial distribution. We show the full algorithm in Appendix E.

We inherit the value function from the baseline stationary equilibrium solved in section V.. Then, we construct several initial distributions for individual states. For simplicity, we only disturb the distribution of health status in the following steps. Firstly, we let the marginal distributions of wealth and productivity be identical to that in the stationary equilibrium. Then for the simplex of epidemiological compartment (s_0, i_0, r_0), we assign 0.5% of initial infection rate and vary the initial recovery rate in $\{0, 34\%, 68\%\}$, while the rest are susceptible to disease. We use the different sizes of initial recovered group to give insights on vaccination. One could interpret that higher initial recovery size implies higher vaccination rate. This is because the recovered group consists of individual with temporal immunity to infection, essentially shares the same characteristics with vaccinated group. Vaccination policy in

Figure 8 – Simulation and Empirical Data



Notes: (a) The red dash lines and black solid lines in this Figure shows dynamics of infection in the model and data respectively. We match the empirical data of the wave of Omicron B.A.1 variant in UK. (b) We set the date when infection began to climb as the reference. (c) In Panel A, we match the social infection rate in the model to the data of positive testing rate. (d) In Panel B, we match the new infection rate to the data. We standardize the observation at the reference date to 1.

fact brings individuals from $\mathcal{S}$ to $\mathcal{R}$ without going through group $\mathcal{I}$ (Federico, Ferrari and Torrente, 2022)[29], and hence expands the size of pre-pandemic recovered group.

The third case with 68% of pre-existing recovery matches the fully-vaccinated rate in UK before the outbreak of Omicron B.A.1 variant. Figure 8 shows how the simulation match the empirical data of this wave of outbreak. We set the date when infection case began to increase as the reference. In Panel A, we match the social infection rate to the data of positive testing rate. The simulation well matches the data except that we underestimate the infection rate at the beginning phase of the outbreak. This is because our simulation does not incorporate infection from other variants like Delta variant before the outbreak.[30] Panel B matches population flow from susceptible group to infected group, i.e. $\int \alpha(m_{\mathcal{P}})\zeta \mathbb{1}(h = \mathcal{S})d\mu$, to the data of new infection case. For comparison, we standardize the data at the referenced date to 1. We find the simulated dynamics roughly fits the data.

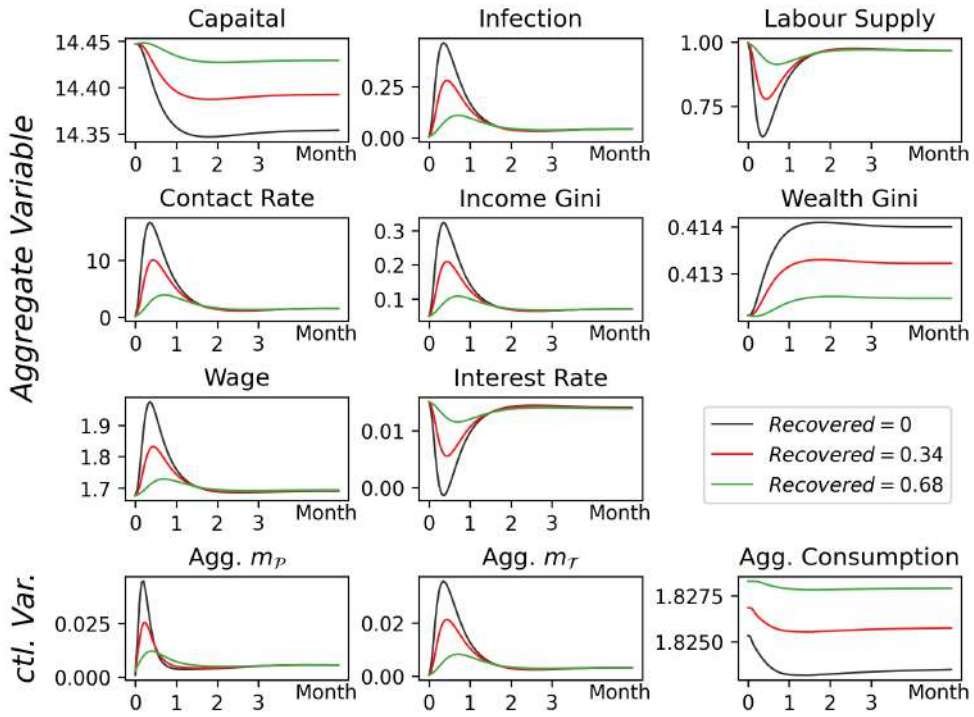
Figure 9 plots the transitional dynamics towards the stationary equilibrium given multiple initial distributions. Throughout the simulation period[31], there would be outbreak that the social infection rate peaks in the first month. We find the in-

[29] Another possible approach is to introduce a new compartment for vaccinated group like e.g. Sun and Yang (2010); Safan, Kretzschmar and Hadeler (2013); Garriga, Manuelli and Sanghi (2022) (another health status in our model). The vaccinated group is mostly similar to the recovered group as they both carries temporary immunity. The only difference is on their reinfection probability. Clinical evidence shows immunity from previous infection lasts longer than from primary series or first booster of vaccination. But the gap is not pronounced (Bobrovitz et al., 2023).

[30] Clinical evidence shows that infection could take place at different variants of virus. Infection cannot provide effective immunity to the new variant (ONS, 2023a).

[31] We let the simulated period to be long enough to ensure convergence to the stationary equilibrium. Here we only plot the first 3 months to visualize the short-run dynamics. Same for other simulations in this session.

Figure 9 – Transitional Dynamics (Convergence Path)



Notes: (a) This Figure plots the transitional dynamics of our baseline model with different initial distributions, where the fraction of pre-existing recovered group are 0 (blue), 16% (yellow) and 32% (green). (b) The first two rows are the dynamics of aggregate variables. The last row shows the control variables aggregated by the time-varying distribution $g_t(a, h)$. (c) The horizontal axis denotes the time (in month) from the simulation.

fection rate converges to the same value for all simulations. However, with smaller initial pre-existing recovered group, the pandemic is worse with larger strike of infection. Correspondingly, labour supply drops significantly during the pandemic and recovers to its steady state. The temporary loss of labour productivity affects wealth accumulation that aggregate capital also sees a subsequent decrease in the first month. Although both capital and labour are negatively shocked, it shows difference in their timing of adjustment. Aggregate labour bounces back very quickly within 2 months as the infected people overcoming illness and obtain immunity. The capital accumulates very slowly and remains in low level for a longer period.

The gap on convergence speed induces factor price fluctuation unlike that in our long-run comparative statics (Table V) where the effect from aggregate K and L offset each other. The change of prices largely follows the motion of infection rate, and is more dynamic with smaller pre-existing recovered group. We observe labour wage is increased while capital interest rate is decreased during the pandemic. The rising wage is due to the shortage of workers, especially in some labour-intensive industries like restaurant, delivery, express etc.. The UK data by ONS (2022b) shows similar pattern during the pandemic. The unit labour cost for firms is significantly

increased in Q1 - Q3 of 2020.[32] The data also reveals a drop in average weekly earning and average output per workers. This observation is also explained by our model that the productivity drop overwhelms the effect of rising labour price in aggregate level. The motion of interest rate also matches our empirical observation during pandemic that average sharecite-keyholder returns declines for most sectors in the first 3 seasons (see e.g. Bradley and Stumpner (2021); Bai et al. (2023) etc.).

One of the key advantages of our HACT setup is that we can track the evolution of the joint distribution $g_t(a, h)$ by the Kolmogolov Forward Equation. This allows us to measure the inequality dynamics by calculating Gini index overtime. In the third row of Figure 9, we find that both income inequality and wealth inequality rise during the outbreak, and are worse for the simulations initiated with smaller recovered group. Specifically, the income Gini index is largely procyclical to the infection rate dynamics. This is because the growing infection rate implies that more individuals would fall into the low-income group. At the same time, those unaffected people with full productivity benefit from the rising labour wage. The growth of wealth Gini index is the deep consequence of expanding income disparity. Differently, we see that the harm on wealth equality is relatively long-lasting as capital accumulation is a slower process to labour supply recovery.

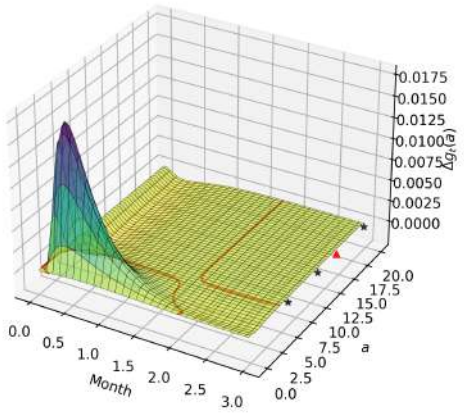
The rising wealth Gini index is induced by the expanding low wealth group. To show this, we break the aggregate measurement of inequality into the change of population density in each wealth level. In Figure 10, we show the evolution of wealth distribution under the simulation with no initial recovery. The surface represents the change of wealth distribution at marginal time, in expression, $\Delta_{dt} g_t(a) = \sum_h g_t(a, h) - \sum_h g_{t-dt}(a, h)$, corresponding to each wealth grid (y-axis) and time after outbreak (x-axis). The blue stars and the red triangle at the y-axis are the 25%, 50%, 75% percentiles and the mean of the initial wealth distribution.[33] The red lines separate areas with different signs of change. We can see that during the first month of the pandemic, population at lower quantile of wealth distribution increases, while population at middle and higher quantile drops. While the economy consists of more poor people and less middle class, the wealth equality would be worsened.

Our model matches the rising wealth Gini qualitatively. However, we find the magnitude for the rise being small, with the change less than 0.01. We believe this is due to the following reasons: (1) The Gini index is naturally small and less dynamic in the Aiyagari-type models, as they could not capture the thick-tail of the wealth distribution. (2) General equilibrium effect is the main cause for this

[32] The largest component of total nominal employment costs is compensation of employees which grew by 11.1% between Quarter 4 2019 and Quarter 4 2021

[33] These percentiles would change during the convergence. We only plot the stationary percentiles here as their changes are sufficiently small.

Figure 10 – Change in Wealth Distribution

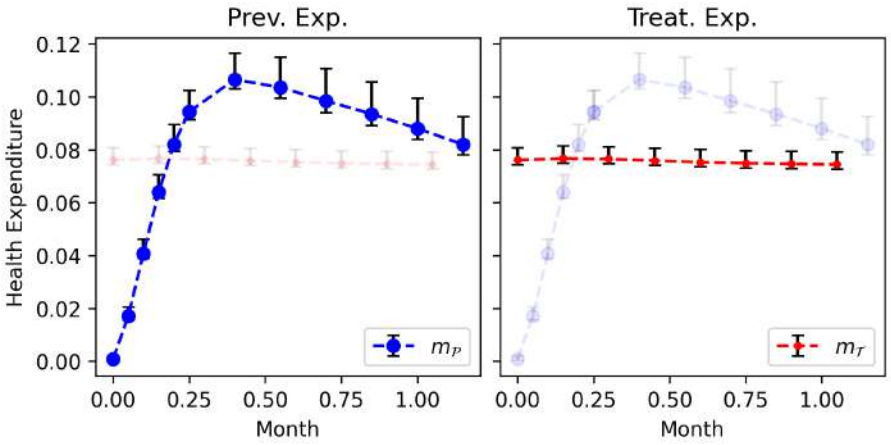


Notes: (a) This Figure shows the change of wealth distribution $\Delta_t g_t(a)$ for the simulation with 16% of pre-existing recovered group, where $g_t(a) = \sum_h g_t(a,h)$. The x-axis is the time after the simulation and the y-axis the wealth level. (b) The blue stars represents the (25%, 0.5%, 0.75%) percentile of wealth distribution. The yellow triangle represents mean of the wealth distribution. (c) The red line is the contour for $\Delta_t g_t(a) = 0$. It separates the increasing and decreasing areas.

outcome. The expanding wealth inequality is mainly induced by the disparity in infection rate among wealth levels, which directly affects their labour income. The assets income disparity among wealth, on the other hand, is reduced because of the decreased interest rate. Therefore, the expansion of income disparity is mitigated by reducing assets return. (3) The lost of productivity by infection is temporary (around one week). One wave of pandemic is not enough to generate substantial change of wealth Gini index in our model, while there exist multiple waves due to either seasonal cause or virus mutation in the empirical data. We note that the virus mutation generate multiple waves of infection could amplify the magnitude of increase of wealth Gini index. To model virus mutation, we simply let part of recovered people become susceptible when the new variant exist. We show the results in Appendix Figure A5. We firstly match the mutation from Omicron B.A.1 to B.A.2 by letting 25% of recovered people be subjected to the new variant. Then we assume no cross-variant immunity in the next few mutations. We find multiple waves of infection is generated and the increase in wealth Gini index from different waves could be stacked up.

We also study the dynamics of optimal health policy. For the aggregate behaviour, in the last row of Figure 9, we average the individual health expenditure by the sequence of joint distributions $g_t(a,h)$. The dynamics of health expenditure

Figure 11 – Change in Health Policy



Notes: (a) This Figure plots the dynamics of health policies. Panel A highlights the dynamics of preventive expenditure ($m_{\mathcal{P}}$). Panel B highlights the treatment expenditure ($m_{\mathcal{T}}$). (b) The connected dots are the health expenditure at 50% of wealth distribution. The lower bars and upper bars are the 10% and 90% of wealth distribution.

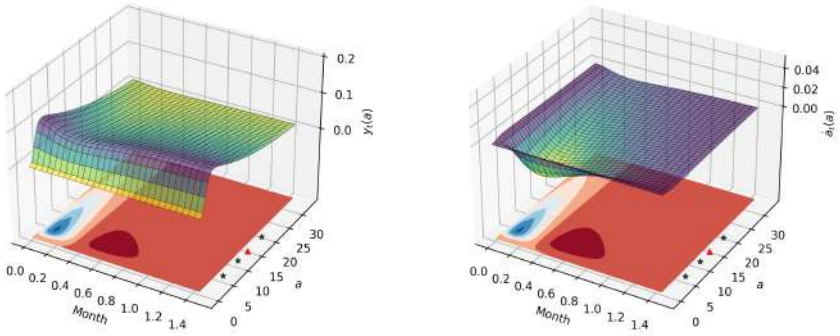
is procyclical with the motion of infection rate in aggregation. We observe substantial increase for both preventive and treatment expenditure in the first month of the pandemic. With slight different between two types of expenditure, the rise of treatment policy is more persistent than preventive one.

When we break the aggregate observations of health expenditure into individual change we find the increase in health expenditure is biased towards higher side of the wealth distribution. As in Figure 11, where we plot the health policy function overtime. Panel A highlights the motion of preventive expenditure. The connected dots, lower and upper bars in the Figure denote the preventive policies at the median, lower 10% and upper 90% of the wealth distribution. We find that the expenditure gap between the rich and poor expands in the first one month of pandemic. The rich people are therefore subjected to lower risk of infection. For treatment policies in Panel B, we still have the observation that the richer have higher expenditure. However, the expenditure gap between the rich and poor does not change too much.

We can now link the transition of distribution (Figure 10) with the motion of optimal policy (Figure 11). The change in wealth is determined by the rate of asset accumulation, that is, savings, which is directly affected by income. In Panel (a) and (b) of Figure 12, the surfaces denote the average income and savings at different wealth level across time.[34] The contour maps at the bottom show the corresponding sign of motion: blue area for decrease; red area for increase. We can see that with decreasing income, the saving rate drops during the first few days of pandemic then slowly goes back. In terms of the change across wealth levels, the decrease in average

[34]The average variables are weighted by distribution $g_t(a,h)$. For example, average value for x with asset a would be $\bar{x}(a) = \sum_h x(a,h)g_t(a,h)$.

Figure 12 – Evolution of Income and Savings in Transition



(a) Income Change **(b)** Savings Change

Notes: (a) This Figure plots the change of income and savings along the wealth distribution. The surface denotes the income and savings. (b) We map the change of the income and savings to the x-y plane. The blue area denotes the time and wealth for decreasing income or savings. The red area denotes the time and wealth for increasing income or savings. The notation of blue star and yellow triangle is inherited from Figure 10.

income and savings is biased towards the lower part of the wealth distribution. This implies that the poor individuals lose more income and spend more net capital on average, and hence get even poorer. Reflected in Figure 10, the population density at lower percentiles increase during the pandemic.

B. The Effect of Income Support Scheme

Support Plan for Infection. We now use our model to explore government's policy. Recall from the empirical evidence in Figure 1 that the wealth Gini index moves differently with or without government subsidy. We add government income support to our framework. We assume that government compensates the infected group by direct transfer. The transfer is lump-sum that government partially cover the lost of labour income due to infection.[35] In practice, let τ to be the transfer per individual[36], individuals' income becomes

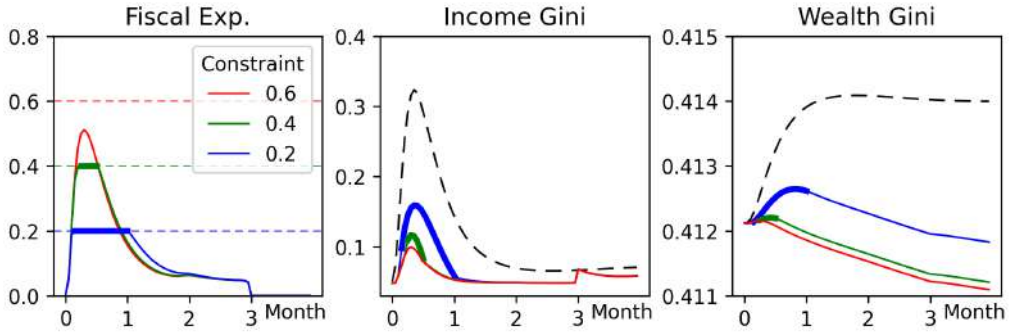
$$ra + wz(h) + \tau \mathbb{1}_{(h=\mathcal{I})}. \tag{23}$$

The income support scheme is financed by government's fiscal expenditure. For

[35]Another way of modelling is by assuming government covers lost of labour income. That is, transfer $\tau = w[z(\mathcal{S}) - z(\mathcal{I})]$. Individuals' total income is hence unrelated to their health status. However, in transitional dynamics, wage rate climbs up during pandemics because of the lack of labour. In this case, government additionally afford the cost of rising price. Thus, transfer could be a more appropriate assumption.

[36]Numerically, we calibrate the transfer τ to be smaller than the wage w, such that the value of infection is always smaller than being healthy.

Figure 13 – Income Support



Notes: (a) In this Figure, we study the policy of targeted income support scheme to infected group. This Figure shows the simulated dynamics for fiscal expenditure, aggregate capital, income and wealth Gini index. (b) The red, green and blue lines represent different trajectories that government faces fiscal constraint of 0.6, 0.4 and 0.2. (c) The black dash lines are the referenced dynamics from the baseline results as in Figure 9. The highlighted parts are the time when the fiscal constraints are binding.

simplicity, we abstract from the source of fiscal expenditure.[37] Instead, we assume that government is subjected to an exogenous fiscal constraint $\mathcal{G}_t \leq \bar{\mathcal{B}}$ where $\mathcal{G}_t$ is the aggregate fiscal expenditure

$$\mathcal{G}_t = \int \tau \mathbb{1}_{(h=\mathcal{I})} g_t(a,h) d\mu. \tag{24}$$

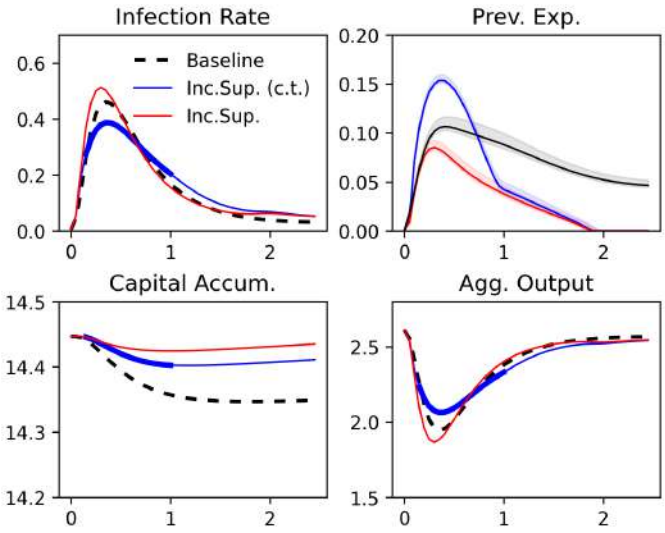
During the period when the fiscal budget constraint is binding, government has to lower the transfer τ such that the budget constraint is satisfied. Furthermore, we assume that income support scheme last for 3 months.

We vary $\bar{\mathcal{B}}$ to study the cross-country difference in fiscal constraint. Figure 9 shows the simulation results. The red lines are the results for economies with the highest fiscal constraint, with its expenditure for income support never hitting the roof. The green and blue lines are the binding economies with different level of constraints. Compared with the simulation without income support, denoted in black dash lines, the spike of income Gini index during pandemic is largely decreased. This is because the income support is targeted at the infected individuals who partially lose their labour income. As a result, the wealth equality with income support is worsened for a short period then improves subsequently.

When it comes to the comparison of fiscal constraints, we find that economies with tighter constraints are less able to relief the increasing income and wealth inequality, especially during the constrained periods when they have to cut the level of subsidy. Our simulations here match with the empirical evidence that richer coun-

[37] We don't model taxation in our model as it is not the main source of income support plan during pandemic. Typically, in UK, the income support package came with tax relief policies during the pandemic.

Figure 14 – Income Support (cont.)



Notes: (a) This Figure shows the infection rate, preventive expenditure, aggregate capital and output under government income support scheme. (b) The red and blue lines denote the simulations under with fiscal constraints of 0.2 and 0.6. The black dash lines are the baseline results without income support. (c) In the top right Panel, the lines denote the preventive expenditure at 50% of the wealth distribution, while the shadowed areas are that at 10%-90% of the wealth distribution.

tries with stronger income support may conversely improve their wealth equality, even though the income support is temporary.

How the income support affect infection rate? We show two scenarios for illustration. In the first Panel of Figure 14, we find the infection rate is increased when the support doesn't hit the fiscal constraint (red line); is decreased when the support is constrained (blue line). This is because the income support affects individual health expenditure in two different ways. First, the direct transfer increases the wealth level that the susceptible could spend more on prevention. Second, subsidizing the infected group mitigates the value loss of being infected and thus discourages health expenditure. Weather people spend more on health depends on the magnitude of the support. To show this, in the top-right Panel of Figure 14, it could be viewed that the preventive expenditure decreases under the unconstrained support scheme and increases under the constrained support.

For capital accumulation in bottom-left Panel of Figure 14, without any counter-intuitive outcome, we find that capital recovers faster under the unconstrained income support, followed by that under the constrained support and without support. We further analyse the aggregate output $Y = AK^{\beta}L^{1-\beta}$ in the bottom-right Panel. The impact of incapacitated labour (infection rate) dominates the motion of output. We find that the constrained income support increases the aggregate output, but the unconstrained support has the opposite effect. Therefore, for the policymaker,

the two-sided effect of income support implies an aggregate trade-off between health and economic equality. The income support by lump-sum transfer could directly improve equality. But it may also induce people to value less on their health and lead to a higher social infection rate, which may reduce aggregate product.

We also discuss other types of support in Appendix Figure A6. Specifically, we evaluate the wealth-targeted support for the poorest 25%[38], i.e. $\tau \mathbb{1}_{(a \leq a_{25\%})}$, and the general support for everyone. They affect equality from different aspects. The infection-targeted support discussed earlier largely improves the income equality. The asset-targeted plan is more effective in improving wealth equality, as it directly rise the income for the poorest 25%. Though asset-targeted and general support plans reduce more wealth Gini index, they put more pressure on fiscal budget as it requires continuous investment.

C. Other Extensions

Weaker Disease. We also explore the convergence path under different calibrations. Specifically, we firstly simulate the dynamics with milder disease. At the later stage of Covid-19 pandemic, the disease mutated towards the variants with lower mortality rate and hospitalized rate. We incorporate this by assuming the infected group are only incapacitated by only 20% while incapacitated by 80% in our previous discussion. The corresponding simulations are shown in Appendix Figure A7. We observe that the infectivity is more severe than the baseline scenario[39] as susceptible people are less precautionary and spend less on health. However, higher infection rate doesn't cut too much aggregate labour supply in this cacse as the disease are weakened. Aggregate capital thus shows smaller decrease and faster recovery. When it comes to inequality measures, both income and wealth Gini index show smaller increase than the baseline because the labour income gap between health group is relatively narrowed.

Long Covid. The second practice in Appendix Figure A8 investigates the effect from Long-Covid, where we lower the productivity of the recovered group that $z(\mathcal{R}) = 0.8$. In this case, individuals' productivity cannot fully recover to the pre-infected stage, thought they still gain immunity. Thus, the value of being infected $v(a, \mathcal{I})$ further reduces as Long Covid cause productivity loss for a longer time. The value for the recovered group $v(a, \mathcal{R})$ also decreases. Consequently, individuals spend more on preventing infection and less on treatment. In aggregation, the peak infection rate is smaller than the baseline. However, as we assume slower

[38] We let the cut-off asset for support to be fixed at the first quantile before the pandemic, i.e. t_0 at the simulation.

[39] We use the simulation with 16% of pre-existing recovered group (yellow lines in Figure 9) as the baseline. Same for the rest of the paper.

productivity recovery process, the decay in aggregate supply of labour and capital is more persistent. Similarly, the income and wealth Gini indices show milder but long-lasting increase.

MIT Shocks. We also study the short-run dynamics of the model by imposing unanticipated temporary shock (MIT shock). We assume the model initiates from the stationary equilibrium in section V.. Then, we impose a temporary increase in infectivity of the disease. As in the first Panel of Appendix Figure A9, the infectivity ϵ_0 is raised by 20%, 40% and 60% respectively then decays back to the initial level. Therefore, the economy will end up with the same stationary equilibrium before the shocks arrived. We find the social infection rate and contact rate increase then adjust back after the shock. Aggregate labour and capital counter the motion of social infection. The magnitude of aggregate capital change is small with only 0.01% to 0.02% deviation from the steady state. But it takes longer for capital to adjust, in lined with our previous observations for convergence paths. The price variables, wage and interest rate, is countercyclical to the motion of labour and capital. For inequality measures, income inequality is procyclical to social infection. Wealth Gini index also increases when infectivity increases. And the impact on wealth inequality is more persistent.

VII. Conclusion

Our paper extends the representative agent epidemiological economic model to a heterogeneous agent framework. We explore individuals' optimal policy and aggregate response to infectious disease. We conclude our paper in following remarks: (1) The policy functions for prevention and treatment expenditure are increasing and more elastic with higher wealth. Health expenditure inelastic to wealth if the disutility of being infected is small. (2) In the stationary equilibrium, infection rate for the poor individuals is higher and thus diverges individuals' labour income. (3) In the short-run, there would be pandemic before the stationary equilibrium is reached. The outbreak is milder with higher level of pre-existing immunity. (4) Income and wealth equality is worsened during the pandemic. The motion of inequality measures are driven by individuals' heterogeneous health policies during the pandemic. (5) Income support scheme can mitigate the negative effect on equality. But it may lead to worse aggregate health outcome.

The heterogeneous agent model in our paper generates increase in income inequality based on optimal policy functions on response to infectious diseases. The wealthy spend more on health because they are endowed with more assets to use. The mechanism is different from Hall and Jones (2007) where health is a superior good with an income elasticity well above one. We abstract from other mechanisms

that can also increase inequality including unemployment; sectorial heterogeneity (Chetty et al., 2023); remote working and digital devices (Stantcheva, 2022); mortality (Goenka, Liu and Nguyen, 2022); furlough scheme (Görtz, McGowan and Yeromonahos, 2021) etc.

References

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

Achdou, Yves, Guy Barles, Hitoshi Ishii, Grigory L Litvinov, and Yves Achdou. 2013. "Finite difference methods for mean field games." Hamilton-Jacobi Equations: Approximations, Numerical Analysis and Applications: Cetraro, Italy 2011, Editors: Paola Loreti, Nicoletta Anna Tchou, 1–47.

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. 2022. "Income and wealth distribution in macroeconomics: A continuous-time approach." The review of economic studies, 89(1): 45–86.

Achdou, Yves, Pierre Cardaliaguet, François Delarue, Alessio Porretta, Filippo Santambrogio, Yves Achdou, and Mathieu Laurière. 2020. "Mean field games and applications: Numerical aspects." Mean Field Games: Cetraro, Italy 2019, 249–307.

Adams-Prassl, Abi, Teodora Boneva, Marta Golin, and Christopher Rauh. 2020. "Work tasks that can be done from home: Evidence on variation within & across occupations and industries."

Aiyagari, S Rao. 1994. "Uninsured idiosyncratic risk and aggregate saving." The Quarterly Journal of Economics, 109(3): 659–684.

Alipour, Jean-Victor, Oliver Falck, and Simone Schüller. 2023. "Germany's capacity to work from home." European Economic Review, 151: 104354.

Alvarez, Fernando, David Argente, and Francesco Lippi. 2021. "A simple planning problem for COVID-19 lock-down, testing, and tracing." American Economic Review: Insights, 3(3): 367–382.

Angelopoulos, Konstantinos, Spyridon Lazarakis, Rebecca Mancy, and Max Schroeder. 2021. "Pandemic-induced wealth and health inequality and risk exposure."

Atkeson, Andrew. 2021. "A parsimonious behavioral SEIR model of the 2020 COVID epidemic in the United States and the United Kingdom." National Bureau of Economic Research.

Bai, Chenjiang, Yuejiao Duan, Xiaoyun Fan, and Shuai Tang. 2023. "Financial market sentiment and stock return during the COVID-19 pandemic." Finance Research Letters, 54: 103709.

Bartik, Alexander W, Zoe B Cullen, Edward L Glaeser, Michael Luca, and Christopher T Stanton. 2020. "What jobs are being done at home during the COVID-19 crisis? Evidence from firm-level surveys." National Bureau of Economic Research.

Beigel, John H, Kay M Tomashek, Lori E Dodd, Aneesh K Mehta, Barry S Zingman, Andre C Kalil, Elizabeth Hohmann, Helen Y Chu, Annie Luetkemeyer, Susan Kline, et al. 2020. "Remdesivir for the treatment of Covid-19." New England Journal of Medicine, 383(19): 1813–1826.

Belot, M, S Choi, JC Jamison, NW Papageorge, E Tripodi, and E Van den Broek-Altenburg. 2020. "Unequal Consequences of Covid 19 Across Age and Income: Representative Evidence from Six Countries: CEPR Discussion Paper No." DP14908.[Google Scholar].

Blundell, Richard, Monica Costa Dias, Robert Joyce, and Xiaowei Xu. 2020. "COVID-19 and Inequalities." Fiscal studies, 41(2): 291–319.

Bobrovitz, Niklas, Harriet Ware, Xiaomeng Ma, Zihan Li, Reza Hosseini, Christian Cao, Anabel Selemon, Mairead Whelan, Zahra Premji, Hanane Issa, et al. 2023. "Protective effectiveness of previous SARS-CoV-2 infection and hybrid immunity against the omicron variant and severe disease: a systematic review and meta-regression." The Lancet Infectious Diseases.

Borgonovi, Francesca, Elodie Andrieu, and SV Subramanian. 2020. "Community level social capital and COVID-19 infections and fatality in the United States." Covid Economics, 32: 110–126.

Bradley, Chris, and Peter Stumpner. 2021. "The impact of COVID-19 on capital markets, one year in." McKinsey & Company.

Brown, Caitlin S, and Martin Ravallion. 2020. "Inequality and the coronavirus: Socioeconomic covariates of behavioral responses and viral outcomes across US counties." national Bureau of economic research.

Cagigi, Alberto, Meng Yu, Björn Österberg, Julia Svensson, Sara Falck-Jones, Sindhu Vangeti, Eric Åhlberg, Lida Azizmohammadi, Anna Warnqvist, Ryan Falck-Jones, et al. 2021. "Airway antibodies emerge according to COVID-19 severity and wane rapidly but reappear after SARS-CoV-2 vaccination." JCI insight, 6(22).

Calvia, Alessandro, Fausto Gozzi, Francesco Lippi, and Giovanni Zanco. 2023. "A simple planning problem for COVID-19 lockdown: a dynamic programming approach." Economic Theory, 1–28.

Cardaliaguet, Pierre, François Delarue, Jean-Michel Lasry, and Pierre-Louis Lions. 2019. The master equation and the convergence problem in mean field games:(ams-201). Princeton University Press.

Carmona, René, François Delarue, et al. 2018. Probabilistic theory of mean field games with applications I-II. Springer.

Chetty, Raj, John N Friedman, Michael Stepner, and Opportunity Insights Team. 2023. "The Economic Impacts of Covid-19: Evidence from a New Public Database Built Using Private Sector Data." The Quarterly Journal of Economics, qjad048.

Coven, Joshua, and Arpit Gupta. 2020. "Disparities in mobility responses to COVID-19." New York Univ, 1.

CreditSuisse. 2023. "Global Wealth Report 2023." Global Wealth Report. https://www.credit-suisse.com/about-us/en/reports-research/global-wealth-report.html.

Di Fusco, Manuela, Xiaowu Sun, Mary M Moran, Henriette Coetzer, Joann M Zamparo, Laura Puzniak, Mary B Alvarez, Ying P Tabak, and Joseph C Cappelleri. 2022. "Impact of COVID-19 and effects of BNT162b2 on patient-reported outcomes: quality of life, symptoms, and work productivity among US adult outpatients." Journal of Patient-Reported Outcomes, 6(1): 123.

Di Matteo, Livio. 2003. "The income elasticity of health care spending: A comparison of parametric and nonparametric approaches." The European Journal of health economics, 4: 20–29.

Dingel, Jonathan I, and Brent Neiman. 2020. "How many jobs can be done at home?" Journal of public economics, 189: 104235.

Fan, Ying, A Yeşim Orhun, and Dana Turjeman. 2020. "Heterogeneous actions, beliefs, constraints and risk tolerance during the COVID-19 pandemic." National Bureau of Economic Research.

Federico, Salvatore, Giorgio Ferrari, and Maria-Laura Torrente. 2022. "Optimal vaccination in a SIRS epidemic model." Economic Theory, 1–26.

Fernández-Villaverde, Jesús, Samuel Hurtado, and Galo Nuno. 2023. "Financial frictions and the wealth distribution." Econometrica, 91(3): 869–901.

Freeman, Donald G. 2003. "Is health care a necessity or a luxury? Pooled estimates of income elasticity from US state-level data." Applied Economics, 35(5): 495–502.

Galasso, Vincenzo, Vincent Pons, Paola Profeta, Michael Becher, Sylvain Brouard, and Martial Foucault. 2020. "Gender differences in COVID-19 attitudes and behavior: Panel evidence from eight countries." Proceedings of the National Academy of Sciences, 117(44): 27285–27291.

Garriga, Carlos, Rody Manuelli, and Siddhartha Sanghi. 2022. "Optimal management of an epidemic: Lockdown, vaccine and value of life." Journal of Economic Dynamics and Control, 140: 104351.

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

Gilboa, M, G Regev-Yochay, M Mandelboim, V Indenbaum, K Asraf, R Fluss, et al. 2022. "Durability of Immune Response After COVID-19 Booster Vaccination and Association With COVID-19 Omicron Infection. JAMA Netw Open. 2022; 5 (9): e2231778." Epub 2022/09/16. https://doi.org/10.1001/jamanetworkopen. 2022.31778 PMID ⋯.

Glover, Andrew, Jonathan Heathcote, Dirk Krueger, and José-Víctor Ríos-Rull. 2023. "Health versus wealth: On the distributional effects of controlling a pandemic." Journal of Monetary Economics, 140: 34–59.

Goenka, Aditya, and Lin Liu. 2020. "Infectious diseases, human capital and economic growth." Economic Theory, 70: 1–47.

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

Goenka, Aditya, Lin Liu, and Manh-Hung Nguyen. 2021. "SIR economic epidemiological models with disease induced mortality." Journal of Mathematical Economics, 93: 102476.

Goenka, Aditya, Lin Liu, and Manh-Hung Nguyen. 2022. "Modelling optimal lockdowns with waning immunity." Economic theory, 1–38.

Görtz, Christoph, Danny McGowan, and Mallory Yeromonahos. 2021. "Furlough and Household Financial Distress during the COVID-19 Pandemic." Oxford Bulletin of Economics and Statistics.

Gottlieb, Charles, Jan Grobovšek, Markus Poschke, and Fernando Saltiel. 2021. "Working from home in developing countries." European Economic Review, 133: 103679.

Hakki, Seran, Jie Zhou, Jakob Jonnerby, Anika Singanayagam, Jack L Barnett, Kieran J Madon, Aleksandra Koycheva, Christine Kelly, Hamish Houston, Sean Nevin, et al. 2022. "Onset and window of SARS-CoV-2 infectiousness and temporal correlation with symptom onset: a prospective, longitudinal, community cohort study." The Lancet Respiratory Medicine, 10(11): 1061–1073.

Hale, Thomas, Samuel Webster, Anna Petherick, Toby Phillips, and B Kira. 2022. "Oxford COVID-19 government response tracker (OxCGRT)." Last updated, 8: 30.

Hall, Robert E, and Charles I Jones. 2007. "The value of life and the rise in health spending." The Quarterly Journal of Economics, 122(1): 39–72.

Hensvik, Lena, Thomas Le Barbanchon, and Roland Rathelot. 2020. "Which jobs are done from home? Evidence from the American Time Use Survey."

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

Huang, Minyi, Roland P Malhamé, and Peter E Caines. 2006. "Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle."

Kaplan, Greg, Benjamin Moll, and Giovanni L Violante. 2020. "The great lockdown and the big stimulus: Tracing the pandemic possibility frontier for the US." National Bureau of Economic Research.

Krueger, Dirk, Kurt Mitman, and Fabrizio Perri. 2016. "Macroeconomics and household heterogeneity." In Handbook of macroeconomics. Vol. 2, 843–921. Elsevier.

Krusell, Per, and Anthony A Smith, Jr. 1998. "Income and wealth heterogeneity in the macroeconomy." Journal of political Economy, 106(5): 867–896.

Laibson, David, Peter Maxted, and Benjamin Moll. 2022. "A simple mapping from mpcs to mpxs." National Bureau of Economic Research.

Lasry, Jean-Michel, and Pierre-Louis Lions. 2007. "Mean field games." Japanese journal of mathematics, 2(1): 229–260.

Lauriere, Mathieu. 2021. "Numerical methods for mean field games and mean field type control." Mean field games, 78: 221.

Lechien, Jerome R, Carlos M Chiesa-Estomba, Sammy Place, Yves Van Laethem, Pierre Cabaraux, Quentin Mat, Kathy Huet, Jan Plzak, Mihaela Horoi, Stéphane Hans, et al. 2020. "Clinical and epidemiological characteristics of 1420 European patients with mild-to-moderate coronavirus disease 2019." Journal of internal medicine, 288(3): 335–344.

Lekfuangfu, Warn N, Suphanit Piyapromdee, Ponpoje Porapakkarm, and Nada Wasi. 2020. "On Covid-19: New implications of job task requirements and spouse's occupational sorting1." Covid Economics, 87.

Lewandowski, Piotr, Katarzyna Lipowska, and Iga Magda. 2021. "The gender dimension of occupational exposure to contagion in Europe." Feminist Economics, 27(1-2): 48–65.

Liu, Ying, and Joacim Rocklöv. 2022. "The effective reproductive number of the Omicron variant of SARS-CoV-2 is several times relative to Delta." Journal of Travel Medicine, 29(3): taac037.

Mongey, Simon, Laura Pilossoph, and Alexander Weinberg. 2021. "Which workers bear the burden of social distancing?" The Journal of Economic Inequality, 19: 509–526.

Moscone, Francesco, and Elisa Tosetti. 2010. "Health expenditure and income in the United States." Health economics, 19(12): 1385–1403.

ONS. 2020. "Deaths involving COVID-19 by local area and socioeconomic deprivation: deaths occurring between 1 March and 31 July 2020."

ONS. 2022a. "Coronavirus (COVID-19) Infection Survey UK funded academic projects, summary results: 2022."

ONS. 2022b. "Labour costs and labour income, UK: 2022."

ONS. 2023a. "Coronavirus (COVID-19) latest insights."

ONS. 2023b. "More information on data sources related to coronavirus (COVID-19)."

Ozkan, Serdar. 2014. "Preventive vs. curative medicine: A macroeconomic analysis of health care over the life cycle." Unpublished. https://sites. google. com/site/serdarozkan/Ozkan_2014. pdf.

Papageorge, Nicholas W, Matthew V Zahn, Michèle Belot, Eline Van den Broek-Altenburg, Syngjoo Choi, Julian C Jamison, and Egon Tripodi. 2021. "Socio-demographic factors associated with self-protecting behavior during the Covid-19 pandemic." Journal of population economics, 34: 691–738.

PHE. 2020. "Disparities in the risk and outcomes of COVID-19 Public Health England."

Safan, Muntaser, Mirjam Kretzschmar, and Karl P Hadeler. 2013. "Vaccination based control of infections in SIRS models with reinfection: special reference to pertussis." Journal of mathematical biology, 67: 1083–1110.

Stachurski, John, and Alexis Akira Toda. 2019. "An impossibility theorem for wealth in heterogeneous-agent models with limited heterogeneity." Journal of Economic Theory, 182: 1–24.

Stantcheva, Stefanie. 2022. "Inequalities in the Times of a Pandemic." Economic Policy, 37(109): 5–41.

Sun, Chengjun, and Wei Yang. 2010. "Global results for an SIRS model with vaccination and isolation." Nonlinear Analysis: Real World Applications, 11(5): 4223–4237.

UK Health Security Agency. 2023. "COVID-19 Omicron variant infectious period and transmission from people with asymptomatic compared with symptomatic infection: a rapid review."

Weill, Joakim A, Matthieu Stigler, Olivier Deschenes, and Michael R Springborn. 2020. "Social distancing responses to COVID-19 emergency declarations strongly differentiated by income." Proceedings of the national academy of sciences, 117(33): 19658–19660.

Yan, Youpei, Amyn A Malik, Jude Bayham, Eli P Fenichel, Chandra Couzens, and Saad B Omer. 2021. "Measuring voluntary and policy-induced social distancing behavior during the COVID-19 pandemic." Proceedings of the National Academy of Sciences, 118(16): e2008814118.

Appendices

A Figures

Figure A1 – Gini Index from 2012 to 2022

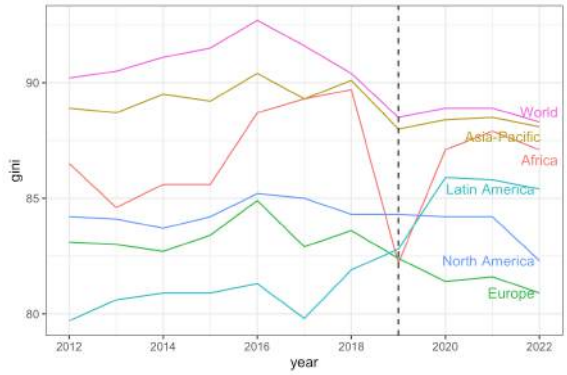


Figure A2 – Lorenz Curve

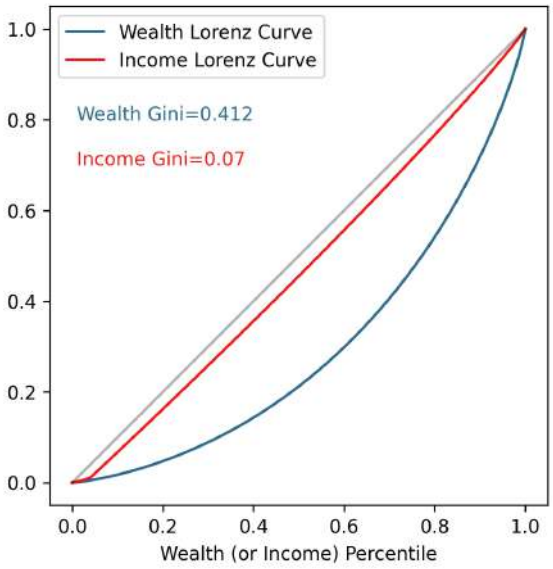


Figure A3 – Labour Income Share

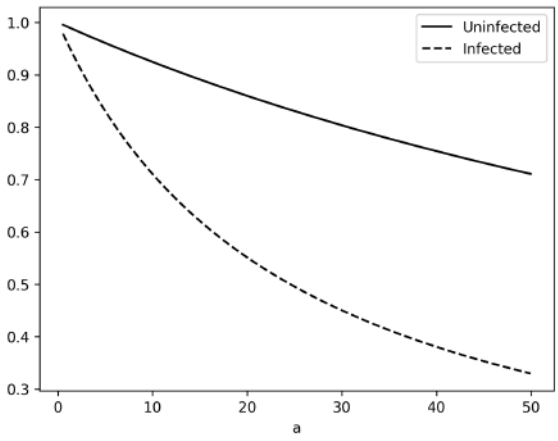


Figure A4 – Health Expenditure Elasticity of Income ($\chi = 0$)

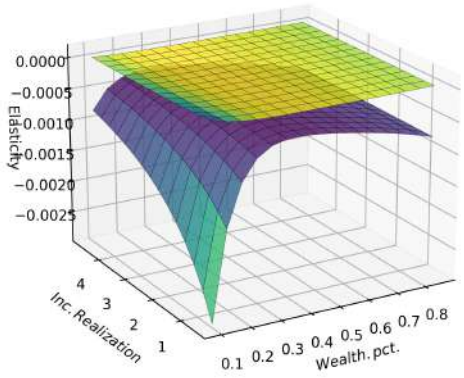


Figure A5 – Multiple Waves of Infection

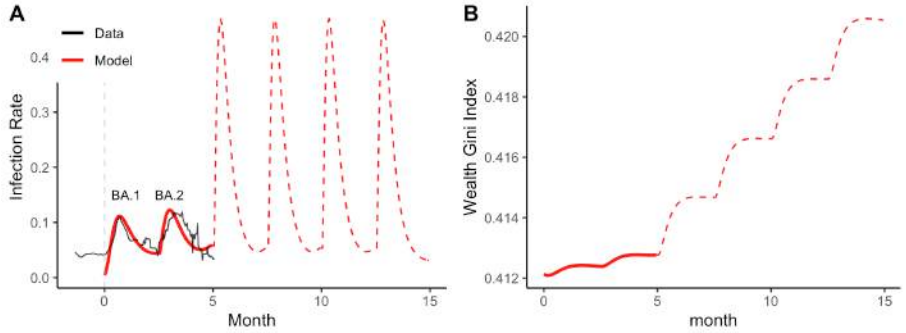


Figure A6 – Other Income Support

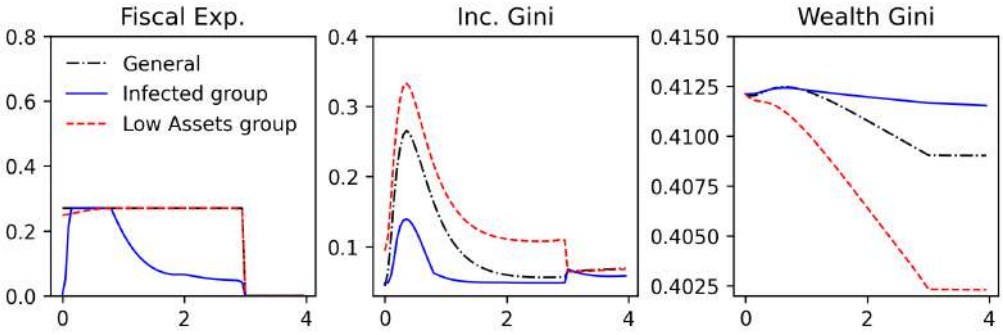


Figure A7 – Transitional Dynamics — Lower Productivity Loss

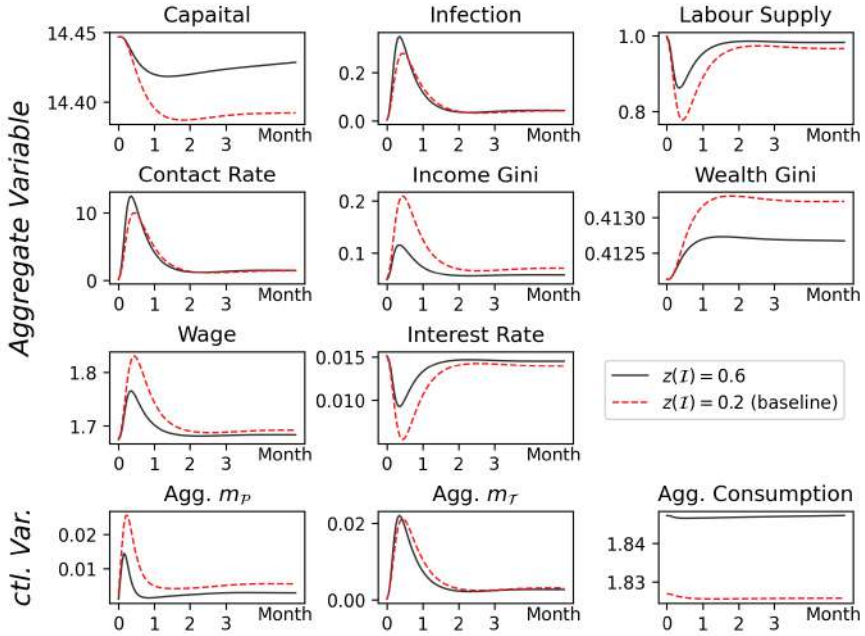


Figure A8 – Transitional Dynamics — Long Covid

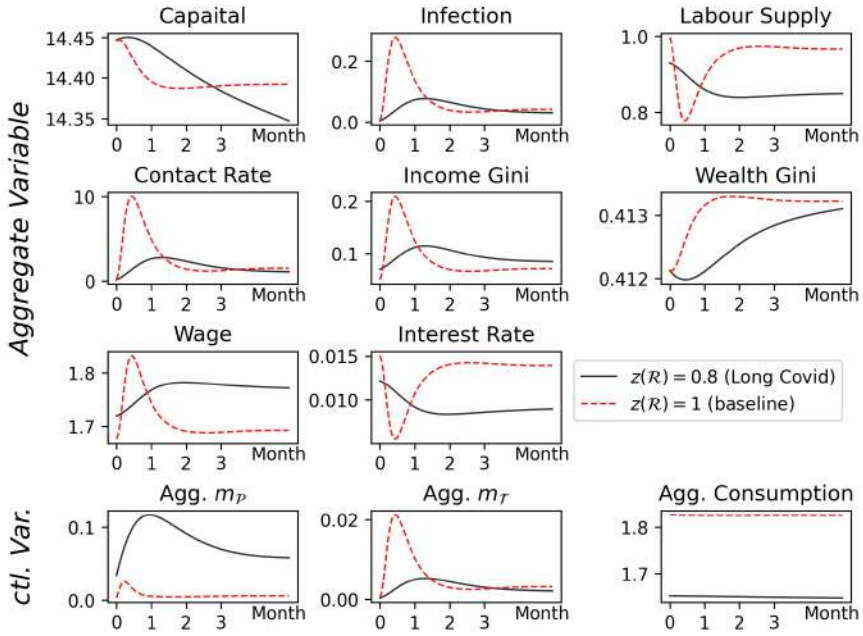
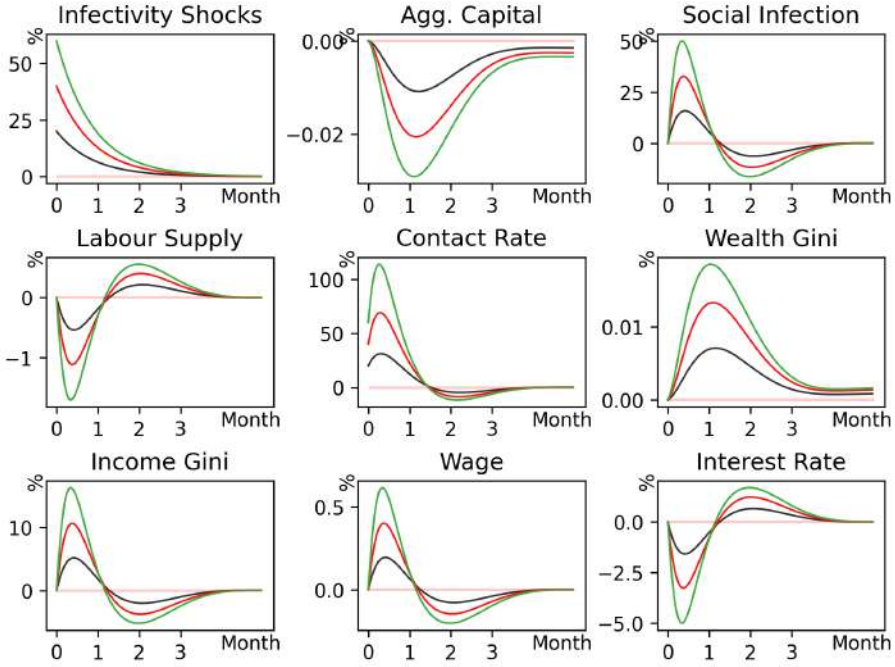


Figure A9 – MIT Shock



B Table

Table B1 – Change of Wealth Gini Index

	gini2016	gini2019	gini2020	gini2022	gini2020	gini2022
(Intercept)	35.5***	7.16	19.2***	22.2***	23.5***	30.8***
	(3.73)	(5.01)	(2.80)	(2.91)	(2.79)	(2.84)
gini2012	0.610***					
	(0.050)					
gini2016		0.849***				
		(0.062)				
gini2019			0.769***	0.730***	0.737***	0.651***
			(0.037)	(0.038)	(0.035)	(0.035)
GovSupp					-2.70***	-4.12***
					(0.546)	(0.580)
Observations	148	158	157	153	148	143
R^2	0.505	0.548	0.738	0.707	0.780	0.789

Table B2 – Comparative Study (changing ϵ_0)

(a) Aggregate Variables

ϵ_0	0.162	0.216	0.252
Infection Rate			
Aggregate	3.998	4.253	4.343
Bottom 25%	4.128	4.378	4.467
Top 25%	3.909	4.177	4.274
diff.	-0.219	-0.201	-0.193
Capital	14.459	14.429	14.419
Prices			
Wage Rate	1.694	1.694	1.694
Interest Rate	0.014	0.014	0.014
Inequality			
Wealth Gini	0.413	0.411	0.41
Income Gini	0.07	0.071	0.072
Wealth Share			
Bottom 25%	6.64	6.72	6.75
Top 25%	53.09	52.94	52.87
diff	46.46	46.22	46.12

(b) Contol Variables

ϵ_0	0.162	0.216	0.252
Consumption			
Aggregate	1.83	1.828	1.828
Bottom 25%	1.681	1.681	1.68
Top 25%	2.054	2.051	2.05
diff.	0.372	0.37	0.37
Preventive Exp.			
Aggregate	0.051	0.053	0.053
Bottom 25%	0.047	0.049	0.05
Top 25%	0.056	0.058	0.059
diff.	0.009	0.009	0.009
Treatment Exp.			
Aggregate	0.075	0.075	0.075
Bottom 25%	0.072	0.072	0.072
Top 25%	0.079	0.079	0.079
diff.	0.007	0.007	0.007

Table B4 – Comparative Study (changing ϵ_0 cont.)

ϵ_0	0.162	0.216	0.252
Expected Income in 3-yr duration			
Bottom 25%	20.547	20.512	20.5
Top 25%	25.099	25.03	25.003
diff.	4.553	4.518	4.504
Labour Income diff.	0.015	0.01	0.009
Capital Income diff.	4.502	4.469	4.456
Expected Savings in 3-yr duration			
Bottom 25%	2.616	2.816	2.887
Top 25%	-0.958	-1.036	-1.063
diff.	-3.574	-3.852	-3.95

Table B5 – Comparative Study (changing ϵ_1)

(a) Aggregate Variables

ϵ_1	-1.1	-1.2	-1.3
Infection Rate			
Aggregate	4.365	4.491	4.551
Bottom 25%	4.49	4.613	4.671
Top 25%	4.295	4.431	4.497
diff.	-0.195	-0.182	-0.174
Capital	14.416	14.401	14.394
Prices			
Wage Rate	1.694	1.694	1.694
Interest Rate	0.014	0.014	0.014
Inequality			
Wealth Gini	0.41	0.409	0.408
Income Gini	0.072	0.072	0.073
Wealth Share			
Bottom 25%	6.75	6.8	6.83
Top 25%	52.92	52.81	52.73
diff	46.16	46.01	45.9

(b) Contol Variables

ϵ_1	-1.1	-1.2	-1.3
Consumption			
Aggregate	1.827	1.827	1.827
Bottom 25%	1.68	1.68	1.68
Top 25%	2.05	2.049	2.048
diff.	0.37	0.369	0.368
Preventive Exp.			
Aggregate	0.056	0.06	0.063
Bottom 25%	0.052	0.056	0.058
Top 25%	0.062	0.066	0.069
diff.	0.01	0.011	0.011
Treatment Exp.			
Aggregate	0.075	0.075	0.075
Bottom 25%	0.072	0.072	0.072
Top 25%	0.079	0.079	0.079
diff.	0.007	0.007	0.007

Table B7 – Comparative Study (changing ϵ_1 cont.)

ϵ_1	-1.1	-1.2	-1.3
Expected Income in 3-yr duration			
Bottom 25%	20.496	20.479	20.472
Top 25%	25.003	24.966	24.944
diff.	4.508	4.486	4.472
Labour Income diff.	0.009	0.006	0.004
Capital Income diff.	4.459	4.44	4.426
Expected Savings in 3-yr duration			
Bottom 25%	2.914	3.013	3.056
Top 25%	-1.071	-1.108	-1.126
diff.	-3.985	-4.121	-4.182

Table B8 – Comparative Study (changing η_1)

(a) Aggregate Variables

η_1	-1.1	-1.2	-1.3
Infection Rate			
Aggregate	4.086	4.069	4.057
Bottom 25%	4.211	4.193	4.179
Top 25%	4.003	3.988	3.976
diff.	-0.208	-0.205	-0.203
Capital	14.449	14.451	14.452
Prices			
Wage Rate	1.694	1.694	1.694
Interest Rate	0.014	0.014	0.014
Inequality			
Wealth Gini	0.413	0.413	0.414
Income Gini	0.07	0.07	0.07
Wealth Share			
Bottom 25%	6.65	6.63	6.61
Top 25%	53.09	53.13	53.17
diff	46.44	46.5	46.56

(b) Contol Variables

η_1	-1.1	-1.2	-1.3
Consumption			
Aggregate	1.83	1.83	1.831
Bottom 25%	1.681	1.682	1.682
Top 25%	2.054	2.055	2.055
diff.	0.372	0.373	0.373
Preventive Exp.			
Aggregate	0.051	0.051	0.051
Bottom 25%	0.048	0.047	0.047
Top 25%	0.057	0.056	0.056
diff.	0.009	0.009	0.009
Treatment Exp.			
Aggregate	0.068	0.062	0.057
Bottom 25%	0.066	0.06	0.055
Top 25%	0.072	0.065	0.06
diff.	0.006	0.005	0.005

Table B10 – Comparative Study (changing η_1 cont.)

η_1	-1.1	-1.2	-1.3
Expected Income in 3-yr duration			
Bottom 25%	20.534	20.535	20.536
Top 25%	25.08	25.087	25.093
diff.	4.546	4.552	4.557
Labour Income diff.	0.013	0.013	0.013
Capital Income diff.	4.497	4.504	4.509
Expected Savings in 3-yr duration			
Bottom 25%	2.657	2.622	2.593
Top 25%	-0.974	-0.96	-0.949
diff.	-3.631	-3.582	-3.541

C R_0 IN HETEROGENEOUS AGENT MODEL

Basic reproduction number R_0 is defined as the average number of secondary infections that occur when one infective is introduced into a completely susceptible host population. The replacement number R (Effective reproduction number) is defined to be the average number of secondary infections produced by a typical infective during the entire period of infectiousness.

In a simple epidemiological model with SIRS dynamics, the motion of infection rate can be written as

$$\begin{aligned} \dot{i} &= \alpha s i - \gamma i \\ &= \gamma i(\frac{\alpha s}{\gamma} - 1) \end{aligned} \tag{25}$$

where α and γ is the contact and recovery rate. This sign of this expression is governed by $\frac{\alpha s}{\gamma} - 1$. Therefore, the varying ratio $\frac{\alpha s_t}{\gamma}$ is the effective reproduction number R. If R is positive, $\dot{i} > 0$ that infection rate increases. Given a complete host population with $s_0 = 1$, the basic reproduction number is

$$R_0 = \frac{\alpha s_0}{\gamma} = \frac{\alpha}{\gamma} \tag{26}$$

Morover, we have the proposition that $R_t = R_0 s_t$

The key difference of solving R_0 in our heterogeneous agent framework is the difficulty that contact rate and recovery rate are endogenized by health expenditure. α, γ vary across the wealth distribution. However, We can find R_0 in a similar way using Kolmogorov Forward Equation. Analogue to Equation 25, the net flow of infectious group in our model is

$$\begin{aligned} \dot{i} &= \int \alpha(m_{\mathcal{P}})\zeta g(a, \mathcal{S}) da - \int \gamma(m_{\mathcal{T}}) g(a, \mathcal{I}) da \\ &= \int \gamma(m_{\mathcal{T}}) g(a, \mathcal{I}) da \left(\frac{\int \alpha(m_{\mathcal{P}})\zeta g(a, \mathcal{S}) da}{\int \gamma(m_{\mathcal{T}}) g(a, \mathcal{I}) da} - 1 \right) \end{aligned} \tag{27}$$

We can similarly define the effective reproduction number and basic reproduction number as

$$\begin{aligned} R_t &= \frac{\int \alpha(m_{\mathcal{P}})\zeta g(a, \mathcal{S}) da}{\int \gamma(m_{\mathcal{T}}) g(a, \mathcal{I}) da} \\ R_0 &= \frac{R_t}{s_t} = \frac{\int \alpha(m_{\mathcal{P}})\zeta g(a, \mathcal{S}) da}{\int \gamma(m_{\mathcal{T}}) g(a, \mathcal{I}) da \int \mathbb{1}(h = \mathcal{S}) g(a, h) d\mu} \end{aligned} \tag{28}$$

D FINITE DIFFERENCING METHOD

To solve the following HJB-KF system
(Optimal HJB)

$$\begin{aligned}\rho v(a,h) = u(c^*) &+ \partial_a v(a,h)[wz^h(h) + ra - c^* - m^*_{\mathcal{P}} - m^*_{\mathcal{T}}] \\ &+ \Lambda^{h'}(m^*_{\mathcal{S}}, m^*_{\mathcal{I}}, h)[v(a,h') - v(a,h)] + \partial_t v(a,h)\end{aligned} \tag{29}$$

where

$$c^* = u'^{-1}(\partial_a v(a,h)) \tag{30}$$

$$m^*_{\mathcal{P}} = \begin{cases} 0; & h = \{\mathcal{I}, \mathcal{R}\} \\ \alpha'^{-1}\left(\frac{\partial_a v(a,\mathcal{S})}{\zeta[v(a,\mathcal{I}) - v(a,\mathcal{S})]}\right); & h = \mathcal{S} \end{cases} \tag{31}$$

$$m^*_{\mathcal{T}} = \begin{cases} 0; & h = \{\mathcal{S}, \mathcal{R}\} \\ \gamma'^{-1}\left(\frac{\partial_a v(a,\mathcal{I})}{v(a,\mathcal{R}) - v(a,\mathcal{I})}\right); & h = \mathcal{I} \end{cases} \tag{32}$$

(KF)

$$\begin{aligned}\frac{\partial g(a,h)}{\partial t} = &- \frac{\partial}{\partial a}[s(a,h)g(a,h)] \\ &- \Lambda^{h'}(m^*,h)g(a,h) + \Lambda^h(m^*,h')g(a,h')\end{aligned} \tag{33}$$

(Assets Market Clearing)

$$K = \int a g(a,h) d\mu \tag{34}$$

(Labour Market Clearing)

$$L = \int z(h) g(a,h) d\mu \tag{35}$$

(Rational Expectation)

$$\zeta = \int \alpha(m^*_{\mathcal{P}}) \mathbb{1}(h = \mathcal{I}) g(a,h) d\mu \tag{36}$$

We discretize the state space (a,h) into $na \times 3$ grids, indexed by i,j. Then, we can express the partial differentiation $\partial_a v(a,h)$ by the forward and backward difference as follow

$$\begin{aligned}\partial_a^F v &= \frac{v_{i+1,j} - v_{i,j}}{da} \\ \partial_a^B v &= \frac{v_{i,j} - v_{i-1,j}}{da}\end{aligned} \tag{37}$$

Then, we can obtain the corresponding optimal policies using forward or backward difference by Equation 30 to Equation 32, listed $c_{i,j}^{*F}, c_{i,j}^{*B}, m_{\mathcal{S}i,j}^{*F}, m_{\mathcal{S}i,j}^{*B}, m_{\mathcal{I}i,j}^{*F}, m_{\mathcal{I}i,j}^{*B}$. Let the motion of asset (savings) using forward and backward difference to be $s_{i,j}^F$ and $s_{i,j}^B$ that

$$\begin{aligned}s_{i,j}^F &= wz_{i,j} + ra_{i,j} - c_{i,j}^{*F} - m_{\mathcal{P}i,j}^{*F} - m_{\mathcal{T}i,j}^{*F} \\ s_{i,j}^B &= wz_{i,j} + ra_{i,j} - c_{i,j}^{*B} - m_{\mathcal{P}i,j}^{*B} - m_{\mathcal{T}i,j}^{*B}\end{aligned} \tag{38}$$

The Finite Differencing Method uses the forward difference if $s_{i,j}^F > 0$, backward difference if $s_{i,j}^B < 0$. By this principal, the Equation 29 could be rewritten as

$$\begin{aligned}\rho v_{i,j} &= u(\tilde{c}_{i,j}) + \partial_a^F v_{i,j} s_{i,j}^+ + \partial_a^B v_{i,j} s_{i,j}^- + \tilde{\Lambda}_{i,j}[v_{i,j'} - v_{i,j}] + \partial_t v_{i,j} \\ &= u(\tilde{c}_{i,j}) + (\frac{v_{i+1,j} - v_{i,j}}{da}) s_{i,j}^+ + (\frac{v_{i,j} - v_{i-1,j}}{da}) s_{i,j}^- + \tilde{\Lambda}_{i,j}[v_{i,j'} - v_{i,j}] + \partial_t v_{i,j} \\ &= u(\tilde{c}_{i,j}) + \left(\frac{s_{i,j}^+}{da}\right) v_{i+1,j} + \left(\frac{s_{i,j}^- - s_{i,j}^+}{da}\right) v_{i,j} - \left(\frac{s_{i,j}^-}{da}\right) v_{i-1,j} + \tilde{\Lambda}_{i,j}[v_{i,j'} - v_{i,j}] + \partial_t v_{i,j}\end{aligned} \tag{39}$$

where $s_{i,j}^+ \equiv s_{i,j}^F \mathbb{1}(s_{i,j}^F > 0)$, $s_{i,j}^- \equiv s_{i,j}^B \mathbb{1}(s_{i,j}^B < 0)$. $\tilde{c}_{i,j}$ is the optimal consumption under FDM scheme, that $\tilde{c}_{i,j} \equiv c_{i,j}^{*F} \mathbb{1}(s_{i,j}^F > 0) + c_{i,j}^{*B} \mathbb{1}(s_{i,j}^B < 0)$. $\tilde{\Lambda}_{i,j}$ is the transition probability under optimal health expenditure defined in the similar way. Let $\tilde{m}_{\mathcal{S}i,j} \equiv m_{\mathcal{S}i,j} \mathbb{1}(s_{i,j}^F > 0) + m_{\mathcal{S}i,j} \mathbb{1}(s_{i,j}^B < 0)$ and $\tilde{m}_{\mathcal{I}i,j} \equiv m_{\mathcal{I}i,j} \mathbb{1}(s_{i,j}^F > 0) + m_{\mathcal{I}i,j} \mathbb{1}(s_{i,j}^B < 0)$, $\tilde{\Lambda}_{i,j}$ could be expressed as.

$$\begin{aligned}\tilde{\Lambda}_{i,1} &= \alpha(\tilde{m}_{\mathcal{P}i,1})\zeta \quad \text{Infection Probability} \\ \tilde{\Lambda}_{i,2} &= \gamma(\tilde{m}_{\mathcal{T}i,2}) \quad \text{Recovery Probability} \\ \tilde{\Lambda}_{i,3} &= \psi \quad \text{Reinfection Probability}\end{aligned} \tag{40}$$

Boundary Conditoins. Our model assumes wealth $a \in [\underline{a}, \bar{a}]$. In the boundaries of the distribution, wealth cannot go further down or up. Hence, the savings must be non-negative at the lower bound; non-positive at the upper bound. Hence, if we have $s_{na,j}^F > 0$ with $a_{na} = \bar{a}$ or $s_{1,j}^B < 0$ with $a_1 = \underline{a}$, we need to activate the boundary conditions.

We use the lower bound as an example. We make use of the F.O.C. and the boundary condition $s_{1,j}^B = 0$. We can write

$$\begin{cases} wz(\mathcal{S}) + r\underline{a} - c_{1,1}^{*B} - m_{\mathcal{S}_{1,1}}^{*B} = 0 \\ wz(\mathcal{I}) + r\underline{a} - c_{1,2}^{*B} - m_{\mathcal{I}_{1,2}}^{*B} = 0 \\ wz(\mathcal{R}) + r\underline{a} - c_{1,3}^{*B} = 0 \end{cases} \tag{41}$$

that is

$$\begin{cases} wz(\mathcal{S}) + r\underline{a} - u'^{-1}(\partial_a^B v_{1,1}) - \alpha'^{-1}\left(\frac{\partial_a v_{1,1}}{\zeta[v_{1,2}-v_{1,1}]}\right) = 0 \\ wz(\mathcal{I}) + r\underline{a} - u'^{-1}(\partial_a^B v_{1,2}) - \gamma'^{-1}\left(\frac{\partial_a v_{1,2}}{v_{1,1}-v_{1,2}}\right) = 0 \\ wz(\mathcal{R}) + r\underline{a} - u'^{-1}(\partial_a^B v_{1,3}) = 0 \end{cases} \tag{42}$$

where we make use of the F.O.C. Equation 31 and Equation 32. Equation 42 are equations for $\partial_a^B v(a,h)$. We can thus solve out backward differentiation at the lower bound of a. Finally, we use these solutions and F.O.C.s to update the value of choice variables and savings at the boundary. Similar strategy could be applied to the upper bound.

Matrix Notation for HJB. We could further express Equation 39 as

$$\begin{aligned} \rho v_{i,1} &= u(\tilde{c}_{i,1}) + \left(\frac{s_{i,1}^+}{da}\right) v_{i+1,1} + \left(\frac{s_{i,1}^- - s_{i,1}^+}{da}\right) v_{i,1} - \left(\frac{s_{i,j}^-}{da}\right) v_{i-1,1} + \tilde{\Lambda}_{i,1}[v_{i,2} - v_{i,1}] + \partial_t v_{i,1} \\ \rho v_{i,2} &= u(\tilde{c}_{i,2}) + \left(\frac{s_{i,2}^+}{da}\right) v_{i+1,2} + \left(\frac{s_{i,2}^- - s_{i,2}^+}{da}\right) v_{i,2} - \left(\frac{s_{i,2}^-}{da}\right) v_{i-1,2} + \tilde{\Lambda}_{i,2}[v_{i,3} - v_{i,2}] + \partial_t v_{i,2} \\ \rho v_{i,3} &= u(\tilde{c}_{i,3}) + \left(\frac{s_{i,3}^+}{da}\right) v_{i+1,3} + \left(\frac{s_{i,3}^- - s_{i,3}^+}{da}\right) v_{i,3} - \left(\frac{s_{i,3}^-}{da}\right) v_{i-1,3} + \tilde{\Lambda}_{i,3}[v_{i,1} - v_{i,3}] + \partial_t v_{i,3} \end{aligned} \tag{43}$$

Next, we stack $v_{i,1}, v_{i,2}, v_{i,3}, i = 1, 2, \ldots na$ into one long vector

$$\mathbf{V} = \begin{bmatrix} v_{1,1} \\ v_{2,1} \\ \cdots \\ v_{na,1} \\ v_{1,2} \\ v_{2,2} \\ \cdots \\ v_{na,2} \\ \cdots \\ v_{na,3} \end{bmatrix} \tag{44}$$

Then, Equation 43 could be written into matrix form as

$$\rho \mathbf{V} = \mathbf{u} + \mathbf{A}\mathbf{V} + \mathbf{B}\mathbf{V} + \partial_t \mathbf{V} \tag{45}$$

$$\mathbf{A} \equiv \begin{bmatrix} \frac{s^-_{1,1}-s^+_{1,1}}{da} & \frac{s^+_{1,1}}{da} & 0 & 0 & \dots & 0 \\ -\frac{s^-_{2,1}}{da} & \frac{s^-_{2,1}-s^+_{2,1}}{da} & \frac{s^+_{2,1}}{da} & 0 & \dots & 0 \\ 0 & -\frac{s^-_{3,1}}{da} & \frac{s^-_{3,1}-s^+_{3,1}}{da} & \frac{s^+_{3,1}}{da} & \dots & 0 \\ 0 & 0 & -\frac{s^-_{4,1}}{da} & \frac{s^-_{4,1}-s^+_{4,1}}{da} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \frac{s^-_{na,3}-s^+_{na,3}}{da} \end{bmatrix} \tag{46}$$

$$\mathbf{B} = \begin{bmatrix} -\mathbf{L}_1 & \mathbf{L}_1 & \mathbf{0} \\ \mathbf{0} & -\mathbf{L}_2 & \mathbf{L}_2 \\ \mathbf{L}_3 & \mathbf{0} & -\mathbf{L}_3 \end{bmatrix} \tag{47}$$

where

$$\begin{aligned} \mathbf{L}_1 &= Diag(\Lambda_{1,1}, \Lambda_{2,1}, \dots \Lambda_{na,1})_{na \times na} \\ \mathbf{L}_2 &= Diag(\Lambda_{1,2}, \Lambda_{2,2}, \dots \Lambda_{na,2})_{na \times na} \\ \mathbf{L}_3 &= Diag(\Lambda_{1,3}, \Lambda_{2,3}, \dots \Lambda_{na,3})_{na \times na} \end{aligned} \tag{48}$$

Thus, define $\mathcal{A} = \mathbf{A} + \mathbf{B}$, the HJB equation could be written as

$$\rho \mathbf{V} = \mathbf{u} + \mathcal{A}\mathbf{V} + \partial_t \mathbf{V} \tag{49}$$

Expressing $\partial_t \mathbf{V} = \frac{V_{t+1}-V_t}{\Delta}$ and using the implicit update method, we can have

$$\begin{aligned} \frac{\mathbf{V}_{t+1} - \mathbf{V}_t}{\Delta} + \rho \mathbf{V}_{t+1} &= \mathbf{u} + \mathcal{A}\mathbf{V}_{t+1} \\ \left((\rho + \frac{1}{\Delta})I - \mathcal{A}\right)\mathbf{V}_{t+1} &= \mathbf{u} + \frac{\mathbf{V}_t}{\Delta} \\ \mathbf{V}_{t+1} &= \left((\rho + \frac{1}{\Delta})I - \mathcal{A}\right)^{-1} \frac{\mathbf{V}_t}{\Delta} \end{aligned} \tag{50}$$

The Kolmogorov Forward Equation Equation 33 can be then expressed asset

$$\dot{\mathbf{g}}_t = \mathcal{A}^* \mathbf{g}_t \tag{51}$$

where $\mathcal{A}^*$ is the adjoint matrix for $\mathcal{A}$ defined before.

E STATIONARY EQUILIBRIUM AND TRANSITIONAL DYNAMICS

A. *Stationary Equilibrium*

Solving the Stationary Equilibrium

In the stationary equilibrium, we have the joint distribution $g(a,h)$ and value function $v(a,h)$ being unchanged over time. That is $\dot{\mathbf{g}}_t = 0$ and $\dot{\mathbf{V}}_t = 0$

Thus, given any initial value function $\mathbf{V}^0$ and price (w,r) [40], we can iterate Equation 50 until $abs(\mathbf{V}_{t+1} - \mathbf{V}_t)$ is sufficiently small. The corresponding value function $\mathbf{V}^*$ would be value function at the stationary equilibrium. Then, we can solve out the corresponding stationary distribution $\mathbf{g}$ by $0 = \mathcal{A}^*\mathbf{g}$.

Next, we calculate the aggregate supply by

$$\begin{aligned} K_S &= \int a g^*(a,h) d\mu \\ i_S &= \int \mathbb{1}(h = \mathcal{I}) g^*(a,h) d\mu \\ \zeta_S &= \int \alpha(m^*_{\mathcal{P}}) \mathbb{1}(h = \mathcal{I}) g^*(a,h) d\mu \end{aligned} \tag{52}$$

We update the aggregate demand and repeat the steps until the aggregate supply and aggregate demand are sufficiently close. The following bullet points summarize the Algorithm

- (i) Given (K_D, i_D, ζ_D), calculate (r,w)
- (ii) Loop the HJB Equation 50 to find $\mathcal{A}$ and $\mathbf{V}^*$
- (iii) Solve out the corresponding stationary equilibrium by $0 = \mathcal{A}^*\mathbf{g}$
- (iv) Use $\mathbf{g}$ to calculate aggregate supply using Equation 52
- Repeat step (i) to (iv) until the gap between aggregate supply and aggregate demand is sufficient small

Income Elasticity of Health Expenditure

The Income Elasticity of Health Expenditure is defined as follow

$$\begin{aligned} \varepsilon_{M_h,y} &= \frac{\partial M_h(a_0,h_0)}{\partial y} \frac{y}{M_h} \\ &= \frac{M_h(a_0+\Delta,h_0) - M_h(a_0,h_0)}{\Delta} \frac{a_0}{M_h(a_0,h_0)} \\ h &= \{\mathcal{P},\mathcal{T}\} \end{aligned} \tag{53}$$

[40] The price determined by the aggregate supply (K_D, L_D, ζ_D)

where

$$\begin{aligned} M_{\mathcal{P}}(a_0, h_0) &= \mathbb{E}\left[\int_0^{\tau} m_{\mathcal{P}}(a_t, h_t)dt \middle| a_0, h_0\right] \\ M_{\mathcal{T}}(a_0, h_0) &= \mathbb{E}\left[\int_0^{\tau} m_{\mathcal{T}}(a_t, h_t)dt \middle| a_0, h_0\right] \end{aligned} \tag{54}$$

To solve the cumulative health expenditure over the period of τ, we apply the Feynman-Kac Formula

Lemma (Feynman-Kac Formula): *Define a time varying function* $\Gamma(a, h, t)$ *satisfying PDE*

$$0 = m_{\mathcal{P}}(a, h) + \partial_a \Gamma(a, h, t)\dot{a}(a, h) + \Lambda^{h'}(h)[\Gamma(a, h', t) - \Gamma(a, h, t)] + \partial_t \Gamma(a, h, t) \tag{55}$$

with terminal condition $\Gamma(a, h, \tau) = 0$*, where* $h = \{\mathcal{S}, \mathcal{I}, \mathcal{R}\}$*. Then,* $M_{\mathcal{P}}(a_0, h_0) = \Gamma(a, h, 0)$.
Similarly, $M_{\mathcal{T}}(a_0, h_0)$ *can be computed as* $\Gamma(a, h, 0)$*, where* $\Gamma(a, h, t)$ *satisfies*

$$0 = m_{\mathcal{T}}(a, h) + \partial_a \Gamma(a, h, t)\dot{a}(a, h) + \Lambda^{h'}(h)[\Gamma(a, h', t) - \Gamma(a, h, t)] + \partial_t \Gamma(a, h, t) \tag{56}$$

Corollary: *The second part of Feynman-Kac Formula is exactly* $\mathcal{A}$ *matrix in the FDM algorithm. Thus, the PDE could be written as*

$$0 = x(a, h) + \mathcal{A}\Gamma(a, h, t) + \partial_t \Gamma(a, h, t) \tag{57}$$

where $x(a, h) = \{c(a, h), ra + wz^h(h)\}$

Computation: We know the terminal condition $\Gamma(a, h, \tau) = 0$, we can solve $\Gamma(a, h, t)$ backward recursively and obtain $\Gamma(a, h, 0)$ by

$$\begin{aligned} 0 &= x(a, h) + \mathcal{A}\Gamma(a, h, t-1) + \frac{\Gamma(a, h, t) - \Gamma(a, h, t-1)}{\Delta t} \\ \left(\frac{1}{\Delta t} I - \mathcal{A}\right)\Gamma(a, h, t-1) &= x(a, h) + \frac{\Gamma(a, h, t)}{\Delta t} \end{aligned} \tag{58}$$

B. Transitional Dynamics

We now simulate the transitional dynamics given initial distribution $g_0(a, h)$ and terminal value function $v_T(a, h)$. Assume an initial path of aggregate demand $\{K_t^D\}, \{i_t^D\}, \{\zeta_t^D\}$, where $t = \{0, 1, ...N\}$ with time interval dt. Using the paths of

aggregate variables, we can compute the paths for price $\{w_t\}$ and $\{r_t\}$.

Then, we can solve the HJB backward by

$$\mathbf{V}_t = u(\mathbf{V}_{t+1}) + \mathcal{A}(\mathbf{V}_{t+1}; r_t, w_t, \zeta_t)\mathbf{V}_t + \frac{\mathbf{V}_{t+1} - \mathbf{V}_t}{dt} \tag{59}$$

where we could obtain $\{\mathbf{V}_t\}$. Then, we can solve KF forward by

$$\frac{\mathbf{g}_{t+1} - \mathbf{g}_t}{dt} = \mathcal{A}^*(\mathbf{V}_t; r_t, w_t, \zeta_t)\mathbf{g}_{t+1} \tag{60}$$

to obtain $\{\mathbf{g}_t\}$. Using the sequence of distribution, we can solve out sequence of aggregate supply $\{K_t^S\}, \{i_t^S\}, \{\zeta_t^S\}$. Finally, we repeat the steps above until the aggregate supply sequence and aggregate demand sequence are sufficiently close to each other. The following bullet points summarize the Algorithm for transitional dynamics

- (i) Given sequence of demand $(\{K_t^D\}, \{i_t^D\}, \{\zeta_t^D\})$, calculate sequence $(\{r_t\}, \{w_t\})$
- (ii) Solve HJB backward from stationary equilibrium $\mathbf{V}^*$ by Equation 59 to find sequence $\{\mathbf{V}_t\}$
- (iii) Solve KF forward from the initial distribution $\{\mathbf{g}_0\}$ by Equation 60 and obtain distribution sequence $\{\mathbf{g}_t\}$
- (iv) Use sequence $\{\mathbf{g}_t\}$ to calculate sequences for aggregate supply $(\{K_t^S\}, \{i_t^S\}, \{\zeta_t^S\})$
- Repeat step (i) to (iv) until the gap between aggregate supply and aggregate demand is sufficient small

For the MIT shock, we can also solve the transitional dynamics using the algorithm above. In the case of MIT shock, the terminal equilibrium is identical to the equilibrium before the shock arrives. Hence, we can just use $v^*(a, h)$ as the terminal value function and $g^*(a, h)$ as the initial distribution.