Epidemics, pandemics and income inequality

Chrys Esseau-Thomas[1], Omar Galarraga[2] and Sherif Khalifa[3*]

Abstract

The novel coronavirus is part of a series of infectious disease outbreaks that include: Ebola, Avian influenza, Middle East respiratory syndrome coronavirus, and Influenza A. This paper addresses the question of how do these epidemics and pandemics affect income inequality in countries around the world during the first two decades of this century. To achieve its objective, the paper develops a model that indicates a positive association between these health crises and income inequality. To empirically test our theoretical predictions, the paper explores the effect on the Gini coefficient of a dummy variable that indicates the occurrence of an epidemic or a pandemic in a country in a given year and the number of deaths per 100,000. To properly address potential endogeneity, we implement a Three-Stage-Least Squares technique. The estimation shows that the number of deaths per 100,000 population variable has a statistically significant positive effect on the Gini coefficient, especially when we incorporate COVID-19 data. This suggests that not only the occurrence, but also the health consequences of COVID-19 have a significant and economically important effect on income inequality.

Background: The purpose of the study is to examine the effect of epidemics and pandemics on income inequality. This has important implications as the outcome of this study can guide policymakers into implementing policies that can mitigate the economic consequences of these health crises.

Methods: The study is a cross country analysis using fixed effects estimation. To address potential endogeneity and determine causality, the paper uses the Three-Stage-Least-Squares estimation.

Results: The paper finds that the number of epidemic deaths per 100,000 population variable has a statistically significant positive effect on the Gini coefficient, especially when we incorporate COVID-19 data.

Conclusions: The paper finds that it is not only the occurrence of an epidemic, captured by the epidemics dummy variable, but also the health consequences, captured by the number of deaths per 100,000 population, that have a significant effect on income inequality. This is especially the case when we incorporate COVID-19 in our analysis.

Trial registration: Not Applicable.

Keywords: Epidemics, Executive, Income inequality

JEL classification: I14, D31

* Correspondence: skhalifa@fullerton.edu
[3]California State University, Fullerton, CA 92834, USA
Full list of author information is available at the end of the article

Introduction

This paper examines the effect of epidemics and pandemics on income inequality. An epidemic is a widespread outbreak of an infectious disease in a community. A pandemic is an epidemic that has spread worldwide. This paper argues that epidemics and pandemics have significant effects on income distribution either through their potential direct disproportionate health effects, or through their indirect economic consequences caused by the intervention of governments to contain these health crises. This question is important as policy makers around the world attempt nowadays to comprehend and address the consequences of emerging disease epidemics and pandemics.

To achieve its objective, the paper develops a simple theoretical setup extending the framework in Moser and Yared [18] to one with low skilled workers and high skilled workers. In this context, policy makers can impose a lockdown to contain an epidemic or a pandemic. The lockdown policy imposes a cap on the supply of low skilled workers and causes more high skilled workers to work remotely. The model shows that there is a threshold lockdown policy beyond which income inequality increases. To empirically test the theoretical findings, the paper uses a panel of 191 countries during the period 2000–2020 to examine the effect of epidemics and pandemics on income inequality. During this period, the world experienced several outbreaks of infectious diseases such as MERS-Cov, H1N1, SARS, Ebola and SARS-Cov-2. In this context, we use a dummy variable to indicate the occurrence of an epidemic or a pandemic in a country in a given year, in addition to the number of deaths per 100,000 population. Panel estimation techniques are used to explore the association between these health indicators and the Gini Coefficient.

In this context, the paper addresses the issue of potential endogeneity. As much as epidemics and pandemics can affect income distribution, the extent of income inequality can determine the disproportionate exposure of various groups to a health crisis causing it to exacerbate to an epidemic or a pandemic. To properly address the issue of causality, we use a Three-Stage-Least Squares estimation technique. The estimation considers several control variables that are identified by the literature as determinants of income inequality, such as the level of economic development captured by the logarithm of GDP per capita, the sectoral structure of the economy, the institutional quality captured by the extent of corruption and the democratic system of governance, in addition to indicators of fiscal policy, monetary policy and trade policy.

The fixed effects estimation shows that the dummy variable has a robust and statistically significant positive effect on income inequality, while the number of deaths per 100,000 population has an insignificant effect. The interpretation is that it is the occurrence of a pandemic, captured by the dummy variable, that is what triggers the policy response that may affect income inequality, rather than the health consequences of these crises, captured by the number of deaths per 100,000. To properly address potential endogeneity, we implement a Three-Stage-Least Squares estimation technique. The 3SLS estimations show that the number of deaths per 100,000 population variable has a statistically significant positive effect on the Gini coefficient, especially when we incorporate the effect of COVID-19. This suggests that not only the occurrence but also the health consequences of COVID-19 have a significant effect on income inequality.

The remainder of the paper is organized as follows: section 2 contains the literature survey, section 3 contains the theoretical model, section 4 discusses the intuition, section 5 on methods includes (a) the description of the data, and (b) the empirical estimation, section 6 presents the baseline results and robustness tests, second 7 includes a final discussion, and section 8 concludes. References, tables and figures are included thereafter.

Literature

This paper contributes to a nascent literature on the economic consequences of epidemics and pandemics. Studies in this literature focus primarily on the effects of these health crises on labor market outcomes, income distribution and aggregate economic conditions. For instance, Lin and Meissner [16] find that stay-at-home orders are only weakly associated with slower growth of COVID-19 cases, while job losses have been no higher in U.S. states that implemented stay-at-home orders during the COVID-19 pandemic than others that did not. Rojas et al. [22] estimate how the spread of COVID-19 and school closures affect labor market conditions at the state level. The authors find that the slowdown in the labor market was due to the nationwide reaction, while state policies and epidemiological conditions have had a modest effect. Papanikolaou and Schmidt [20] analyze the supply-side disruptions of COVID-19. The authors find that sectors in which a larger fraction of the workforce is not able to work remotely experienced greater declines in employment, greater declines in expected revenue growth, worse stock market performance, and higher expected likelihood of default.

Borjas and Cassidy [5] show that the employment decline due to COVID-19 was severe for immigrants who had previous employment advantage in the labor market. This is because immigrants were less likely to work in jobs that could be performed remotely during the lockdown. Montenovo et al. [17] show greater declines in employment during the COVID-19 pandemic for

Hispanics, young workers, those with high school degrees and some college, and those with occupations that require more interpersonal contact. Forsythe et al. [11] find that the collapse in job vacancies and the spike in unemployment insurance claims due to Covid-19 affected all the states of the United States, irrespective of the intensity of the spread of the virus or the timing of stay-at-home policies. The authors also show that nearly all industries and occupations saw contraction in job postings and increases in unemployment insurance claims, with the exception of essential front line jobs.

Sumner et al. [24] show that COVID-19 poses a challenge to the United Nations Sustainable Development Goal of ending poverty by 2030 as "global poverty could increase for the first time since 1990 demonstrating a reversal of a decade of progress in reducing world poverty." The authors show that under the scenario of a 20% income or consumption contraction, the number of people living in poverty could increase by 420–580 million, compared to 2018. Fenichel [10] compare social distancing outcomes under decentralized, full control social planner, and constrained social planner. The author finds that constrained social planner decision making can in some instances make society worse off than decentralized decision making with no intervention.

Some studies also discuss the distributional consequences of epidemics. For instance, Bell and Gersbach [3] find that during an outbreak, programs under which supported families enjoy the benefits of spending on health and education simultaneously are superior to those under which the benefits are sequenced. The authors also find that these superior programs restrict support to fewer families initially and thus increases income inequality. Glover et al. [14] develop a model that predicts that young workers in sectors deemed non-essential would benefit from ending the shutdown, while others would lose. The authors conclude that redistribution can render the welfare impacts less unequal. Furceri et al. [12] show that the COVID-19 pandemic could exert an adverse effect on income inequality in the absence of efforts to protect the vulnerable segments of society. The authors conclude that their finding that inequality increases with the adverse effect of a pandemic on economic activity "suggests that, all else equal, the distributional consequences of COVID-19 may be larger than those in previous pandemic episodes." Palomino et al. [19] explore the effect of lockdowns on income inequality and poverty in Europe. The authors find that a two months' lockdown scenario causes an average increase in the headcount poverty index of 4.9 percentage points, and an increase in the Gini coefficient by 3.5 percentage point. The authors also show that these estimates increase under longer lockdown scenarios and that both between-countries and within-countries

income inequality significantly increase, with the change of the latter being more important.

Our paper's contribution to the literature is threefold. It is the first to have a global analysis of the effects of epidemics and pandemics on income inequality, it is the first to address the issue of causality between these health crises and income inequality, and it is also the first to develop a simple theoretical framework that helps us understand the effect of these health crises on income inequality.

Model

Consider an infinite-horizon model of an economy during a pandemic. Periods are indexed by $t = 0, 1, \ldots$. The economy is populated by a continuum of low skilled agents of unit mass, and a continuum of high skilled agents of unit mass. In each period, policy makers choose a lockdown policy $\{L_t, H_t\}$. L_t determines the fraction of low skilled labor who cannot work, while H_t determines the fraction of high skilled labor who can work remotely. This lockdown policy affects economic output negatively but curbs the spread of disease. Second, labor is supplied such that the wage equals the marginal product of labor. Third, production takes place and proceeds are paid to firms and workers. Finally, the pandemic evolves according to a model of disease spread, which depends on the lockdown policy. Thus, this model is an extension of that of Moser and Yared [18] to a setup where we introduce two types of labor and with a primary focus on the effect of pandemic lockdown policy on income inequality.

The production function states that output depends on the supply of low skilled labor l_t, high skilled labor h_t, and Ω_t which is the health state at date t. Thus, the production function is as follows

$$y_t = f(l_t, h_t, \Omega_t) \qquad (1)$$

We assume that the function $f(.)$ is continuously differentiable, increasing, and globally concave in l_t and h_t. Thus, $lim_{l \to 0}(\frac{\partial f}{\partial l}) = lim_{h \to 0}(\frac{\partial f}{\partial h}) = \infty$ and $lim_{l \to \infty}(\frac{\partial f}{\partial l}) = lim_{h \to \infty}(\frac{\partial f}{\partial h}) = 0$. The dependence of the production function on the health state reflects the possibility that the pandemic could decrease output by debilitating the labor force, by changing the share of the labor force working from office and from home, and by adopting social distancing efforts that may affect production.

The policy makers choose a lockdown policy $L_t \in [0, 1]$ which reflects the fraction of low skilled labor that is prohibited from working. If $L_t = 0$, there is no lockdown and all low skilled agents can go to work. If $L_t = 1$, there is maximal lockdown and no agent is allowed to work. Low skilled agents inelastically supply effective labor l_t up to an upper bound $(1-L_t)\bar{l}(\Omega_t)$. The lockdown policy

also determines H_t which is the fraction of high skilled labor that will work remotely. If $H_t = 0$, there is no lock-down and all high skilled agents can go to work in-person. If $H_t = 1$, there is maximal lockdown and all high skilled agents are only allowed to work remotely. High skilled agents inelastically supply effective labor h_t up to an upper bound $(1-H_t)\bar{h}(\Omega_t) + H_t\tilde{h}(\Omega_t)$. In this context, those working in-person supply $\bar{h}(\Omega_t)$ while those working remotely supply $\tilde{h}(\Omega_t)$, where $\tilde{h}(\Omega_t) > \bar{h}(\Omega_t)$ $\forall t$. This is because those who work remotely will be able to supply more labor as they save on the time they spend commuting to work in the absence of a lockdown. Thus, the labor market clearing conditions are as follows

$$l_t = (1-L_t)\bar{l}(\Omega_t) \tag{2}$$

$$h_t = (1-H_t)\bar{h}(\Omega_t) + H_t\tilde{h}(\Omega_t) \tag{3}$$

Firms maximize profits π_t such that

$$\pi_t = y_t - w_t^l l_t - w_t^h h_t \tag{4}$$

where the price of output is normalized to 1, w_t^l is the wage of low skilled labor, and w_t^h is the wage of high skilled labor. Labor is competitively supplied so wages equal the marginal product of labor as follows

$$w_t^l = \frac{\partial f(l_t, h_t, \Omega_t)}{\partial l_t} \tag{5}$$

$$w_t^h = \frac{\partial f(l_t, h_t, \Omega_t)}{\partial h_t} \tag{6}$$

In this context, workers consume their wage income

$$c_t^l = w_t^l l_t \tag{7}$$

$$c_t^h = w_t^h h_t \tag{8}$$

where c_t^l is the aggregate consumption of low skilled labor, and c_t^h is the aggregate consumption of high skilled labor. In this context, social welfare equals the discounted sum of utility streams as follows

$$\sum_{t=0}^{\infty} \beta^t u\left(c_t^l, c_t^h, \Omega_t\right) \tag{9}$$

where $\beta \in (0, 1)$ is the discount factor, and $u(.)$ is a strictly increasing and strictly concave utility function of consumption c_t^l and c_t^h and the health state Ω_t. In this economy, income inequality can be captured by the share of income of high skilled labor to the total income of low skilled and high skilled labor as follows

$$\frac{w_t^h h_t}{\left(w_t^l l_t + w_t^h h_t\right)} \tag{10}$$

Finally, we model disease spread following a Susceptible-Infected-Recovered-Dead (SIRD) model as

in Moser and Yared [18]. This model defines a law of motion of the health state as follows

$$\Omega_{t+1} = \Gamma(H_t, L_t, \Omega_t) \tag{11}$$

which depends on the degree of lockdown at date t. The initial health state Ω_0 is taken as given. This setup allows us to derive some conclusions on the relationship between the pandemic lockdown policy and income inequality.

Proposition 1 *There is a threshold H^* such that $\left(\frac{\partial w_t^h h_t}{\partial h_t}\right) > 0$, $\forall H > H^*$ if and only if $\left(\frac{\partial w_t^h}{\partial h_t}\right) < 1$ at $H = 1$.*

Proof. At $H = 0$, we have to consider three possibilities: either $\left(\frac{\partial w_t^h}{\partial h_t}\right) > 1$, $\left(\frac{\partial w_t^h}{\partial h_t}\right) = 1$ or $\left(\frac{\partial w_t^h}{\partial h_t}\right) < 1$. (1) If $\left(\frac{\partial w_t^h}{\partial h_t}\right) > 1$ at $H = 0$ and $\left(\frac{\partial w_t^h}{\partial h_t}\right) < 1$ at $H = 1$. Given that the marginal product of high skilled labor decreases with an increase in the supply of high skilled labor since $lim_{h\to\infty}\left(\frac{\partial f}{\partial h}\right) = 0$, then there exists a threshold $H^* \in (0, 1)$ where $\left(\frac{\partial w_t^h}{\partial h_t}\right) = 1$. This implies that $\left(\frac{\partial w_t^h h_t}{\partial h_t}\right) > 0$, $\forall H > H^*$. (2) If $\left(\frac{\partial w_t^h}{\partial h_t}\right) = 1$ at $H = 0$ then $H^* = 0$. (3) If $\left(\frac{\partial w_t^h}{\partial h_t}\right) < 1$ at $H = 0$ then $\left(\frac{\partial w_t^h h_t}{\partial h_t}\right) > 0$ $\forall H$.

Proposition 2 *There is a threshold L^* such that $\left(\frac{\partial w_t^l l_t}{\partial l_t}\right) < 0$, $\forall L > L^*$ if and only if $\left(\frac{\partial w_t^l}{\partial l_t}\right) < 1$ at $L = 1$.*

Proof. At $L = 0$, we have to consider three possibilities: either $\left(\frac{\partial w_t^l}{\partial l_t}\right) > 1$, $\left(\frac{\partial w_t^l}{\partial l_t}\right) = 1$ or $\left(\frac{\partial w_t^l}{\partial l_t}\right) < 1$. (1) If $\left(\frac{\partial w_t^l}{\partial l_t}\right) > 1$ at $L = 0$ and $\left(\frac{\partial w_t^l}{\partial l_t}\right) < 1$ at $L = 1$. Given that the marginal product of low skilled labor increases with a decrease in the supply of low skilled labor since $lim_{l\to 0}\left(\frac{\partial f}{\partial l}\right) = \infty$, then there exists a threshold $L^* \in (0, 1)$ where $\left(\frac{\partial w_t^l}{\partial l_t}\right) = 1$. This implies that $\left(\frac{w_t^l l_t}{\partial l_t}\right) < 0$, $\forall L > L^*$. (2) If $\left(\frac{\partial w_t^l}{\partial l_t}\right) = 1$ at $L = 0$ then $L^* = 0$. (3) If $\left(\frac{\partial w_t^l}{\partial l_t}\right) < 1$ at $L = 0$ then $\left(\frac{w_t^l l_t}{\partial l_t}\right) < 0$ $\forall L$.

Proposition 3 *The pandemic lockdown policy increases income inequality if $H > H^*$ and $L > L^*$.*

Proof. If $H > H^*$, then $\left(\frac{\partial w_t^h h_t}{\partial h_t}\right) > 0$ and if $L > L^*$, then $\left(\frac{\partial w_t^l l_t}{\partial l_t}\right) < 0$. This implies that the lock down policy increases h_t and decreases l_t, then $w_t^h h_t$ increases if $H > H^*$ while $w_t^l l_t$ decreases if $L > L^*$. This causes an increase in income inequality captured by $\left(\frac{w_t^h h_t}{w_t^l l_t + w_t^h h_t}\right)$.

Proposition 4 *H^* decreases with an increase in the slope $\frac{\partial h_t}{\partial H_t} = (\tilde{h}-\bar{h})$.*

Proof. At H^*, $\left(\frac{\partial w_t^h h_t}{\partial h_t}\right) = 0$. We have

$$w_t^h h_t = \left[\frac{\partial f(.)}{\partial h_t}\right] h_t = \frac{\partial f(.)}{\partial h_t} \left[(1-H_t)\bar{h} + H_t\tilde{h}\right]$$

In this case,

$$\left(\frac{\partial w_t^h h_t}{\partial h_t}\right) = \frac{\partial^2 f(.)}{\partial h_t} \left[\bar{h} + (\tilde{h}-\bar{h})H_t\right] + \frac{\partial f(.)}{\partial h_t}$$

which is equal to 0 at H^*. Thus,

$$\left(\frac{\partial w_t^h h_t}{\partial h_t}\right) = \frac{\partial^2 f(.)}{\partial h_t} \left[\bar{h} + (\tilde{h}-\bar{h})H_t^*\right] + \frac{\partial f(.)}{\partial h_t} = 0$$

This implies that in order to maintain the equality, H_t^* is lower with a higher slope $(\tilde{h}-\bar{h})$. A lower H_t^* implies that income inequality can increase with a smaller fraction of high skilled labor working remotely.

Proposition 5 L^* *decreases with an increase in the slope* $\frac{\partial l_t}{\partial L_t} = -\bar{l}$.

Proof. At L^*, $\left(\frac{\partial w_t^l l_t}{\partial l_t}\right) = 0$. We have

$$w_t^l l_t = \left[\frac{\partial f(.)}{\partial l_t}\right] l_t = \frac{\partial f(.)}{\partial l_t} \left[(1-L_t)\bar{l}\right]$$

In this case,

$$\left(\frac{\partial w_t^l l_t}{\partial l_t}\right) = \frac{\partial^2 f(.)}{\partial l_t} \left[(1-L_t)\bar{l}\right] + \frac{\partial f(.)}{\partial l_t}$$

which is equal to 0 at L^*. Thus,

$$\left(\frac{\partial w_t^l l_t}{\partial l_t}\right) = \frac{\partial^2 f(.)}{\partial l_t} \left[(1-L_t^*)\bar{l}\right] + \frac{\partial f(.)}{\partial l_t} = 0$$

This implies that in order to maintain the equality, L_t^* is lower with a higher slope $-\bar{l}$. A lower L_t^* implies that income inequality can increase with a smaller fraction of low skilled labor not working.

Proposition 6 *A higher level of income inequality decreases social welfare if and only if* $H > H^*$ *and* $L > L^*$ *and if* $\left|\frac{\partial w_t^l l_t}{\partial l_t}\right| > \left|\frac{\partial w_t^h h_t}{\partial h_t}\right|$

Proof. Social welfare depends on the consumption of both low skilled and high skilled workers. The level of consumption in turn depends on the total wages. If $H > H^*$, then $\left(\frac{\partial w_t^h h_t}{\partial h_t}\right) > 0$. This causes an increase in the consumption of high skilled agents. If $L > L^*$, $\left(\frac{\partial w_t^l l_t}{\partial l_t}\right) < 0$. This causes a decrease in the consumption of low skilled agents. If $\left|\frac{\partial w_t^l l_t}{\partial l_t}\right| > \left|\frac{\partial w_t^h h_t}{\partial h_t}\right|$ then the decrease in the consumption of low skilled agents is larger than the increase in the consumption of high skilled agents. This decreases social welfare.

Intuition

This section elaborates on the intuition of the relationship between these health crises and income inequality. We expect that epidemics and pandemics cause an increase in income inequality. This is because these health crises cause authorities and policy makers to interfere in an attempt to contain an outbreak before it strains or overwhelms the health care system. Examples of these interventions are stay-at-home orders, shelter-in-place orders, restrictions imposed on in-person transactions, lockdowns and social distancing. These policies can cause the loss of jobs that cannot be done remotely but require in-person presence in the workplace. Many of these jobs are likely to be low-skilled jobs. On the other hand, many high-skilled jobs can be done remotely in an easier manner. Not to mention that high-skilled workers are more capable than low-skilled ones to use the technology that allows them to work remotely. Accordingly, more low-skilled jobs are lost during an epidemic or a pandemic than high-skilled ones. This can cause an increase in income inequality between low-skilled and high-skilled workers.

The health effects of epidemics and pandemics are also disproportionate between income categories. The fatality rate of epidemics or pandemics is usually highest amongst those who are less healthy due to obesity, worse diets or preexisting chronic conditions. These conditions are more common amongst the poor than the affluent. Thus, epidemics and pandemics have disproportionate health consequences that can exacerbate income disparities as well. This is because those who are affected can either die leaving their families without sufficient financial support, or survive but are unable to work for a period of time which adversely affects the family's income. On the other hand, preventive care and public health education have steadily tilted towards the educated and the well-off. These groups tend to react to the epidemics in a manner that better enables them to lessen its spread compared to those who are poorer and less educated. These factors contribute to an increase in income inequality. Epidemics also tend to spread faster in densely populated areas which is typical of poorer communities, compared to the sparsely populated suburbs which only the affluent can afford. This implies that infectious diseases hit poorer areas harder than others. This causes the existing income inequality to exacerbate. The fatality rate of these health crises are also higher amongst the elderly who are more likely to have a larger accumulation of wealth and who may leave bequests to their children. This can increase wealth disparities especially if they have fewer offsprings on average.

Epidemics and pandemics also have an effect on the school system which is likely to affect worker's compensation that is usually dependent on the level of educational

attainment. School closures during an epidemic or a pandemic can widen the student achievement gap. Students from a poorer background are less likely to catch-up in an online educational or remote learning environment compared to their more affluent peers. This could be due to the lack of access to internet or the technological devices needed, the lack of knowledge of the technology used, or the need for both parents to work. This achievement gap in education causes a subsequent income gap. In addition, the policy response to epidemics and pandemics usually causes severe budget deficits that may lead to spending cuts on public education. This adversely affects the prospects of social upward mobility through education, and thus can cause persistence in income inequality. This implies that epidemics and pandemics can also have a long term effect on income inequality.

The economic conditions during epidemics and pandemics also favor larger corporations who can adjust their operations to serve their customers online and to deliver their commodities to the consumers' doorsteps. These big businesses, compared to smaller ones, are also the ones who have sufficient cash buffer to ride out the economic repercussions caused by these health crises. On the other hand, closures of non-essential businesses during lockdowns can cause the shutdown of many small businesses that cannot serve their customers online. This may increase income inequality between shareholders of large corporations and small business owners.

The lower interest rates, due to monetary expansion in response to an epidemic-induced economic slowdown, can lower mortgage rates and increase mortgage applications. This increase in demand for houses can lead to a subsequent increase in house prices. The epidemics and pandemics can also increase the need to relocate to the suburbs away from densely populated urban areas that usually have higher rates of transmission. This also increases the demand for houses. The subsequent appreciation in the value of houses and real estate prices, which is one of the main sources of wealth, can increase wealth inequality between home owners and others. Lower interest rates can also lead to an increase in demand for stocks which could boost stock prices. This can also exacerbate wealth inequality between portfolio owners and others.

Methods

This paper aims to empirically test our theoretical model on the effect of epidemics and pandemics on income inequality. The main study design is a cross-country econometric comparison using panel data. The setting of the study is global in nature as we use data for most countries in the world. In this section we describe the data for the dependent variable (income inequality) and the main variable of interest (epidemics dummy and deaths per 100,000 population), as well as control variables. Then we describe the econometrics methods including fixed effects and three-stage least squares regression analyses.

Data

This paper uses a panel of 191 countries during the period 2000–2020. The countries included in the analysis are: Afghanistan, Albania, Algeria, Angola, Anguilla, Antigua and Barbuda, Argentina, Armenia, Australia, Austria, Azerbaijan, Bahamas, The Bahrain, Bangladesh, Barbados, Belarus, Belgium, Belize, Benin, Bhutan, Bolivia, Bosnia and Herzegovina, Botswana, Brazil, Brunei Darussalam, Bulgaria, Burkina Faso, Burundi, Cabo Verde, Cambodia, Cameroon, Canada, Central African Republic, Chad, Chile, China, Colombia, Comoros, Congo, Democratic, Republic of the Congo, Republic of Costa Rica, Croatia, Cyprus, Czech Republic, Cote d'Ivoire, Denmark, Djibouti, Dominica, Dominican Republic, Ecuador, Egypt, El Salvador, Equatorial Guinea, Estonia, Ethiopia, Fiji, Finland, France, Gabon, Gambia, The Georgia, Germany, Ghana, Greece, Grenada, Guatemala, Guinea, Guinea-Bissau, Guyana, Haiti, Honduras, Hong Kong SAR, Hungary, Iceland, India, Indonesia, Iran, Iraq, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Kiribati, Korea, Kosovo, Kuwait, Kyrgyz Republic, Lao P.D.R., Latvia, Lebanon, Lesotho, Liberia, Libya, Lithuania, Luxembourg, Macedonia FYR, Madagascar, Malawi, Malaysia, Maldives, Mali, Malta, Mauritania, Mauritius, Mexico, Micronesia, Moldova, Mongolia, Montenegro, Rep. of, Morocco, Mozambique, Myanmar, Namibia, Nauru, Nepal, Netherlands, New Zealand, Nicaragua, Niger, Nigeria, Norway, Oman, Pakistan, Palau, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Poland, Portugal, Puerto Rico, Qatar, Romania, Russia, Rwanda, Samoa, Saudi Arabia, Senegal, Serbia, Seychelles, Sierra Leone, Singapore, Slovak Republic, Slovenia, Solomon Islands, Somalia, South Africa, South Sudan, Spain, Sri Lanka, St. Kitts and Nevis, St. Lucia, St. Vincent and the Grenadines, Sudan, Suriname, Swaziland, Sweden, Switzerland, Syria, Sao Tome and Principe, Taiwan, Tajikistan, Tanzania, Thailand, Timor-Leste, Togo, Tonga, Trinidad and Tobago, Tunisia, Turkey, Turkmenistan, Tuvalu, Uganda, Ukraine, United Arab Emirates, United Kingdom, United States, Uruguay, Uzbekistan, Vanuatu, Venezuela, Vietnam, West Bank and Gaza, Yemen, Zambia, and Zimbabwe. The summary statistics of the variables used in the analysis are included in Table 1.

Epidemics and pandemics

Data on epidemics was compiled from the disease profiles in the World Health Organization.[1] The World

---

[1]https://www.who.int/data/collections

Table 1 Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Gini coefficient	2,707	45.44204	6.847209	0	69.4
Epidemic dummy	4,011	0.0870087	0.2818859	0	1
Epidemic deaths per 100k population	3,969	261.7834	9670.335	1	541274.8
Corruption	3,820	-0.0575628	0.9898911	-1.9	2.5
GDP per capita	3,890	11432.45	17116.64	0	118824
Inflation rate	3,534	5.779259	15.57113	-60.5	513.9
Agriculture/GDP	3,726	11.99187	11.74738	0	79
Agriculture/labor	3,540	28.51299	23.66144	0	92.3
Democracy	3,404	2.682726	16.1293	-88	10
Income tax	3,014	6.556578	4.619059	0	32.2331
Population density	3,907	4.253121	1.421133	0.4338433	8.992747
Urbanisation	3,948	56.1293	23.17549	8.246	100
Individualism	1,499	43.61107	23.83992	0	91

Health Organization Data and Statistics are used to identify the occurrence, location and time of the epidemic. This dataset includes, SARS (2002–04),[2] H1N1 (2009),[3] MERS-Cov (2012),[4] Ebola (2013–2016)[5] and SARS-Cov-2 (2020), the virus that causes COVID-19. Using this information, we construct a dummy variable as follows: if a country had a disease profile for that particular epidemic, a value of 1 was assigned; if it did not have a disease profile for the epidemic, a value of zero was assigned. Thus, the dummy variable takes the value of 1 for countries which had reported an epidemic during a year, and zero otherwise. For each of the epidemics we consider, only countries which had been affected by a given epidemic would have a disease profile. These are infectious disease epidemics that are highly contagious and have the potential to become pandemics. For COVID-19, all countries in the sample are affected. We also calculate the number of deaths per 100,000 population due to the epidemic or pandemic as another health indicator.[6]

Income inequality

The data for income inequality is derived from the Standardized World Income Inequality SWIID Database, Version 8.[7] The SWIID's income inequality estimates are based on reported Gini indices from published sources, including the OECD Income Distribution Database, the Socio-Economic Database for Latin America and the Caribbean generated by CEDLAS and the World Bank, Eurostat, the World Bank's PovcalNet, the UN Economic Commission for Latin America and the

Caribbean, national statistical offices around the world, and academic studies. The SWIID incorporates comparable Gini indices of disposable and market income inequality for 198 countries for as many years as possible from 1960 to the present.

Controls

Several control variables are used in the analysis to check the robustness of our results. These are factors that are identified by the literature as potential determinants of income inequality. The first variable used is Real Gross Domestic Product GDP per capita which is derived from the Penn World Tables version 8.0. The logarithm of real Gross Domestic Product per capita is used in the analysis. According to the Kuznets curve, there is an association between the level of economic development and the extent of income inequality.

We also use indicators that reflect the sectoral structure of the economy. In particular, we use the share of agriculture in the labor force and the share of agriculture in GDP. The first variable is defined as Agriculture, forestry, and fishing, value added (% of GDP). The second variable is defined as Agriculture, forestry, and fishing, value added per worker (constant 2010 US$). This data is derived from the World Development Indicators. We expect that an economy that is more agricultural in nature to have a lower level of income inequality.

We use various institutional quality indicators. Democracy reflects the quality of political institutions. The democracy variable is extracted from the Polity IV Project. The Polity score captures a country's political regime on a 21-point scale ranging from – 10 (strongly autocratic) to + 10 (strongly democratic). The paper uses the Polity2 variable which is a modified version of the Polity variable by applying a simple treatment to convert instances of "standardized authority scores" (–66, –77, –88) to conventional polity scores within the range between – 10 to + 10. Some studies argue that a transition to democracy is expected to give greater weight to the

---

[2]https://www.who.int/publications/m/item/summary-of-probable-sars-cases-with-onset-of-illness-from-1-november-2002-to-31-july-2003
[3]https://www.who.int/csr/disease/swineflu/updates/en/
[4]http://www.emro.who.int/pandemic-epidemic-diseases/mers-cov/mers-situation-update-january-2020.html
[5]https://apps.who.int/ebola/current-situation/ebola-situation-report-16-march-2016
[6]https://coronavirus.jhu.edu/data/mortality
[7]https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/LM4OWF

preferences of the poor in collective decision-making. The poor may use the political process to their benefit and influence policy makers to implement inequality-reducing policies. Acemoglu et al. [1] survey the existing empirical literature on why democracy is expected to increase redistribution and decrease income inequality. Another institutional indicator is the corruption perception index derived from Transparency International. Some studies find an association between corruption and income inequality, such as Pedersen [21] and Spinesi [23].

We also include some policy indicators that reflect fiscal policy, monetary policy and trade policy. We include the tax rate in a country, derived from the World Development Indicators. The higher the tax rate the larger is the welfare state, which is expected to decrease income inequality. Another policy indicator is the inflation rate, consumer prices (%) derived from the World Development Indicators. This indicator reflects the efficiency of the conduct of monetary policy by central banks. Several studies examine the effect of inflation on income inequality. For instance, Albanesi [2] provide cross-country evidence on a positive association between inflation and income inequality. We also include trade openness, which is the sum of exports and imports of goods and services as a share of Gross Domestic Product. This is derived from the World Development Indicators. Some studies find an association between trade liberalization and income inequality, such as Goldberg and Pavcnik [15].

Estimation
Results and robustness
This section empirically estimates the effect of epidemics and pandemics on income inequality for the period 2000–2020 as follows

$$Gini_{it} = \alpha + \beta EpiPan_{it-1} + X_{it-1}\gamma + \delta_i + \varepsilon_t + u_{it} \tag{12}$$

where $Gini_{it}$ is the Gini coefficient in country i in year t. $EpiPan_{it-1}$ is the indicator for epidemics or pandemics in country i in year $t - 1$. X_{it-1} is a vector of control variables identified by the literature as determinants of income inequality in country i in year $t - 1$. The δ_i denotes a full set of country dummies, the ε_t denotes a full set of time effects that capture common shocks to income inequality of all countries, and u_{it} is an error term capturing all other omitted factors, with $E(u_{it}) = 0$ for all i and t.

The fixed effects estimation results are included in Table 2 when our variable of interest is the epidemics dummy. Column 1 of Table 2 shows the results without control variables. We add the interaction term between the dummy variable and population density in the second column. This is to distinguish between the effect of epidemics and pandemics on income inequality in low and high levels of population density. The intuition is that the effect of these health crises is exacerbated with high population density. We also add the control variables in subsequent columns. The results in column 1 show that the dummy variable has a statistically significant and positive effect on income inequality. However, when we add the control variables, the dummy variable loses its statistical significance.

Table 3 includes the results when our variable of interest is the number of deaths per 100,000 population. The results in column 1 show that this variable has a weak statistically significant positive effect on income inequality. When we add the control variables in subsequent columns, the number of deaths per 100,000 population

Table 2 Effect of epidemics dummy on the Gini coefficient (Fixed Effects estimation). Notes: Standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

LAGGED VARIABLES	(1) gini	(2) gini	(3) gini	(4) gini	(5) gini	(6) gini	(7) gini	(8) gini
Epidemics dummy	0.3704***	0.0244	0.0678	0.5558	0.9842	1.1051	1.1576	0.9771
	(0.1052)	(0.5688)	(0.5362)	(0.5281)	(0.8453)	(0.8794)	(0.8807)	(0.9176)
Dummy x density		0.0767	0.0691	-0.0449	-0.1109	-0.1174	-0.1271	-0.0911
		(0.1214)	(0.1136)	(0.1062)	(0.1685)	(0.1718)	(0.1726)	(0.1789)
Corruption		0.0938	0.1342	-0.0871	0.0235	0.1409	0.1797	0.1713
		(0.4193)	(0.4309)	(0.5053)	(0.5374)	(0.5577)	(0.5806)	(0.5709)
GDP/capita			-0.0000**	-0.0000*	-0.0000	-0.0000	-0.0000	-0.0000
			(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Democracy				-0.0020	-0.0038	-0.0027	-0.0026	-0.0026
				(0.0040)	(0.0038)	(0.0044)	(0.0043)	(0.0044)
Income tax					-0.1410	-0.1603*	-0.1218	-0.1013
					(0.0921)	(0.0929)	(0.1008)	(0.1031)
Inflation rate						0.0387**	0.0344**	0.0344**
						(0.0157)	(0.0155)	(0.0152)
Agriculture/Labor							0.0532	0.0509
							(0.0604)	(0.0719)
Agriculture/GDP								0.0067
								(0.0537)
Constant	45.4132***	45.5204***	45.9780***	46.0620***	47.1848***	47.1015***	45.4039***	45.2346***
	(0.0061)	(0.0230)	(0.2490)	(0.2379)	(0.6812)	(0.7195)	(2.1278)	(1.8982)
Observations	2,554	2,508	2,491	2,234	1,940	1,833	1,830	1,801
R-squared	0.0008	0.0009	0.0070	0.0053	0.0065	0.0105	0.0138	0.0127
Number of ifscode	189	184	182	158	148	142	141	139

Table 3 Effect of the epidemics deaths per 100,000 population (logs) on the Gini coefficient (Fixed Effects estimation). Notes: Standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

LAGGED VARIABLES	(1) gini	(2) gini	(3) gini	(4) gini	(5) gini	(6) gini	(7) gini	(8) gini	(9) gini
Deaths per 100k (log)	0.1358*	0.3421	0.3406	0.3925	1.4575*	1.9239	2.1191	2.1008	2.0497
	(0.0772)	(0.5936)	(0.5922)	(0.5650)	(0.7728)	(1.3477)	(1.4756)	(1.4710)	(1.4894)
Deaths x density (log)		-0.0391	-0.0387	-0.0464	-0.2254	-0.2811	-0.3143	-0.3123	-0.2408
		(0.1133)	(0.1130)	(0.1071)	(0.1560)	(0.2943)	(0.3167)	(0.3174)	(0.3177)
Corruption			0.1022	0.1442	-0.0816	0.0508	0.1765	0.2154	0.2081
			(0.4207)	(0.4325)	(0.5060)	(0.5424)	(0.5656)	(0.5879)	(0.5788)
GDP/capita				-0.0000**	-0.0000*	-0.0000	-0.0000	-0.0000	-0.0000
				(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Democracy					-0.0020	-0.0039	-0.0027	-0.0027	-0.0027
					(0.0039)	(0.0037)	(0.0043)	(0.0042)	(0.0043)
Income tax						-0.1408	-0.1599*	-0.1222	-0.1023
						(0.0913)	(0.0920)	(0.0998)	(0.1017)
Inflation rate							0.0376**	0.0333**	0.0336**
							(0.0155)	(0.0154)	(0.0151)
Agriculture/Labor								0.0522	0.0502
								(0.0603)	(0.0719)
Agriculture/GDP									0.0065
									(0.0535)
Constant	45.5556***	45.5416***	45.5363***	45.9951***	46.0733***	47.1961***	47.1223***	45.4565***	45.2806***
	(0.0025)	(0.0028)	(0.0219)	(0.2515)	(0.2388)	(0.6757)	(0.7137)	(2.1186)	(1.8857)
Observations	2,527	2,508	2,508	2,491	2,234	1,940	1,833	1,830	1,801
R-squared	0.0002	0.0002	0.0002	0.0063	0.0053	0.0065	0.0101	0.0133	0.0128
Number of ifscode	187	184	184	182	158	148	142	141	139

and the interaction term do not have a statistically significant effect on income inequality.

In Table 4, we include both the epidemics dummy and the number of deaths per 100,000 population. The estimation results show that the dummy variable has a statistically significant and positive coefficient in all specifications, while the number of deaths per 100,000 population is not statistically significant. This implies that it is the occurrence of the pandemic, as captured by the dummy variable, that triggers the policy response that may have an effect on income inequality, rather than its health consequences captured by the number of deaths per 100,000 population.

In Table 5 we include both health variables with the exclusion of Ebola. This is because Ebola has different transmission mechanisms than the other epidemics considered in our paper. Other epidemics are diseases of the respiratory system, with similar transmission mechanisms that require similar policy responses. On the other hand, Ebola is transmitted through contact with bodily fluids. Thus, the reponse is different than those required for respiratory system illnesses. The results in Table 5 confirm our previous finding that the dummy variable has a statistically significant and positive coefficient, while the number of deaths per 100,000 population is not statistically significant in all specifications.

In order to incorporate the effect of the COVID-19 pandemic, we consider the contemporaneous effect of these health indicators on income inequality. This allows for the inclusion of the available data on COVID-19 in the year 2020. Table 6 shows that the dummy variable has a statistically significant and positive effect on income inequality in all specifications, while the death per 100,000 population indicator does not have a robust effect. Table 7 shows the estimation of the contemporaneous effect without Ebola. The results also confirm our finding of a statistically significant positive coefficient for the dummy variable, but not for the death per 100,000 population.

Causality

This section attempts to address the issue of causality between these health crises and income inequality. As much as epidemics and pandemics can affect income distribution as stated in our theoretical setup and the intuition sections, the extent of income inequality can determine the disproportionate exposure of various groups to a health crisis causing it to exacerbate to an epidemic or a pandemic. At high levels of income inequality, there is a large group of low income individuals that experience a higher level of exposure, transmission and mortality in a health crisis. This could be due to the fact that this group typically have worse pre-existing health conditions, have less health insurance coverage, and have less education and awareness of health issues. This can cause income inequality to lead to an exacerbation of a health crisis causing it to turn into an epidemic or a pandemic. An article in the New York Times supports this view and suggests that[8]: *"those in lower economic strata are likelier to catch the disease...... At the same*

---

[8]https://www.nytimes.com/2020/03/15/world/europe/coronavirus-inequality.html

Table 4 Effect of epidemics dummy and epidemics deaths per 100,000 population (logs) on the Gini coefficient (Fixed Effects estimation). Notes: Standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

LAGGED VARIABLES	(1) gini	(2) gini	(3) gini	(4) gini	(5) gini	(6) gini	(7) gini	(8) gini	(9) gini
Epidemics dummy	0.3483***	0.3465***	0.3453***	0.3484***	0.2744***	0.3789***	0.4772***	0.4828***	0.4098***
	(0.1020)	(0.1041)	(0.1031)	(0.0992)	(0.0870)	(0.1248)	(0.1406)	(0.1429)	(0.1424)
Death by 100k (log)	0.0576	0.1093	0.1086	0.1561	1.2374*	1.7940	1.9529	2.0088	1.9237
	(0.0484)	(0.5538)	(0.5523)	(0.5184)	(0.6931)	(1.2717)	(1.3766)	(1.3595)	(1.3853)
Death x density (log)		-0.0097	-0.0094	-0.0162	-0.2087	-0.3143	-0.3552	-0.3731	-0.3033
		(0.1043)	(0.1040)	(0.0972)	(0.1413)	(0.2876)	(0.3072)	(0.3032)	(0.3160)
Corruption			0.0957	0.1366	-0.0846	0.0395	0.1609	0.2015	0.1935
			(0.4203)	(0.4318)	(0.5062)	(0.5417)	(0.5639)	(0.5861)	(0.5773)
GDP/capita				-0.0000**	-0.0000*	-0.0000	-0.0000	-0.0000	-0.0000
				(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Democracy					-0.0020	-0.0039	-0.0027	-0.0027	-0.0027
					(0.0040)	(0.0038)	(0.0044)	(0.0043)	(0.0044)
Income tax						-0.1414	-0.1609*	-0.1218	-0.1025
						(0.0912)	(0.0919)	(0.0997)	(0.1017)
Inflation rate							0.0388**	0.0345**	0.0346**
							(0.0158)	(0.0156)	(0.0153)
Agriculture/labor								0.0531	0.0505
								(0.0605)	(0.0719)
Agriculture/GDP									0.0066
									(0.0536)
Constant	45.5381***	45.5244***	45.5195***	45.9784***	46.0612***	47.1845***	47.1009***	45.3759***	45.2490***
	(0.0064)	(0.0065)	(0.0233)	(0.2496)	(0.2381)	(0.6764)	(0.7149)	(2.1233)	(1.8884)
Observations	2,527	2,508	2,508	2,491	2,234	1,940	1,833	1,830	1,801
R-squared	0.0008	0.0008	0.0009	0.0069	0.0057	0.0071	0.0111	0.0162	0.0135
Number of ifscode	187	184	184	182	158	148	142	141	139

Table 5 Effect of epidemics dummy and epidemics deaths per 100,000 population (logs) on the Gini coefficient (Fixed Effects estimation without Ebola). Notes: Standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

LAGGED VARIABLES	(1) giny	(2) giny	(3) giny	(4) giny	(5) giny	(6) giny	(7) giny	(8) giny	(9) giny
Epidemics dummy	0.3593***	0.3577***	0.3566***	0.3596***	0.2857***	0.3789***	0.4772***	0.4911***	0.4098***
	(0.1028)	(0.1049)	(0.1038)	(0.0999)	(0.0878)	(0.1248)	(0.1406)	(0.1434)	(0.1424)
Death by 100k (log)	0.0551	0.1018	0.1011	0.1485	1.2283*	1.7940	1.9529	1.9253	1.9237
	(0.0476)	(0.5524)	(0.5509)	(0.5168)	(0.6905)	(1.2717)	(1.3766)	(1.3674)	(1.3853)
Death x density (log)		-0.0088	-0.0085	-0.0152	-0.2081	-0.3143	-0.3552	-0.3526	-0.3033
		(0.1039)	(0.1037)	(0.0968)	(0.1409)	(0.2876)	(0.3072)	(0.3063)	(0.3160)
Corruption			0.0960	0.1369	-0.0843	0.0395	0.1609	0.1991	0.1935
			(0.4203)	(0.4318)	(0.5062)	(0.5417)	(0.5639)	(0.5861)	(0.5773)
GDP/capita				-0.0000**	-0.0000*	-0.0000	-0.0000	-0.0000	-0.0000
				(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Democracy					-0.0019	-0.0039	-0.0027	-0.0027	-0.0027
					(0.0040)	(0.0038)	(0.0044)	(0.0043)	(0.0044)
Income tax						-0.1414	-0.1609*	-0.1229	-0.1025
						(0.0912)	(0.0919)	(0.0997)	(0.1017)
Inflation rate							0.0388**	0.0346**	0.0346**
							(0.0158)	(0.0156)	(0.0153)
Agriculture/labor								0.0527	0.0505
								(0.0603)	(0.0719)
Agriculture/GDP									0.0066
									(0.0536)
Constant	45.5378***	45.5241***	45.5193***	45.9781***	46.0607***	47.1845***	47.1009***	45.4191***	45.2490***
	(0.0064)	(0.0064)	(0.0233)	(0.2496)	(0.2381)	(0.6764)	(0.7149)	(2.1208)	(1.8884)
Observations	2,527	2,508	2,508	2,491	2,234	1,940	1,833	1,830	1,801
R-squared	0.0009	0.0009	0.0009	0.0070	0.0057	0.0071	0.0111	0.0144	0.0135
Number of ifscode	187	184	184	182	158	148	142	141	139

Table 6 Contemporaneous effect of epidemics dummy and epidemics deaths per 100,000 population (logs) on the Gini coefficient (Fixed Effects estimation). Notes: Standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

VARIABLES	(1) gini	(2) gini	(3) gini	(4) gini	(5) gini	(6) gini	(7) gini	(8) gini	(9) gini
Epidemics dummy	0.3445**	0.3805***	0.3064***	0.3044***	0.2583***	0.2630**	0.3458***	0.3513***	0.2651**
	(0.1424)	(0.1377)	(0.0906)	(0.0892)	(0.0846)	(0.1144)	(0.1273)	(0.1279)	(0.1281)
Death by 100k (log)	-0.6869**	-1.2242**	0.0655	0.0935	0.9027	1.9634*	2.1480*	2.1135*	2.1053*
	(0.2761)	(0.5570)	(0.4353)	(0.4143)	(0.6258)	(1.1847)	(1.2747)	(1.2691)	(1.2696)
Death x density (log)		0.1112	-0.0037	-0.0077	-0.1671	-0.3453	-0.3905	-0.3870	-0.3313
		(0.1204)	(0.0828)	(0.0783)	(0.1255)	(0.2654)	(0.2825)	(0.2822)	(0.2869)
Corruption			0.2468	0.2830	0.1160	0.2568	0.3418	0.3874	0.3502
			(0.4503)	(0.4650)	(0.5426)	(0.5886)	(0.6206)	(0.6414)	(0.6352)
GDP/capita				-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
				(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Democracy					-0.0022	-0.0050	-0.0039	-0.0038	-0.0039
					(0.0039)	(0.0037)	(0.0043)	(0.0042)	(0.0043)
Income tax						-0.1607	-0.1750	-0.1357	-0.1174
						(0.1041)	(0.1069)	(0.1145)	(0.1196)
Inflation rate							0.0328**	0.0282*	0.0283**
							(0.0147)	(0.0146)	(0.0142)
Agriculture/labor								0.0516	0.0517
								(0.0546)	(0.0649)
Agriculture/GDP									-0.0011
									(0.0499)
Constant	45.5774***	45.5618***	45.5409***	45.8719***	45.9427***	47.2228***	47.1410***	45.4872***	45.3637***
	(0.0115)	(0.0114)	(0.0234)	(0.2406)	(0.2247)	(0.7558)	(0.8131)	(2.0144)	(1.8638)
Observations	2,679	2,658	2,650	2,631	2,363	2,050	1,940	1,936	1,906
R-squared	0.0071	0.0077	0.0009	0.0041	0.0027	0.0067	0.0096	0.0132	0.0121
Number of ifscode	187	184	184	182	158	148	142	141	140

time, inequality itself may be acting as a multiplier on the coronavirus's spread and deadliness."

This implies an issue of reverse causality. To deal with potential endogeneity, we use a system of simultaneous equations that can be jointly estimated using Three-Stage-Least-Squares (3SLS). Simultaneous equations are a statistical model in which the dependent variables are functions of other dependent variables, rather than just independent variables. In our context, both the Gini coefficient and the epidemic and pandemic indicator can be determined jointly as follows

$$Gini_{it} = \alpha + \beta EpiPan_{it-1} + X_{it-1}\gamma + u_{it}$$

$$EpiPan_{it} = \lambda + \tau Gini_{it-1} + Z_{it-1}\rho + e_{it} \qquad (13)$$

where Z_{it-1} is a vector of control variables identified by the literature as determinants of an epidemic or a pandemic in country i in year $t-1$. This vector includes population density, urbanization, democracy and individualism. The identification strategy is based on the intuition that a high level of population density in a country can contribute to a faster spread of a disease causing it to turn into an epidemic, and can also exacerbate the health effects of these epidemics. Population density is captured by the number of people per square kilometer of land area, derived from the World

Development Indicators. Infectious diseases also spread faster in urban concentrations compared to sparsely-populated rural areas. Thus, the urbanization rate is expected to have a positive association with the epidemics and pandemics indicators. Urbanization is measured by the urban population as a percentage of total population, derived from the World Development Indicators. We also include our indicator for democracy, the Polity score. Democratic governments are more likely to be held accountable for how they deal with health crises and accordingly can be more proactive in their attempt to contain the spread of infectious diseases. On the other hand, authoritarian countries tend to cover up the spread of a disease, which makes them more impotent to deal with an epidemic once out of control. Finally, we include the individualism score from the Geert Hofstede dataset.[9] This is because those in collectivist cultures are more likely to sacrifice their personal freedom to follow the rules imposed for the common good of the group. While those in individualistic cultures value more their individual freedoms and personal rights even in the face of a calamity that requires the collective involvement of every individual to be able to deal with it. Thus, we expect that individualism to have a positive association

---

[9]https://geerthofstede.com/research-and-vsm/dimension-data-matrix/

Table 7 Contemporaneous effect of epidemics dummy and epidemics deaths per 100,000 population (logs) on the Gini coefficient (Fixed Effects estimation without Ebola). Notes: Standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

VARIABLES	(1) giny	(2) giny	(3) giny	(4) giny	(5) giny	(6) giny	(7) giny	(8) giny	(9) giny
Epidemics dummy	0.3554**	0.3950***	0.3095***	0.3071***	0.2365***	0.2630**	0.3458***	0.3513***	0.2651**
	(0.1457)	(0.1393)	(0.0922)	(0.0906)	(0.0819)	(0.1144)	(0.1273)	(0.1279)	(0.1281)
Death by 100k (log)	-0.7159**	-1.3042**	0.1881	0.2227	1.3134**	1.9634*	2.1480*	2.1135*	2.1053*
	(0.2890)	(0.5433)	(0.4920)	(0.4671)	(0.6640)	(1.1847)	(1.2747)	(1.2691)	(1.2696)
Death x density (log)		0.1206	-0.0235	-0.0286	-0.2239*	-0.3453	-0.3905	-0.3870	-0.3313
		(0.1182)	(0.0921)	(0.0870)	(0.1329)	(0.2654)	(0.2825)	(0.2822)	(0.2869)
Corruption			0.2464	0.2825	0.1163	0.2568	0.3418	0.3874	0.3502
			(0.4503)	(0.4651)	(0.5428)	(0.5886)	(0.6206)	(0.6414)	(0.6352)
GDP/capita				-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
				(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Democracy					-0.0021	-0.0050	-0.0039	-0.0038	-0.0039
					(0.0039)	(0.0037)	(0.0043)	(0.0042)	(0.0043)
Income tax						-0.1607	-0.1750	-0.1357	-0.1174
						(0.1041)	(0.1069)	(0.1145)	(0.1196)
Inflation rate							0.0328**	0.0282*	0.0283**
							(0.0147)	(0.0146)	(0.0142)
Agriculture/labor								0.0516	0.0517
								(0.0546)	(0.0649)
Agriculture/GDP									-0.0011
									(0.0499)
Constant	45.5772***	45.5616***	45.5403***	45.8716***	45.9419***	47.2228***	47.1410***	45.4872***	45.3637***
	(0.0116)	(0.0114)	(0.0234)	(0.2406)	(0.2248)	(0.7558)	(0.8131)	(2.0144)	(1.8638)
Observations	2,679	2,658	2,650	2,631	2,363	2,050	1,940	1,936	1,906
R-squared	0.0073	0.0081	0.0009	0.0042	0.0029	0.0067	0.0096	0.0132	0.0121
Number of ifscode	187	184	184	182	158	148	142	141	140

with the spread of infectious diseases which can transform them into epidemics or pandemics. In our estimation, we include an interaction term between population density and individualism.

In this context, we do not consider the effect of the epidemic dummy since this can complicate the estimation when a dummy is the dependent variable in the second equation. The results of the Three-Stage-Least-Squares estimation are included in Table 8 when our variable of interest is the number of deaths per 100,000 population. In the table, column 1 shows the results of the first equation where the Gini coefficient is our dependent variable while column 2 shows the results of the second equation where the health indicator is the dependent variable. The results show that the health indicator does not have a statistically significant effect on the Gini coefficient, while income inequality has a statistically significant negative effect on the health indicator in the 2020 sample. The results also show that the interaction term is not statistically significant either.

We also conduct the Three-Stage-Least-Squares estimation to examine the contemporaneous effect in Table 9, with and without Ebola. This is to examine the effect of COVID-19 given the data limitations. Columns 1 and 2 include the results with Ebola, while columns 3 and 4 include the results without Ebola. In each case,

the first column shows the results of the first equation where the Gini coefficient is our dependent variable while the second column shows the results of the second equation where the health indicator is the dependent variable. The results show that the deaths per 100,000 population has a positive effect on the Gini coefficient. This effect is statistically significant and economically important as can be seen from the magnitude of the coefficient.

This implies that before COVID-19 it was the policy response to these health crises that mattered in its effect on income inequality, When we include the effect of COVID-19, and addressing potential endogeneity, the health consequences of these epidemics and pandemics also matter for income inequality. This suggests that not only the occurrence, captured by the dummy variable, but also the health consequences of COVID-19, captured by the number of deaths, have a significant effect on income inequality.

Discussion
Our findings are consistent with other recent research. For example, years spent in poverty were estimated for the pandemic: it had generated at least 68 million additional poverty years in 150 countries [8]. That result is consistent with within-country inequality, even if it may not be consistent with cross-country global income inequality [7]. Similarly, research from Italy concludes that

Table 8 Effect of epidemic deaths per 100,000 population (logs) on the Gini Coefficient (Three-Stage-Least-Squares). Notes: Standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

2020		
	(1)	(2)
LAGGED VARIABLES	gini	deaths per 100k (log)
Deaths per 100k (log)	4.4368	
	(3.4414)	
Deaths x density (log)	-0.9086	
	(0.7823)	
Corruption	0.6923**	
	(0.3084)	
GDP/capita	-0.0001***	
	(0.0000)	
Inflation rate	0.0124	
	(0.0258)	
Agriculture/GDP	-0.4492***	
	(0.0715)	
Agriculture/Labor	0.0560*	
	(0.0291)	
Democracy	0.1045**	0.0004
	(0.0415)	(0.0027)
Income tax	0.2157***	
	(0.0467)	
gini		-0.0047**
		(0.0021)
Density (log)		0.0089
		(0.0082)
Urbanisation		0.0021***
		(0.0007)
Individualism		-0.0007
		(0.0005)
Constant	47.0497***	0.1017
	(0.6392)	(0.1145)
Observations	977	977
R-squared	0.1706	0.0137

the positive shift in working from home (WFH) feasibility is associated with an increase in average labor income, yet the benefit is not equally distributed among employees: the increased opportunity of WFH favors male, older, high-educated, and high-paid employees [4].

Our 3SLS approach allows for causal interpretation from epidemics and pandemics to income inequality. This is important because the reverse association has also been documented for OECD countries: A 1% increase in the Gini coefficient is associated with a 4% increase in COVID-19 cases per-million and a 5% increase in deaths per-million [25]. Similar results were found for

French departments, where the authors conclude "inequality kills: departments with higher income inequality experience more deaths, more discharged (gravely ill) patients and more infections" [13]. And similarly for U.S. counties, where income inequality increases the infection rate, yet the effects are largely accounted by the correlation with racial composition [6].

Future research should consider the structural factors leading to the observed results. This is particularly pertinent to areas of high income inequality, such as Latin America, where recent research found that most of the dispersion in labor income loss across countries is

Table 9 Contemporaneous effect of epidemic deaths per 100,000 population (logs) on the Gini coefficient (Three-Stage-Least-Squares with and without Ebola). Notes: Standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

with Ebola			without Ebola		
VARIABLES	(1) Gini	(2) Deaths per 100k (log)	VARIABLES	(1) Gini	(2) Deaths per 100k (log)
Deaths per 100k (log)	55.2767*		Deaths per 100k (log)	55.2767*	
	(31.6694)			(31.6694)	
Corruption	0.6787**		Corruption	0.6787**	
	(0.3418)			(0.3418)	
GDP/capita	-0.0001***		GDP/capita	-0.0001***	
	(0.0000)			(0.0000)	
Inflation rate	0.0329		Inflation rate	0.0329	
	(0.0337)			(0.0337)	
Agriculture/GDP	-0.3673***		Agriculture/GDP	-0.3673***	
	(0.0812)			(0.0812)	
Agriculture/Labor	0.0460		Agriculture/Labor	0.0460	
	(0.0365)			(0.0365)	
Democracy	0.0941	-0.0006	Democracy	0.0941	-0.0006
	(0.0573)	(0.0011)		(0.0573)	(0.0011)
Income tax	0.2020***		Income tax	0.2020***	
	(0.0779)			(0.0779)	
gini		0.0028	gini		0.0028
		(0.0025)			(0.0025)
Density (log)		-0.0062*	Density (log)		-0.0062*
		(0.0035)			(0.0035)
Urbanisation		-0.0000	Urbanisation		-0.0000
		(0.0003)			(0.0003)
Individualism		0.0003	Individualism		0.0003
		(0.0002)			(0.0002)
Constant	45.6406***	-0.0891	Constant	45.6406***	-0.0891
	(1.4034)	(0.1162)		(1.4034)	(0.1162)
Observations	1,056	1,056	Observations	1,056	1,056
R-squared	-1.5771	-0.0200	R-squared	-1.5771	-0.0200

explained by sectoral/occupational employment structure of the economies [9]. Future research can also include additional data on income inequality to examine the effect of lagged health indicators including COVID-19 on income inequality.

Conclusion

This paper tests a theoretical model that predicts that epidemics and pandemics increase income inequality. To achieve its objective, the paper explores the effect on the Gini coefficient of a dummy variable that indicates the occurrence of an epidemic or a pandemic in a country in a given year and the number of deaths per 100,000 population. The fixed effects estimation shows a robust statistically significant positive effect of the dummy variable but an insignificant effect of the number of deaths per 100,000 population. The interpretation is that it is the occurrence of a pandemic, captured by the dummy variable, that is what triggers the policy response that may affect income inequality, rather than the health consequences of these crises captured by the number of deaths per 100,000.

To properly address potential endogeneity, we implement a Three-Stage-Least Squares estimation technique. The 3SLS estimation shows that the number of deaths per 100,000 population variable has a statistically significant contemporaneous positive effect on the Gini coefficient, especially when we incorporate the data on COVID-19. This suggests that not only the occurrence but also the health consequences of COVID-19 have a significant effect on income inequality.

Authors' contributions
"All authors worked on the different sections of the paper. All authors read and approved the final manuscript.

Declarations

Ethics approval and consent to participate
"Not applicable"

Consent for publication
"Not applicable"

Competing interests
"The authors declare that they have no competing interests"

Author details
[1]International Monetary Fund, Washington, DC, USA. [2]Brown University, Providence, RI 02912, United States. [3]California State University, Fullerton, CA 92834, USA.

References

1. Acemoglu D, Naidu S, Restrepo P, Robinson J. Democracy, redistribution, and inequality. Handb Income Distrib. 2015;2:1885–966. https://doi.org/10.1016/B978-0-444-59429-7.00022-4.
2. Albanesi S. Inflation and inequality. J Monet Econ. 2007;54(4):1088–114. https://doi.org/10.1016/j.jmoneco.2006.02.009.
3. Bell C, Gersbach H. The macroeconomics of targeting: the case of an enduring epidemic. J Health Econ. 2009;28(1):54–72. https://doi.org/10.1016/j.jhealeco.2008.07.011.
4. Bonacini L, Gallo G, Scicchitano S, Zimmermann KF. Working from home and income inequality: risks of a 'new Normal' with COVID-19. J Popul Econ. 2021;34(1):303–60. https://doi.org/10.1007/s00148-020-00800-7.
5. Borjas G, Cassidy H. The adverse effect of the COVID-19 labor market shock on immigrant employment. National Bureau of Economic Research Working Paper. 2020;27243.
6. Brown C, Ravallion M. Inequality and the Coronavirus: Socioeconomic Covariates of Behavioral Responses and Viral Outcomes across US Counties. National Bureau of Economic Research Working Paper. 2020;27549.
7. Deaton A. COVID-19 and Global Income Inequality 2021. National Bureau of Economic Research Working Paper. 28392.
8. Decerf B, Ferreira F, Mahler D, Sterck O. Lives and Livelihoods: Estimates of the Global Mortality and Poverty Effects of the Covid-19 Pandemic. World Development, forthcoming.
9. Delaporte I, Escobar J, Peña W. The distributional consequences of social distancing on poverty and labour income inequality in Latin America and the Caribbean. J Popul Econ. 2021;34(4):1385–443. https://doi.org/10.1007/s00148-021-00854-1.
10. Fenichel E. Economic considerations for social distancing and behavioral based policies during an epidemic. J Health Econ. 2013;32(2):440–51. https://doi.org/10.1016/j.jhealeco.2013.01.002.
11. Forsythe E, Kahn L, Lange F, Wiczer D. Labor demand in the time of COVID-19: evidence from vacancy postings and UI claims. J Public Econ. 2020;2020.
12. Furceri D, Loungani P, Ostry J, Pizzuto P. Will Covid-19 affect inequality? Evidence from past pandemics. Covid Economics. 2020;12:138–57.
13. Ginsburgh V, Magerman G, Natali I. COVID-19 and the role of inequality in French regional departments. Eur J Health Econ. 2021;22(2):311–27. https://doi.org/10.1007/s10198-020-01254-0.
14. Glover A, Heathcote J, Kreuger D, Rios-Rull J-V. Health versus wealth: on the distributional effects of controlling a pandemic. National Bureau of Economic Research Working Paper. 2020;27046.
15. Goldberg P, Pavcnik N. Distributional effects of globalization in developing countries. J Econ Lit. 2007;45(1):39–82. https://doi.org/10.1257/jel.45.1.39.
16. Lin Z, Meissner C. Health vs. wealth? Public health policies and the economy during Covid-19. National Bureau of Economic Research Working Paper. 2020;27099.
17. Montenovo L, Jiang X, Rojas F, Schmutte I, Simon K, Weinberg B, et al. Determinants of disparities in Covid-19 job losses. National Bureau of Economic Research Working Paper. 2020;27132.
18. Moser C, Yared P. Pandemic Lockdown: The Role of Government Commitment. National Bureau of Economic Research Working Paper. 2021; 27062.
19. Palomino J, Rodriguez J, Sebastian R. Wage inequality and poverty effects of lockdown and social distancing in Europe. Eur Econ Rev. 2020;129:1–25. https://doi.org/10.1016/j.euroecorev.2020.103564.
20. Papanikolaou D, Schmidt L. Working remotely and the supply-side impact of Covid-19. National Bureau of Economic Research Working Paper. 2020; 27330.
21. Pedersen K. Rent-seeking, political influence and inequality: a simple analytical example. Public Choice. 1995;82(3-4):281–305. https://doi.org/10.1007/BF01047698.
22. Rojas F, Jiang X, Montenovo L, Simon K, Weinberg B, Wing C. Is the cure worse than the problem itself? Immediate labor market effects of COVID-19 case rates and school closures in the U.S. National Bureau of Economic Research Working Paper. 2020;27127.
23. Spinesi L. Rent-seeking bureaucracies, inequality, and growth. J Dev Econ. 2009;90(2):244–57. https://doi.org/10.1016/j.jdeveco.2008.09.009.
24. Sumner A, Hoy C, Ortiz-Juarez E. Estimates of the impact of COVID-19 on global poverty. WIDER Working Paper. 2020;43. https://doi.org/10.35188/UNU-WIDER/2020/800-9.
25. Wildman J. COVID-19 and income inequality in OECD countries. Eur J Health Econ. 2021;22(3):455–62. https://doi.org/10.1007/s10198-021-01266-4.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.