



Will COVID-19 Have Long-Lasting Effects on Inequality? Evidence from Past Pandemics

Davide Furceri^{1,2} · Prakash Loungani^{1,3} · Jonathan D. Ostry⁴ · Pietro Pizzuto²

Received: 14 July 2021 / Accepted: 23 March 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

This paper provides evidence on the impact of major epidemics from the past two decades on income distribution. The pandemics in our sample, even though much smaller in scale than COVID-19, have led to increases in the Gini coefficient, raised the income share of higher-income deciles, and lowered the employment-to-population ratio for those with basic education compared to those with higher education. We provide some evidence that the distributional consequences from the current pandemic may be larger than those flowing from the historical pandemics in our sample, and larger than those following typical recessions and financial crises.

Keywords COVID-19 · Pandemics · Inequality

JEL Codes E52 · E58 · D43 · L11

The views expressed in this paper are those of the authors and do not necessarily represent those of the IMF or its member countries. A previous draft of the paper was published in Covid Economics with the title “Will COVID-19 Affect Inequality? Evidence from Past Pandemics”.

✉ Davide Furceri
dfurceri@imf.org

Prakash Loungani
ploungani@imf.org

Jonathan D. Ostry
jostry@imf.org

Pietro Pizzuto
pietro.pizzuto02@unipa.it

¹ International Monetary Fund, RCEA, 700 19th Street NW, Washington, DC 20431, USA

² University of Palermo, Palermo, Italy

³ Johns Hopkins University, Baltimore, MD, USA

⁴ CEPR, Washington, DC, USA

1 Introduction

As of February 15, 2022, deaths from the COVID-19 pandemic have reached about 5.8 million worldwide according to official statistics. This tragic cost has been accompanied by the upending of millions of other lives as governments take necessary steps to limit the spread of the virus. For instance, at the beginning of 2021, the International Labor Organization (ILO) estimated an unprecedented worldwide loss of 255 million jobs as a result of the COVID-19 pandemic, with the unemployment rate rising by 1.1 percentage points to 6.5%, and 81 million workers pushed out of the labor market (ILO 2021), far more than were lost over the entire Great Recession of 2008–09 (for a comparison in the case of the United States see, for example, Coibion and Weber 2020).

While most income groups are adversely affected by the pandemic, it is possible that lower-income deciles and those with lower skills end up being disproportionately hurt. Indeed, there is already evidence of such effects, raising the prospect at least of a persistent increase in inequality in the absence of forceful policy interventions. Using data from a large-scale survey of U.K. households, Crossley et al. (2020) show that those in the lowest quintiles of income and those from minority ethnic groups have experienced the largest job losses. Similarly, using transaction data from a large Fintech company, Hacıoglu et al. (2020) and Surico et al. (2020) document a surge in market income inequality in the United Kingdom since the beginning of the COVID-19 crisis. Aspachs et al. (2020), using high-frequency data on bank records, wages and public transfers for Spain, provide evidence of increasing income inequality due to severe job losses for low-income households. Additional preliminary evidence (see Stantcheva 2021 and references cited therein) for selected countries (mostly in the EU) suggests a regressive effect caused by the COVID-19 pandemic outbreak. The increasing effect on market Gini ranges from about 0.7% (Italy) to 20% (Ireland) with the short-term policy support provided in response to the crisis, more than offsetting the negative distributional effects caused by the pandemic.¹

The socio-economic impact of the pandemic, moreover, is not limited to income-related losses. Blundell et al. (2020) documents adverse effects on health, education, labor market access and other socio-demographic indicators in United Kingdom. In the case of Finland, the situation room report by the Helsinki (2021) shows striking increases in unemployment benefits applications especially in April and May 2020, with the number of applications falling clearly below the levels of 2020 and 2019 only in the final months of 2021. Similarly, using survey data, Aucejo et al. (2020) show that the pandemic is widening achievement gaps in higher education, with lower-income students being 55% more likely than their higher-income peers to delay graduation. There are also direct and immediate effects from lower-income groups being more prone to the disease: Schmitt-Grohé et al. (2020) find that, in New York City, poor people are less likely to test negative for COVID-19: moving from the richest to the poorest zip codes is associated with a decline in the fraction of negative test results from 65 to 38%.

To shed light on possible medium-term distributional impacts of COVID-19, this paper uses data from major epidemics (referred to interchangeably below as pandemics) over the past two decades and their links to: income inequality; income shares of the top and bottom deciles; and employment prospects of people with low education levels (using educational attainment

¹ Stantcheva (2021) argues that the inequality is likely to rise in the medium term due to the broad adoption of remote work.

as a proxy for skills). The results suggest that pandemics have a negative causal effect on income distribution: past pandemics, even though much smaller in scale, have led to increases in the Gini coefficient, raised the income shares of higher-decile income groups, and lowered the employment-to-population ratio of those with basic education compared to those with higher education. Our evidence suggests that the distributional consequences from the current pandemic are likely to be significantly larger than those flowing from the average historical pandemic in our sample, and larger than those following typical recessions and financial crises. In this context, our results can also directly inform us about the distributional consequences of potential future epidemics.

This paper relates to two main strands of literature. The first is the literature on the economic effects of pandemics: Atkeson 2020; Barro et al. 2020; Eichenbaum et al. 2020; Jordà et al. 2020; Ma et al. 2020a. This literature provides evidence of large and persistent effects on economic activity from pandemics. Ma et al. (2020a) examined the same set of episodes used here and found that real GDP is 2.6% lower on average across 210 countries in the year the outbreak is officially declared and remains 3% below the pre-shock level five years later. The second strand of the literature relates to the effects of crises and recessions on inequality and employment including of the less skilled and youth: Camacho and Palmieri 2019; de Haan and Sturm 2017.

The remainder of the paper is structured as follows. Section II describes our data and econometric method and Section III presents our results. The last section concludes and outlines avenues for future work on this topic.

2 Data and Econometric Method

2.1 Income Distribution

Our data on various measures of distribution come from three sources. Table 1 provides summary statistics on the variables used in the analysis.

- Gini coefficients are from the Standardized World Income Inequality Database (SWIID 8.3), which combines information from the Luxembourg Income Study (LIS) and other sources (such as the OECD Income Distribution Database, the Socio-Economic Database

Table 1 Data sources and descriptive statistics

Variable	Source	Obs	Mean	Std. Dev.	No. of Countries
Gini Market	SWIID 8.3	5472	45.39	6.59	177
Gini Net	SWIID 8.3	5472	38.38	8.73	177
Top 40% Income Share	WDI	1444	67.77	6.65	64
Top 20% Income Share	WDI	1444	46.28	7.83	64
Top 10% Income Share	WDI	1444	30.85	7.31	64
Bottom 40% Income Share	WDI	1444	17.12	4.56	64
Bottom 20% Income Share	WDI	1444	6.31	2.19	64
Bottom 10% Income Share	WDI	1443	2.44	1.02	64
<i>Employment/Population (E/P) ratios</i>					
E/P ratio – Basic Education	ILO	1340	42.51	16.22	76
E/P ratio – Non-basic Education	ILO	1340	57.49	16.22	76
Financial Crises	Laeven and Valencia (2020)	289 episodes			177

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). SWIID provides comparable estimates of market (pre-tax, pre-transfer) and net (post-tax, post-transfer) income inequality using LIS survey data as the standard and adopting a multiple imputation procedure to offer the widest possible coverage across countries and over time (Solt 2009). Our sample includes 177 countries from 1960 to the present.²

- Income shares by decile are from the World Bank's World Development Indicators. Data are based on primary household survey data obtained from government statistical agencies and World Bank country departments and includes both labor (salaries, own-business, and self-employment income) as well as non-labor incomes. This source provides internationally comparable statistics for a large number of economies; however, for many countries the time series is rather short, so in the end our results on income deciles are for a limited sample of 64 countries from 1981 to the present.
- Comparable data on employment by skill levels are difficult to obtain for a large group of countries. The ILO notes that "statistics on levels of educational attainment remain the best available indicators of labor force skill levels." Hence, we use ILO data on employment-to-population ratios for different education levels—advanced, tertiary and basic—for a limited sample of 76 countries from 1990 to the present.

2.2 Pandemic Events

As in Ma et al. 2020a, we focus on five major events: SARS in 2003; H1N1 in 2009; MERS in 2012; Ebola in 2014; and Zika in 2016. The countries affected by each event are presented in Table 2 and Table A1 in the Appendix (we exclude countries for which income inequality data are unavailable). We construct a dummy variable, the pandemic event, which takes the value 1 when the WHO declares a pandemic for the country and 0 otherwise. Our baseline results estimate the evolution of inequality in the aftermath of the pandemic event. However, we also take account of how the severity of the pandemic affects distributional outcomes. The most widespread pandemic in our sample is H1N1 (Swine Flu Influenza), with more than 6,000,000 confirmed cases across 148 countries (about 1 case per thousand people) and about 19,000 fatalities. While H1N1 spread across all regions, the other four events are mostly confined to specific regions: (i) SARS and MERS in Asia; (ii) Ebola in Africa and (iii) Zika in the Americas (Fig. 1). In terms of average mortality rates (deaths/confirmed cases), MERS and Ebola were the most severe (around 35%), followed by SARS and H1N1.

For the sake of comparison, as of February 2022, COVID-19 infections were confirmed in 227 countries, areas, or territories with more than 400 million confirmed cases (about 50 cases per thousand people) and a total mortality rate a bit larger than H1N1 (about 1.40%). The median country in terms of cases to population ratio for COVID-19 is about 30 cases per 1000 inhabitants—roughly corresponding to twice the severity of a pandemic episode at the 99th percentile of the severity distribution in our sample.³

² Income aggregates comprises labor and non-labor incomes. We use data from SWIID as baseline because of the larger country and time coverage compared to other commonly-used sources, such as WIDER and POVCAL. In the robustness checks, we show that our results hold when using data from these alternative sources.

³ For COVID-19 data see <https://www.who.int/emergencies/diseases/novel-coronavirus-2019> (accessed on February 15, 2022)

Table 2 List of pandemic and epidemic episodes

Starting year	Announced month	Event Name	Affected Countries	Number of countries	Total Deaths	Total Cases	Total Mortality rate (%)	Average Cases/Pop (*100,000)	Average Mortality rate (%)
2003	2	SARS	AUS, CAN, CHE, CHN, DEU, ESP, FRA, GBR, HKG, IDN, IND, IRL, ITA, KOR, MNG, MYS, NZL, PHL, ROU, RUS, SGP, SWE, THA, TWN, USA, VNM, ZAF	27	774	8094	9.56	1.25	9.77
2009	4	H1N1	AFG, AGO, ALB, ARG, ARM, AUS, AUT, BDI, BEL, BGD, BGR, BHS, BIH, BLR, BLZ, BOL, BRA, BRB, BTN, BWA, CAN, CHE, CHL, CHN, CIV, CMR, COD, COG, COL, CPV, CRI, CYP, CZE, DEU, DJI, DMA, DNK, DOM, DZA, ECU, EGY, ESP, EST, ETH, FIN, FJI, FRA, FSM, GAB, GBR, GEO, GHA, GRC, GTM, HND, HRV, HTI, HUN, IDN, IND, IRL, IRN, IRQ, ISL, ISR, ITA, JAM, JOR, JPN, KAZ, KEN, KHM, KNA, KOR, LAO, LBN, LCA, LKA, LSO, LTU, LUX, LVA, MAR, MDA, MDG, MDV, MEX, MKD, MLI, MLT, MNE, MNG, MOZ, MUS, MWI, MYS, NAM, NGA, NIC, NLD, NOR, NPL, NZL, PAK, PAN, PER, PHL, PLW, PNG, POL, PRI, PRY, QAT, ROU, RUS, RWA, SAU, SDN, SGP, SLB, SLV, STP, SUR, SVK, SVN, SWE, SWZ, SYC, TCD, THA, TJK, TON, TUN, TUR, TUV, TZA, UGA, UKR, URY, USA, VEN, VNM, VUT, WSM, YEM, ZAF, ZMB, ZWE	149	19,091	6,502,779	0.29	122.69	4.02
2012	3	MERS	AUT, CHN, DEU, EGY, FRA, GBR, GRC, IRN, ITA, JOR, KOR, LBN, MYS, NLD, PHL, QAT, SAU, THA, TUN, TUR, USA, YEM	22	572	1453	39.37	0.24	35.95
2014	8	Ebola	ESP, GBR, ITA, LBR, SLE, USA	6	8767	24,809	35.34	74.37	16.34
2016	2	Zika	ARG, BOL, BRA, BRB, CAN, CHL, COL, CRI, DOM, ECU, HND, JAM, LCA, PAN, PER, PRI, PRY, SLV, SUR, URY, USA	21	20	198,122	0.01	76.21	0.03
225									
Total Pandemic and Epidemic Events									

Sources: WHO, Ma and others (2020); ECDC, CDC; PAHO; Wikipedia. Information in the table refers to countries for which data on Net Gini are available (i.e. for Ebola not all countries affected by the epidemic event are included in our analysis due to data constraints). The sources of the number of cases/deaths are as follows (accessed on June 24, 2020). Data on Population are from the World Bank's World Development Indicator Database

SARS: https://www.who.int/csr/sars/country/table2004_04_21/en/;

H1N1: https://en.wikipedia.org/wiki/2009_swine_flu_pandemic_by_country_and_https://www.ecdc.europa.eu/en/seasonal-influenza/2009-influenza-h1n1;

MERS: <https://www.ecdc.europa.eu/en/news-events/epidemiological-update-middle-east-respiratory-syndrome-coronavirus-mers-cov-1-0>;

EBOLA: <https://www.cdc.gov/vhf/ebola/history/2014-2016-outbreak/index.html>;

ZIKA: https://www.paho.org/hq/index.php?option=com_content&view=article&id=12390:zika-cumulative-cases&Itemid=42090&lang=en

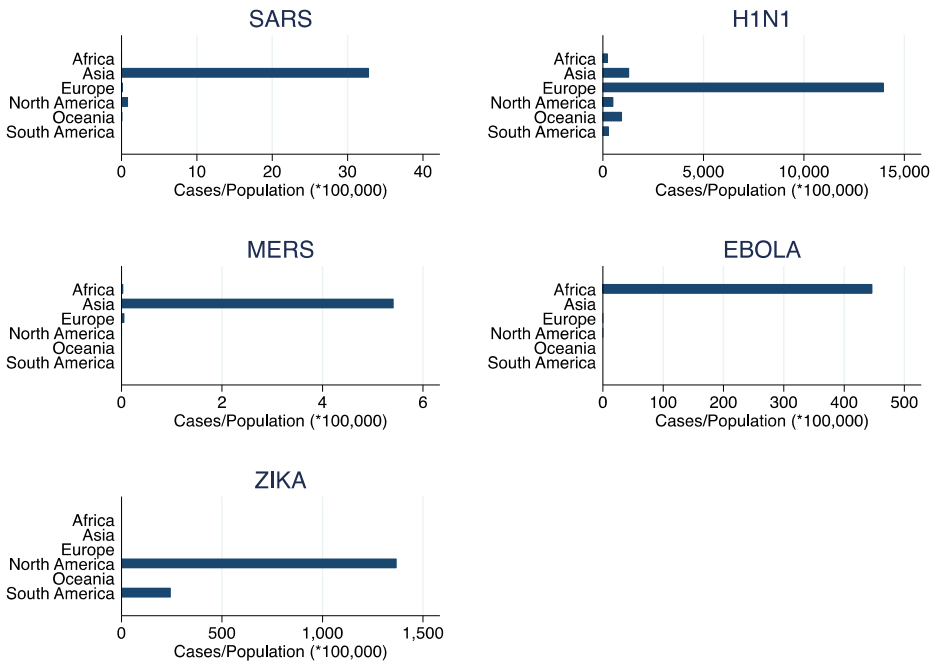


Fig. 1 Pandemic cases by regions. Sources: WHO, Ma and others (2020); ECDC, CDC; PAHO; Wikipedia. Information in the table refers to countries for which data on Net Gini are available (i.e. for Ebola not all countries affected by the epidemic event are included in our analysis due to data constraints). The sources of the number of cases/deaths are as follows (accessed on June 24, 2020). Data on Population are from the World Bank's World Development Indicator Database. SARS: https://www.who.int/csr/sars/country/table2004_04_21/en/; H1N1: https://en.wikipedia.org/wiki/2009_swine_flu_pandemic_by_country and <https://www.ecdc.europa.eu/en/seasonal-influenza/2009-influenza-h1n1>; MERS: <https://www.ecdc.europa.eu/en/news-events/epidemiological-update-middle-east-respiratory-syndrome-coronavirus-mers-cov-1-0>; EBOLA: <https://www.cdc.gov/vhf/ebola/history/2014-2016-outbreak/index.html>; ZIKA: https://www.paho.org/hq/index.php?option=com_content&view=article&id=12390:zika-cumulative-cases&Itemid=42090&lang=en

2.3 Empirical Methodology

To estimate the distributional impact of pandemics, we follow the method proposed by Jordà (2005) and estimate impulse response functions directly from local projections:

$$y_{i,t+k} = \alpha_i^k + \gamma_t^k + \beta^k D_{i,t} + \theta^k X_{i,t} + \varepsilon_{i,t+k} \quad (1)$$

where $y_{i,t}$ is a distribution variable (e.g. the Gini coefficient) for country i in year t ; α_i are country fixed effects, included to take account of differences in countries' average income distribution; γ_t are time fixed effects, included to take account of global shocks such as shifts in oil prices or the global business cycle; $D_{i,t}$ is a dummy variable indicating a pandemic event in country i at year t . $X_{i,t}$ is a vector that includes two lags of the dependent variable and of the pandemic dummy. In the baseline, we do not include other controls on the grounds that the date of the pandemic event is likely to be exogenous to the economy. Indeed, as shown in Table 4, the dates of pandemic events are uncorrelated with past levels and changes of inequality. Nonetheless, we consider, subsequent to presenting our baseline findings, possible concerns arising from our empirical strategy (i.e. omitted variable bias and reverse causality): we present a wide range of robustness checks including an Augmented Inverse Probability

Weighting (AIPW) estimation as in Jordà and Taylor (2016) and an Instrumental Variable approach.

Equation (1) is estimated for an unbalanced panel of 177 countries over the period 1960–2019 (country and time coverage are dependent on data availability for Gini coefficients), for each horizon (year) $k = 0, \dots, 5$. Impulse response functions are computed using the estimated coefficients β^k , and the confidence bands associated with the estimated impulse-response functions are obtained using the estimated standard errors of the coefficients β^k , based on robust standard errors clustered at the country level.

3 Distributional Impacts of Pandemics

3.1 Impacts on Gini Coefficients

Figure 2 shows the estimated dynamic response of net Gini to a pandemic event over the five-year period following the event, together with the 90% confidence interval around the point estimate. Table 3 reports the associated regressions. Pandemics lead to a persistent increase in inequality with a peak effect of about 0.4 five years after the pandemic—that is an average increase of 1.1%. Given that the Gini is a slow-moving variable, these are quantitatively important effects: peak effects correspond to about a 1½ standard deviation of the average change of the Gini in the sample.⁴

3.2 Robustness Checks

We have carried out several robustness checks of these findings. First, we check the sensitivity of our results to an alternative measure of inequality from SWIID, namely the market Gini. The results in Fig. 3 confirm our main findings suggesting that pandemics also lead to an increase in the Gini index computed before taxes and transfers.

Second, we check the sensitivity of our results to measures of inequality taken from alternative databases. Although the SWIID database allows one of the widest possible coverage across countries and over time, such Gini data can be affected by measurement errors since they are based on a multiple imputations method, rather than relying only on survey data (major critiques relate to data comparability, quality and also, about the imputation model; see Jenkins 2015 and Chapter 2 in Ostry et al. 2019, for a broader discussion of the pros and cons of SWIID data set relative to others). We use Ginis from the World Bank POVCAL database—which covers the period 1978–2017 and includes 171 countries (1711 observations)—and the World Institute for Development Research WIDER (WIID) dataset—which covers the period 1948–2014 and includes 166 countries (1386 observations).⁵ Reassuringly, the results in Fig. 4 confirm our main findings. In particular, results based on the POVCAL/WIDER datasets point to even higher medium-term effects: about 1.5–2.0—Gini points and statistically significant at 1 and 5% level, respectively.

⁴ The Gini coefficient on net income has increased cumulatively by about 10% in the US during the period 1980–2010 (from about 0.45 in 1980 to about 0.5 in 2010—see, among others, Coibion et al. 2017).

⁵ Data are taken from the All the Ginis (ALG) Database (<https://stonecenter.gc.cuny.edu/research/all-the-ginis-alg-dataset-version-february-2019/>).

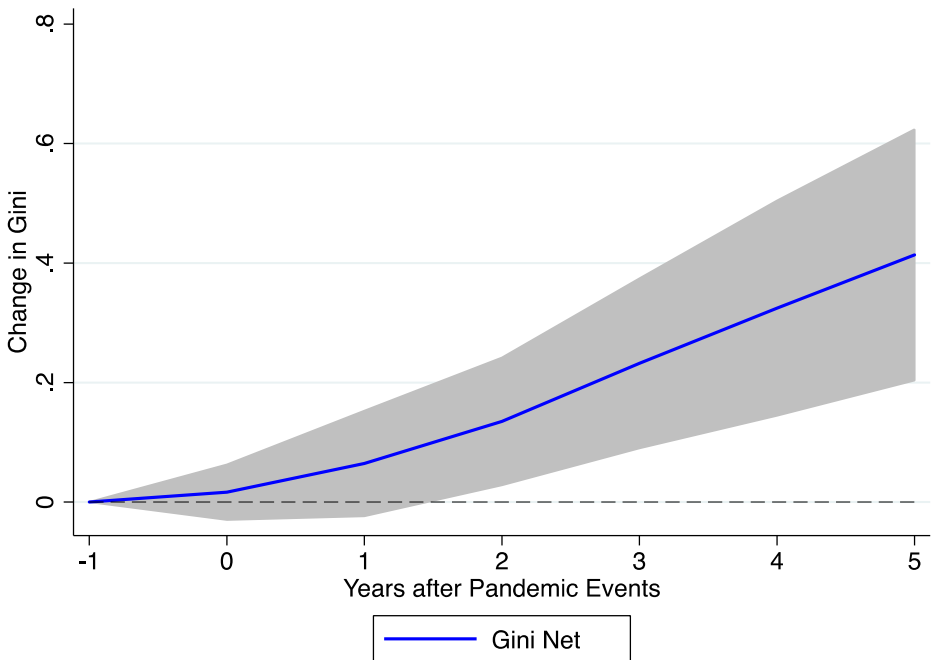


Fig. 2 Impact of pandemics on net Gini. Notes: Impulse response functions are estimated using a sample of 177 countries over the period 1960–2019. The graph shows the response and 90% confidence bands. The x-axis shows years (k) after pandemic events; $t = 0$ is the year of the pandemic event. Estimates based on $y_{i,t+k} = \alpha_i^k + \gamma_t^k + \beta^k D_{i,t} + \theta^k X_{i,t} + \varepsilon_{i,t+k}$. $y_{i,t}$ is the Gini coefficient for country i in year t ; α_i are country fixed effects; γ_t are time fixed effects; $D_{i,t}$ is a dummy variable indicating a pandemic event that affects country i in year t . $X_{i,t}$ is a vector that includes two lags of the dependent variable and two lags of the pandemic dummy. See Table A1 for the full list of pandemic events

Third, as an alternative empirical strategy, we present results from the autoregressive distributed lag (ADL) approach of Romer and Romer (2010) and Furceri et al. (2019). Fourth, since the episodes in our sample occurred in the latest two decades, we replicate the analysis using a restricted sample that begins in 1990. Fifth, in order to mitigate omitted variable bias, we include several control variables that could be related to inequality—such as proxies for the level of economic development, demographics, measures of trade and financial globalization and country-specific time trends. The results presented in Figs. 5, 6, 7 and 8 are similar to, and not statistically different from, the baseline.

Sixth, we check the sensitivity of our results to the fact that some countries are affected by several pandemics over time and the periods following each episode sometimes overlap. As in Teulings and Zubanov (2014), we re-estimate our model with a correction that augments the local projection regression with dummy variables for pandemics occurring within the forecast horizon, i.e. between t and $t + k$. In addition, in order to further isolate the causal effect of pandemics, we define a new dummy variable that takes a value of 1 only for the first pandemic affecting a specific country (i.e. if a country was affected both by H1N1 and MERS the dummy takes the value 1 only in the first case). Reassuringly, Fig. 9 shows that, with both of these adjustments, the results remain very similar to, and not statistically different from, the baseline.

We also checked the validity of the *parallel trend assumption*—that is, the assumption that the inequality in the treatment and counterfactual were following a parallel trend before the

Table 3 Impact of pandemics on market Gini and net Gini coefficients

Panel A: Net Gini						
	k=0	k=1	k=2	k=3	k=4	k=5
$D_{i,t}$	0.017 (0.028)	0.065 (0.054)	0.135** (0.065)	0.232*** (0.086)	0.325*** (0.109)	0.414*** (0.128)
$D_{i,t-1}$	0.038 (0.029)	0.105** (0.046)	0.237*** (0.067)	0.309*** (0.100)	0.400*** (0.119)	0.534*** (0.150)
$D_{i,t-2}$	0.037 (0.028)	0.115* (0.060)	0.233*** (0.080)	0.316*** (0.103)	0.346*** (0.130)	0.401** (0.166)
$y_{i,t-1}$	1.629*** (0.029)	2.068*** (0.058)	2.372*** (0.088)	2.510*** (0.127)	2.578*** (0.163)	2.624*** (0.193)
$y_{i,t-2}$	-0.651*** (0.029)	-1.127*** (0.056)	-1.480*** (0.083)	-1.675*** (0.119)	-1.803*** (0.152)	-1.911*** (0.181)
Observations	5110	4933	4756	4579	4403	4228
R ²	1.000	0.998	0.997	0.994	0.992	0.990
Panel B: Market Gini						
	k=0	k=1	k=2	k=3	k=4	k=5
$D_{i,t}$	0.050 (0.036)	0.113** (0.052)	0.182*** (0.070)	0.219** (0.090)	0.314*** (0.119)	0.413*** (0.152)
$D_{i,t-1}$	0.041 (0.034)	0.111** (0.055)	0.203*** (0.078)	0.256** (0.111)	0.335** (0.134)	0.539*** (0.186)
$D_{i,t-2}$	0.040 (0.031)	0.070 (0.061)	0.151* (0.088)	0.214* (0.119)	0.338** (0.155)	0.477** (0.212)
$y_{i,t-1}$	1.608*** (0.031)	2.053*** (0.059)	2.389*** (0.089)	2.582*** (0.120)	2.734*** (0.154)	2.843*** (0.182)
$y_{i,t-2}$	-0.625*** (0.0311)	-1.099*** (0.060)	-1.473*** (0.088)	-1.710*** (0.117)	-1.910*** (0.149)	-2.072*** (0.176)
Observations	5110	4933	4756	4579	4403	4228
R ²	0.999	0.997	0.994	0.989	0.985	0.980

Note: Estimates are obtained using a sample of 177 countries over the period 1960–2019, and based on $y_{i,t+k} = \alpha_i^k + \gamma_t^k + \beta^k D_{i,t} + \theta^k X_{i,t} + \varepsilon_{i,t+k}$. $y_{i,t}$ is the Gini coefficient for country i in year t ; α_i are country fixed effects; γ_t are time fixed effects; $D_{i,t}$ is a dummy variable indicating a pandemic event that affects country i in year t . $X_{i,t}$ is a vector that includes two lags of the dependent variable and two lags of the pandemic dummy. See Table A1 for the full list of pandemic events. Standard errors in parentheses are clustered at the country level. Country and time fixed effects included but not reported

pandemic—in the evolution of inequality before the pandemic between countries by running a placebo test. Reassuringly, the impulse response functions obtained by attributing randomly pandemic dates across the whole sample do not point to significant results (Fig. 10). Similarly, estimations of lags of Gini (i.e., Gini prior to the pandemic outbreak) on the contemporaneous pandemic dummy do not point to significant results as well and suggest that pandemics at time t are not associated with past changes in inequality—that is, Gini was not statistically different between a country affected by the pandemic and a country not-affected, before the occurrence of the pandemic (Table A2).

3.3 Addressing Endogeneity

In order to further address endogeneity, we adapt the approach proposed by Jordà and Taylor (2016) to estimate the causal effect of austerity, and we use the Augmented Inverse Probability weighting (AIPW). The rationale of this approach is to address potential endogeneity in the measure of treatment (the pandemic event in our case). Indeed, pandemics may not be fully exogenous events and be related to pre-existing country characteristics. We therefore construct a predictive model for the likelihood of pandemics using various specifications including the

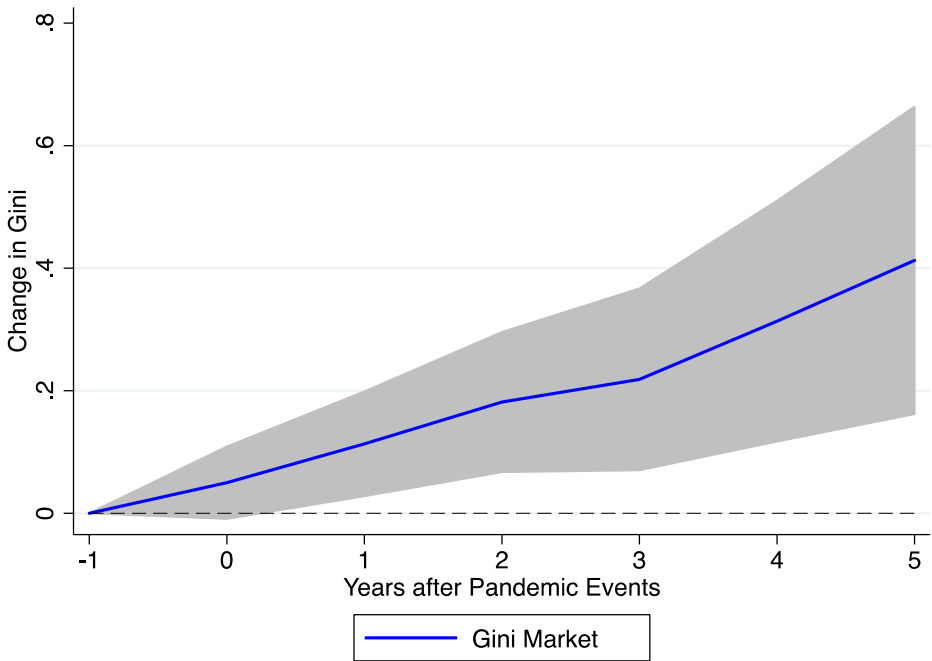


Fig. 3 Impact of pandemics on market Gini. Notes: Impulse response functions are estimated using a sample of 177 countries over the period 1960–2019. The graph shows the response and 90% confidence bands. The x-axis shows years (k) after pandemic events; $t = 0$ is the year of the pandemic event. Estimates based on $y_{i,t+k} = \alpha_i^k + \gamma_t^k + \beta^k D_{i,t} + \theta^k X_{i,t} + \varepsilon_{i,t+k}$. $y_{i,t}$ is the Gini coefficient for country i in year t ; α_i are country fixed effects; γ_t are time fixed effects; $D_{i,t}$ is a dummy variable indicating a pandemic event that affects country i in year t . $X_{i,t}$ is a vector that includes two lags of the dependent variable and two lags of the pandemic dummy. See Table A1 for the full list of pandemic events

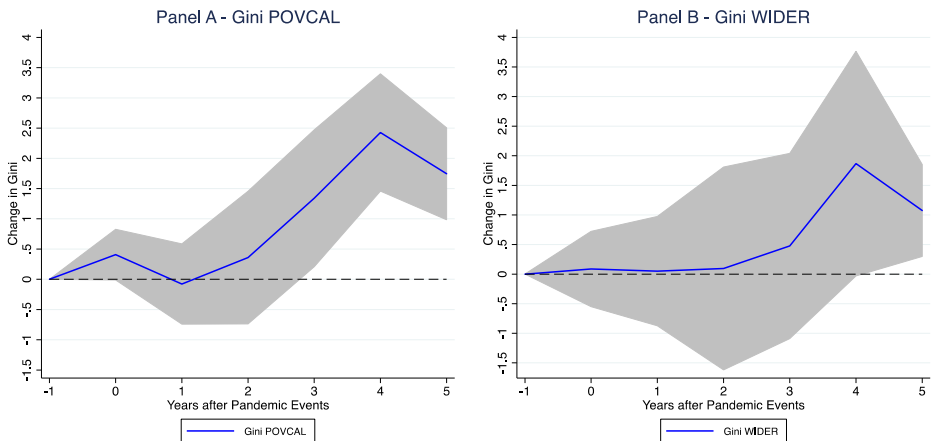


Fig. 4 Impact of pandemics on net Gini—different measures of inequality. Notes: Impulse response functions are estimated using, alternatively, Gini POVCAL (69 countries – 763 observations) and Gini WIDER (66 countries – 641 observations). The graph shows the responses and 90% confidence bands. The x-axis shows years (k) after pandemic events; $t = 0$ is the year of the pandemic event. Estimates based on $y_{i,t+k} = \alpha_i^k + \gamma_t^k + \beta^k D_{i,t} + \theta^k X_{i,t} + \varepsilon_{i,t+k}$. $y_{i,t}$ is the Gini coefficient for country i in year t ; α_i are country fixed effects; γ_t are time fixed effects; $D_{i,t}$ is a dummy variable indicating a pandemic event that affects country i in year t . $X_{i,t}$ is a vector that includes two lags of the dependent variable and two lags of the pandemic dummy

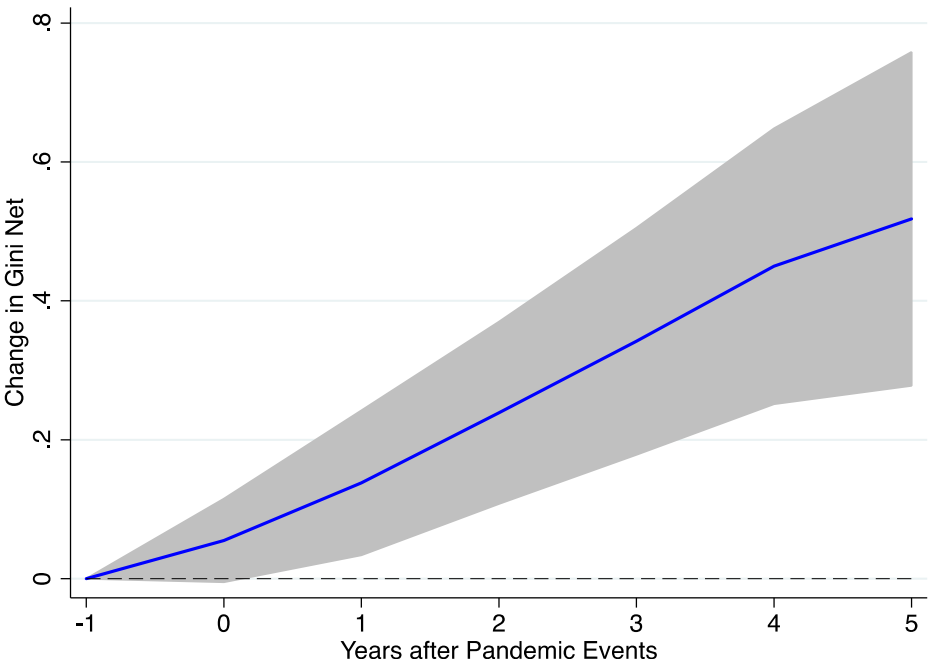


Fig. 5 Impact of pandemics on net Gini—ADL. Notes: Impulse response functions are estimated using a sample of 177 countries over the period 1960–2019. The graph shows the response and 90% confidence bands. The x-axis shows years (k) after pandemic events; $t = 0$ is the year of the pandemic event. Estimates based on $\Delta y_{i,t} = \alpha_i + \gamma_t + \beta_k(l)D_{i,t} + \varepsilon_{i,t}$. $y_{i,t}$ is the Gini coefficient for country i in year t ; α_i are country fixed effects; γ_t are time fixed effects; $D_{i,t}$ is a dummy variable indicating a pandemic event that affects country i in year t . See Table A1 for the full list of pandemic events

level of GDP level its growth rate, average country temperature, total health expenditures, government final expenditures, mortality rate, and other controls. The predictive model: “serves to reallocate probability mass from the regions of the distributions in the treatment/control subpopulations that are oversampled to those regions that are under-sampled, thus enabling identification in the framework of the Rubin Causal Model” (Jordà and Taylor 2016). Table 4 reports the Probit regression results. As shown in the table, the dates of pandemic events are uncorrelated with past levels and changes of inequality, but depend on some country characteristics such as temperature and GDP per capita.

Since we are interested in estimating the Average Treatment Effect (ATE) we use an augmented regression-adjusted estimation instead, denoted AIPW, which combines IPW with regression control and adjusts the estimator to achieve semi-parametric efficiency. Specifically, we estimate the following model

$$y_{i,t+k} = \alpha_i^k + \gamma_t^k + \Lambda^k D_{i,t} + \theta^k X_{i,t} + \varepsilon_{i,t+k} \tag{2}$$

with:

$$\hat{\Lambda}_{AIPW}^k = \frac{1}{n} \sum_t \left\{ \left[\frac{D_t(y_{t+k})}{\hat{p}_t} - \frac{(1-D_t)(y_{t+k})}{(1-\hat{p}_t)} \right] - \frac{(D_t - \hat{p}_t)}{\hat{p}_t(1-\hat{p}_t)} \left[(1-\hat{p}_t)m_1^h(X_t, \hat{\theta}_1^h) + \hat{p}_tm_0^h(X_t, \hat{\theta}_0^h) \right] \right\}$$

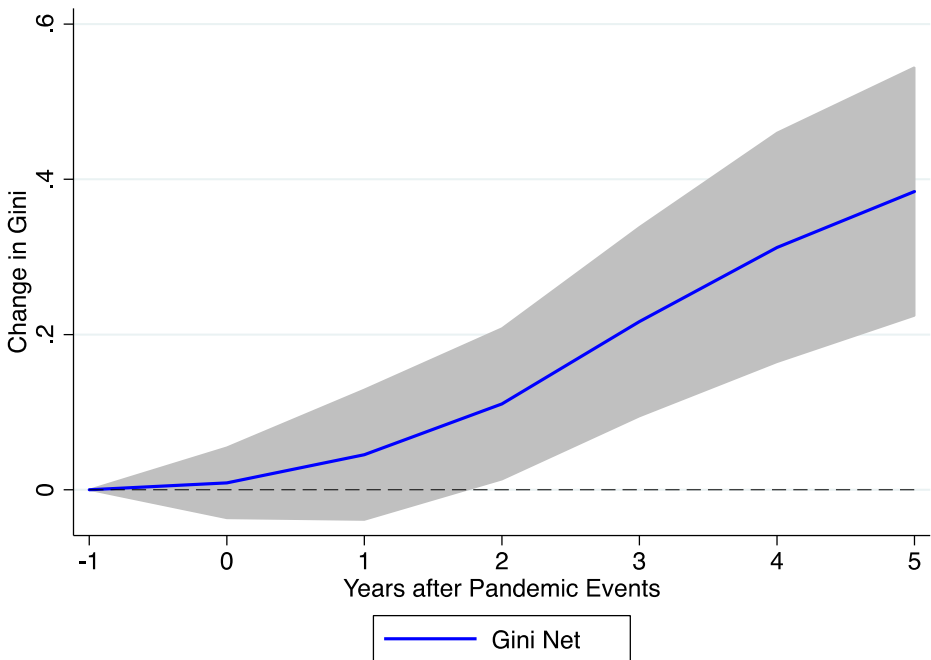


Fig. 6 Impact of pandemics on net Gini—Restricted sample (1990–2019). Notes: Impulse response functions are estimated using a sample of 177 countries over the period 1990–2019. The graph shows the response and 90% confidence bands. The x-axis shows years (k) after pandemic events; $t = 0$ is the year of the pandemic event. Estimates based on $y_{i,t+k} = \alpha_i^k + \gamma_t^k + \beta^k D_{i,t} + \theta^k X_{i,t} + \varepsilon_{i,t+k} \cdot y_{i,t}$ is the Gini coefficient for country i in year t ; α_i are country fixed effects; γ_t are time fixed effects; $D_{i,t}$ is a dummy variable indicating a pandemic event that affects country i in year t . $X_{i,t}$ is a vector that includes two lags of the dependent variable and two lags of the pandemic dummy. See Table A1 for the full list of pandemic events

where: \hat{p}_t is the propensity score obtained from estimating the Probit models as in Table 4; $m_j^h(X_t, \hat{\theta}_j^h)$ for $j = 1, 0$ is the conditional mean from the first-step regression of $(y_{i,t+k})$ on X_t ; $\hat{\theta}_j^h$ is the parameter accounting for the differential effect of the treatment conditioned to the value of the X_t ; $y_{i,t}$ is our distribution variable (e.g. the Gini coefficient) for country i in year t ; α_i are country fixed effects, included to take account of differences in countries' average income distribution; γ_t are time fixed effects, included to take account of global shocks such as shifts in oil prices or the global business cycle; $D_{i,t}$ is a dummy variable indicating a pandemic event that affects country i in year t . $X_{i,t}$ is a vector that includes two lags of the dependent variable and two lags of the pandemic dummy. The results are reported in Table 5. Regardless of the specification chosen, they point to statistically significant impact of pandemics on income inequality with effects being quantitatively close to those shown in Fig. 1.

3.4 Impact on Other Indicators of Distribution

To shed light on the channels through which pandemics affect inequality, we explore the impact of pandemic events on income shares and employment outcomes by educational groups. These results are for a smaller set of countries given data availability.

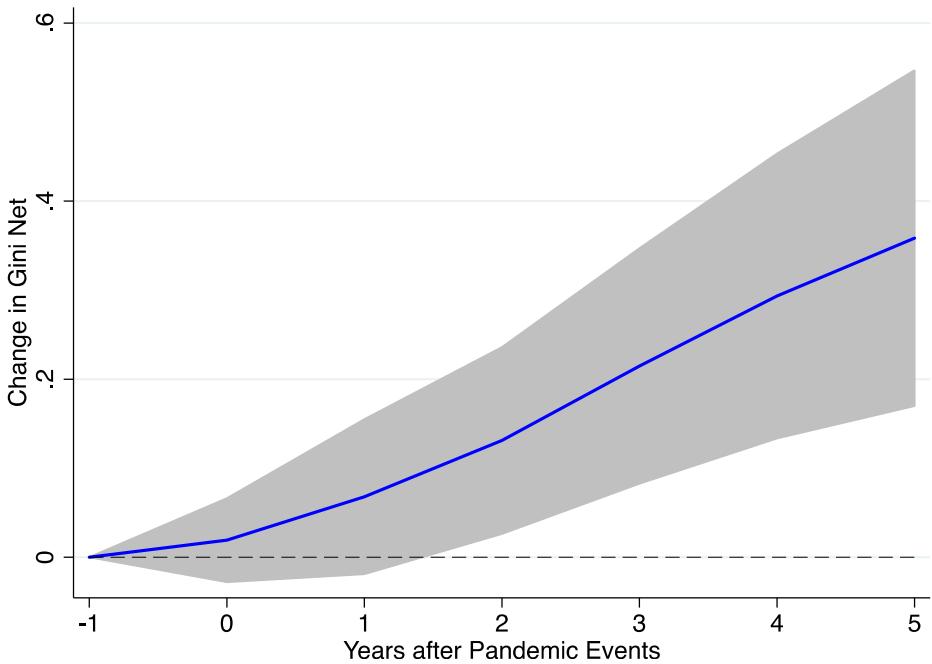


Fig. 7 Impact of pandemics on net Gini—Additional controls. Notes: Impulse response functions are estimated using a sample of 175 countries over the period 1960–2017. The graph shows the response and 90% confidence bands. The x-axis shows years (k) after pandemic events; $t = 0$ is the year of the pandemic event. Estimates based on $y_{i,t+k} = \alpha_i^k + \gamma_t^k + \beta^k D_{i,t} + \theta^k X_{i,t} + \varepsilon_{i,t+k}$. $y_{i,t}$ is the Gini coefficient for country i in year t ; α_i are country fixed effects; γ_t are time fixed effects; $D_{i,t}$ is a dummy variable indicating a pandemic event that affects country i in year t . $X_{i,t}$ is a vector that includes two lags of the dependent variable, two lags of the pandemic dummy, the level of GDP, the level of GDP square, population density, the share of population in urban area, the KOF index of trade globalization and the KOF index of financial globalization. See Table A1 for the full list of pandemic events

The results for the impact of pandemics on the income shares held by the top (bottom) 20% are shown in Fig. 11. It is evident that the impact is to raise the shares of the upper-income quintile and reduce those of the lower-income quintile. The impacts are statistically significant and quantitatively sizable. For instance, the share of income going to the top two deciles is 46% on average; five years after the pandemic, this share increases to nearly 48%. The share of income going to the bottom two deciles is 6%; five years after the pandemic, this share falls to 5.5%. We find similar effects when looking at the top (bottom) 10 and 40% (Fig. 12).

Figure 13 shows the disparate impact on the employment of people with different levels of educational attainment. Those with non-basic levels of education are scarcely affected, whereas the employment to population ratio of those with basic levels of education falls significantly, by more than 5% in the medium term—the effect is statistically significant at 5%.

3.5 Pandemics Vs. Financial Crises and Other Recessions

Are pandemics different from other recessions and crises? To answer this question, we augment our framework to include financial crises (taken from Laeven and Valencia 2020) and recession episodes—defined as years of negative real GDP growth (rather than in terms of output gaps, which are poorly measured in the case of developing countries). This exercise

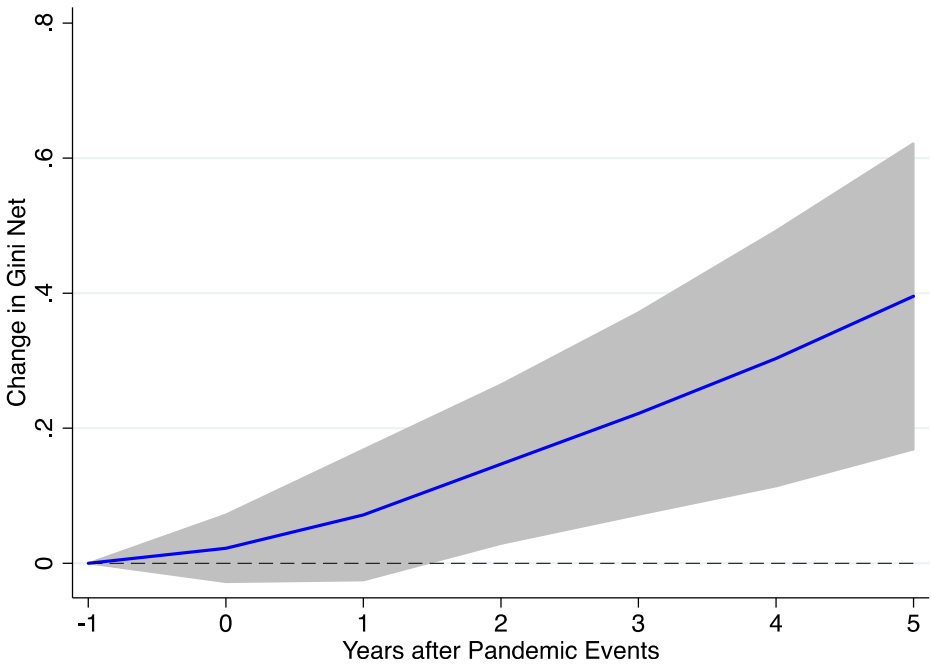


Fig. 8 Impact of pandemics on net Gini—adding country-specific time trends as control. Notes: Impulse response functions are estimated using a sample of 177 countries over the period 1960–2019. The graph shows the response and 90% confidence bands. The x-axis shows years (k) after pandemic events; $t = 0$ is the year of the pandemic event. Estimates based on $y_{i,t+k} = \alpha_i^k + \gamma_t^k + \beta^k D_{i,t} + \theta^k X_{i,t} + \varepsilon_{i,t+k}$. $y_{i,t}$ is the Gini coefficient for country i in year t ; α_i are country fixed effects; γ_t are time fixed effects; $D_{i,t}$ is a dummy variable indicating a pandemic event that affects country i in year t . $X_{i,t}$ is a vector that includes two lags of the dependent variable, two lags of the pandemic dummy and country-specific time trends. See Table A1 for the full list of pandemic events

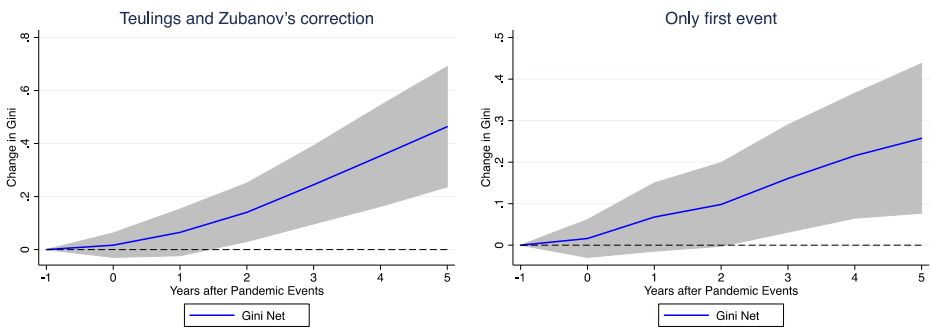


Fig. 9 Impact of pandemics on net Gini—Teulings and Zubanov’s correction and no overlapping events. Notes: Impulse response functions are estimated using a sample of 177 countries over the period 1960–2019. The graph shows the response and 90% confidence bands. The x-axis shows years (k) after pandemic events; $t = 0$ is the year of the pandemic event. Estimates based on $y_{i,t+k} = \alpha_i^k + \gamma_t^k + \beta^k D_{i,t} + \theta^k X_{i,t} + \varepsilon_{i,t+k}$. $y_{i,t}$ is the Gini coefficient for country i in year t ; α_i are country fixed effects; γ_t are time fixed effects; $D_{i,t}$ is a dummy variable indicating a pandemic event that affects country i in year t . $X_{i,t}$ is a vector that includes two lags of the dependent variable, two lags of the pandemic dummy. See Table A1 for the full list of pandemic events. For the left panel, the equation is augmented with dummy variables for pandemics occurring within the forecast horizon. For the right panel, $D_{i,t}$ is takes the value of 1 only for the first pandemic that affect country i , and 0 otherwise

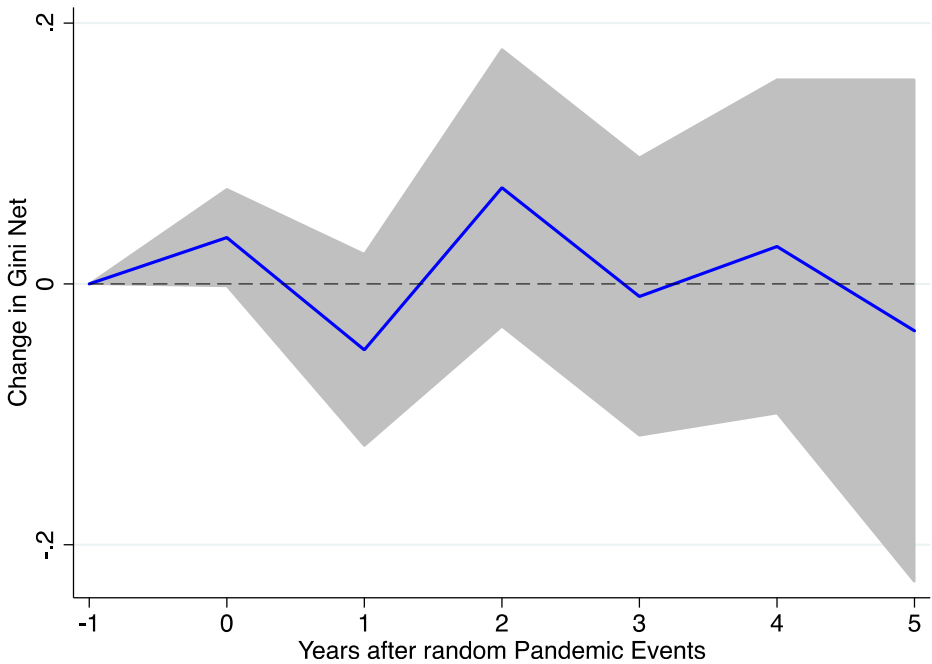


Fig. 10 Impact of pandemics on net Gini—Placebo test. Notes: Impulse response functions are estimated using a sample of 177 countries over the period 1960–2019. The graph shows the response and 90% confidence bands. The x-axis shows years (k) after pandemic events; $t = 0$ is the year of the pandemic event. Estimates based on $y_{i,t+k} = \alpha_i^k + \gamma_t^k + \beta^k D_{i,t} + \theta^k X_{i,t} + \varepsilon_{i,t+k}$. $y_{i,t}$ is the Gini coefficient for country i in year t ; α_i are country fixed effects; γ_t are time fixed effects; $D_{i,t}$ is a dummy variable indicating a pandemic event that affects country i in year t . $X_{i,t}$ is a vector that includes two lags of the dependent variable and two lags of the pandemic dummy. The placebo test is conducted estimating the impulse responses attributing the values of our measure of the shock, $D_{i,t}$, randomly, across the whole sample

also allows us to address the concern that some pandemic events in our sample may have occurred also during a period of crisis or recession. We estimate the following equation:

$$y_{i,t+k} = \alpha_i^k + \gamma_t^k + \beta^k D_{i,t} + \vartheta^k C_{i,t} + \theta^k M_{i,t} + \varepsilon_{i,t+k} \tag{3}$$

where C denotes the year of occurrence of a financial crisis or a year in which growth was negative, and M includes our earlier set of control variable X augmented by two lags of the financial crisis or recession dummy.⁶

The results in Fig. 14 suggest that the distributional effects of pandemics are larger than those associated with financial crises or recessions: financial crises do not have a significant effect on inequality, and the Gini increases by about 0.05 following a typical recession—compared to more than 0.4 for pandemics.⁷ We find similar results when looking at income shares. While recessions seem to result in higher top income shares, top shares tend to decline in the medium term following a financial crisis (see Fig. 16) consistent with the fact the

⁶ Qualitatively similar results are obtained including financial crises and recessions at the same time.

⁷ This result is consistent with Camacho and Palmieri (2019) who did not find significant positive impacts of economic downturns and financial crises on income distribution. Consistent with the insignificant effect of financial crises on inequality, we also find that the effect of the H1N1 pandemic during the Global Financial Crisis is lower than that in other pandemic episodes (see Fig. 15).

Table 4 Pandemic dummy regression, pooled probit estimator (average marginal effects)

Probit model of treatment at time $t+1$ (pandemic dummy)				
	(1)	(2)	(3)	(4)
Total Health Expenditures	-0.022 (0.054)	-0.031 (0.055)		
General government final consumption expenditure (% of GDP)	-0.042 (0.029)	-0.046 (0.031)	-0.030 (0.023)	-0.032 (0.024)
Age dependency ratio (% of working-age population)	-0.046*** (0.016)	-0.046*** (0.017)	-0.048*** (0.011)	-0.053*** (0.013)
GDP per capita (constant 2010 US\$) (log)	0.965** (0.390)	0.851** (0.408)	1.686*** (0.302)	1.648*** (0.316)
Growth rate of GDP	-0.016 (0.013)	-0.016 (0.013)	-0.014 (0.011)	-0.016 (0.012)
Mortality rate, adult, (per 1000 adults)	0.002* (0.001)	0.001 (0.001)	0.001 (0.001)	0.000 (0.001)
Temperature (Year average)	0.253*** (0.080)	0.298*** (0.083)	0.328*** (0.074)	0.375*** (0.077)
Gini Disposable - level		-0.0691 (0.0459)		-0.002 (0.034)
Gini Disposable - change		-0.0215 (0.190)		-0.101 (0.159)
Observations	1912	1850	3485	3413

Note: Country fixed effects included but not reported. Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 5 Average treatment effect of pandemics, AIPW estimates

Panel A						
	k=0	k=1	k=2	k=3	k=4	k=5
ATE, restricted $\theta_1^h = \theta_0^h$	0.06*** (0.02)	0.13*** (0.03)	0.25*** (0.04)	0.26*** (0.05)	0.17*** (0.06)	0.25*** (0.06)
ATE, unrestricted $\theta_1^h \neq \theta_0^h$	0.02 (0.02)	0.16*** (0.04)	0.27*** (0.05)	0.36*** (0.06)	0.38*** (0.08)	0.62*** (0.08)
Observations	1826	1826	1745	1630	1511	1393
Panel B						
	k=0	k=1	k=2	k=3	k=4	k=5
ATE, restricted $\theta_1^h = \theta_0^h$	0.08*** (0.01)	0.15*** (0.03)	0.28*** (0.04)	0.35*** (0.05)	0.38*** (0.06)	0.47*** (0.06)
ATE, unrestricted $\theta_1^h \neq \theta_0^h$	-0.01 (0.02)	0.09** (0.04)	0.16*** (0.06)	0.15** (0.07)	0.13 (0.09)	0.27*** (0.10)
Observations	3298	3298	3213	3094	2971	2849

Notes: Empirical sandwich standard errors (clustered by country) in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. AIPW estimates based on the baseline model as in eq. (5). Results for Panels A and B are based on the propensity scores obtained using the model of column 2 and 4 of Table 4, respectively. When imposing $\theta_1^h = \theta_0^h$ (i) the effect of the controls X_i on the outcomes is assumed to be stable across the treated and control subpopulations (i.e. countries experiencing a pandemic event and countries not experiencing a pandemic event); (ii) the expected value of X_i in each subpopulation is assumed to be the same. When imposing $\theta_1^h \neq \theta_0^h$ these assumptions are relaxed. For further details see the methodological section and Jordà and Taylor (2016)

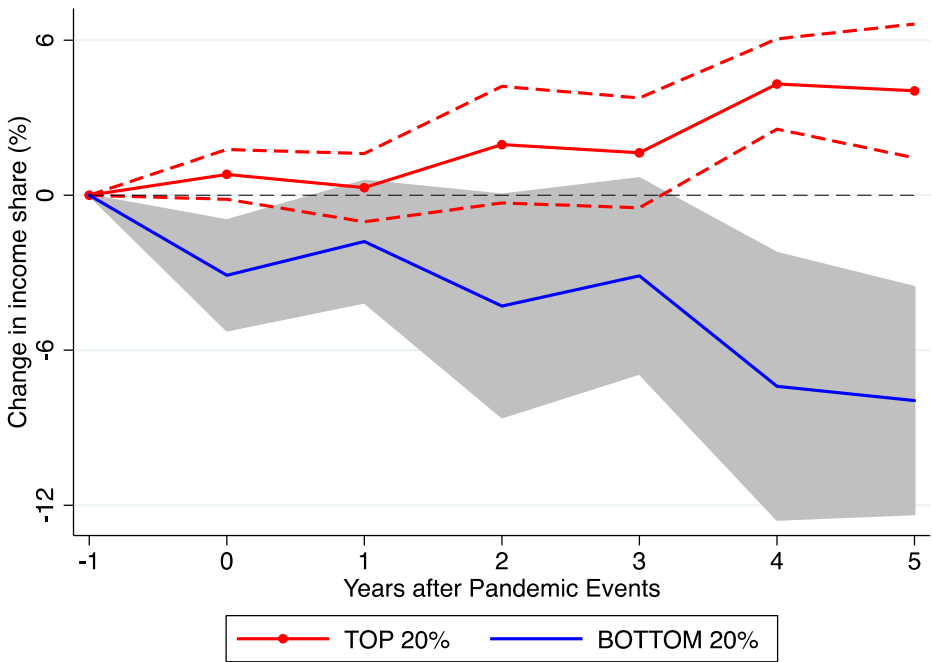


Fig. 11 Impact of pandemics on shares of income. Notes: Impulse response functions are estimated using a sample of 64 countries over the period 1981–2017. The graph shows the response and 90% confidence bands. The x-axis shows years (k) after pandemic events; $t = 0$ is the year of the pandemic event. Estimates are based on $y_{i,t+k} - y_{i,t-1} = \alpha_i^k + \gamma_t^k + \beta^k D_{i,t} + \theta^k X_{i,t} + \varepsilon_{i,t+k}$. $y_{i,t}$ is the log of the income share held by the top (bottom) 20% for country i in year t ; α_i are country fixed effects; γ_t are time fixed effects; $D_{i,t}$ is a dummy variable indicating a pandemic event that affects country i in year t . $X_{i,t}$ is a vector that includes two lags of the dependent variable and the pandemic dummy. See Table A1 for the full list of pandemic events

financial income tends to be highly concentrated in the upper part of the income distribution.⁸ Finally, the medium-term effects on the employment to population ratio of those with basic education falls significantly in all types of crises, suggesting that the difference in distributional effects between pandemics and other recessions is not due to this channel (see Fig. 17).

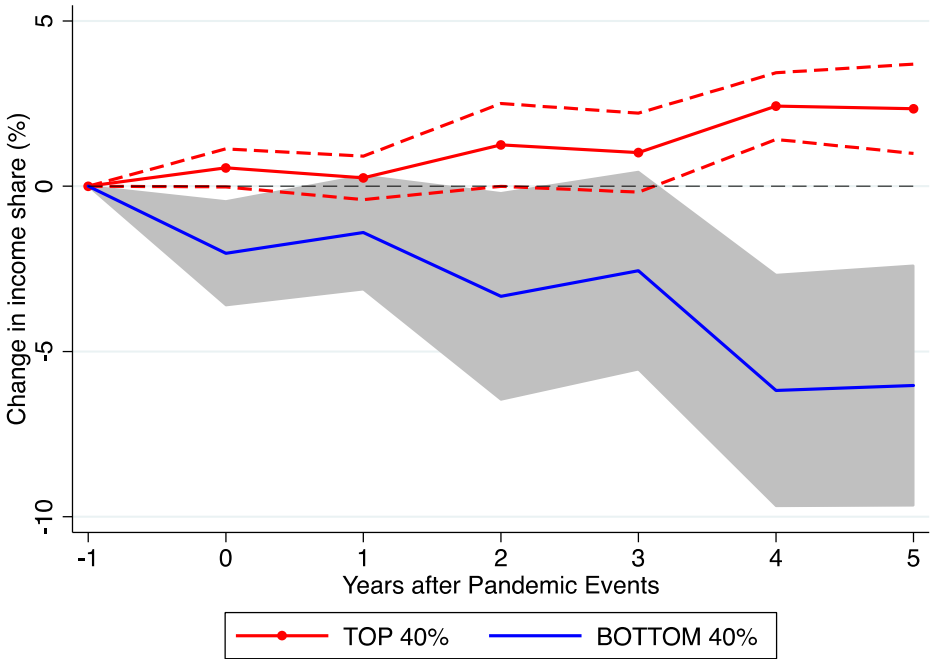
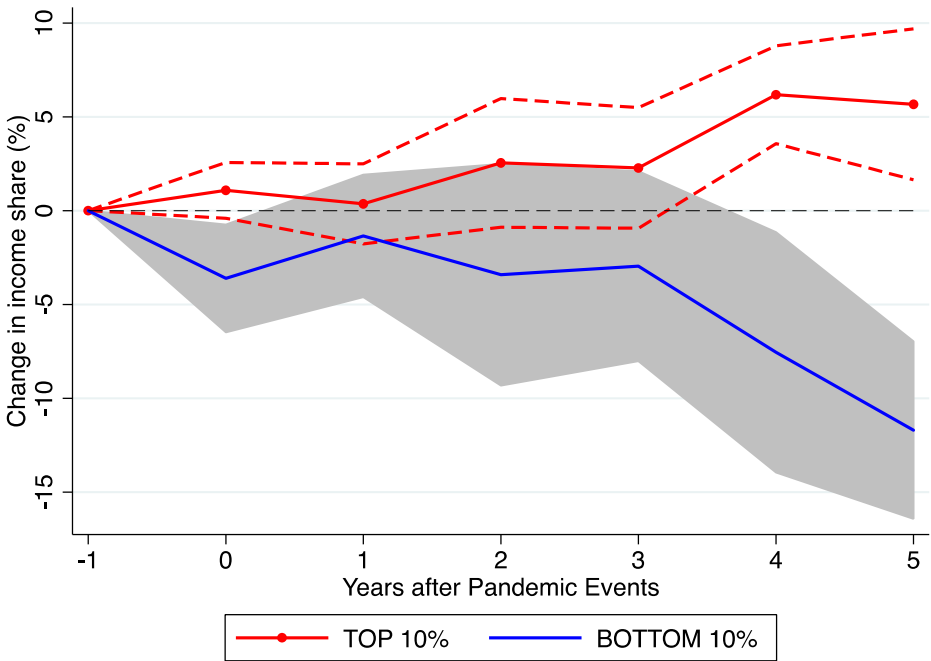
3.6 Heterogeneity across Episodes Depend on the Severity of Pandemics

The average response of inequality to pandemic events may mask significant heterogeneity across episodes, based on the severity of the pandemic event, both in terms of confirmed cases and its economic effects. To probe further, we use two approaches. In the first, we replace pandemic dummy with a continuous variable using the information of the number of cases (Emmerling et al. 2021). Specifically, we estimate the following equation:

$$y_{i,t+k} = \alpha_i^k + \gamma_t^k + \beta_c^k \text{cases} + \theta^k X_{i,t} + \epsilon_{i,t}^k \tag{4}$$

where, the variable proxying the severity of the pandemic is alternatively, $\text{cases} = \log_{10}(1 + x)$ or $\text{cases} = \ln(x + (x^2 + 1)^{1/2})$ with $x = \left(\frac{1000 \cdot \text{confirmed_cases}_{i,t}}{\text{population}_{i,t}}\right)$. The latter (i.e., the inverse

⁸ Income comprises labor, business financial income and transfers.



hyperbolic sine transformation – IHS) is particularly useful to transform skewed variables that include zero or negative values.

Fig. 12 Impact of pandemics on shares of income, by deciles – robustness checks. Notes: Impulse response functions are estimated using a sample of 64 countries over the period 1981–2017. The graph shows the response and 90% confidence bands. The x-axis shows years (k) after pandemic events; $t = 0$ is the year of the pandemic event. Estimates are based on $y_{i,t+k} - y_{i,t-1} = \alpha_i^k + \gamma_t^k + \beta^k D_{i,t} + \theta^k X_{i,t} + \varepsilon_{i,t+k}$. $y_{i,t}$ is, in turn, the log of the income share held by the top (bottom) 10% (40%) for country i in year t ; α_i are country fixed effects; γ_t are time fixed effects; $D_{i,t}$ is a dummy variable indicating a pandemic event that affects country i in year t . $X_{i,t}$ is a vector that includes two lags of the dependent variable and two lags of the pandemic dummy. See Table A1 for the full list of pandemic events

While the use of this continuous variable has the advantage of differentiating episodes based on their severity, it has two important drawbacks. First, it may be more prone to reverse causality as higher initial levels of inequality may increase the number of infections due to the higher economic and health vulnerability of marginalized people. Second, measurement errors related to total cases detected is likely to be non-negligible. To address these concerns, we resort to an instrumental variable approach. Following Nunn and Qian (2014), our Instrumental Variable (IV) approach consists of interacting a time-varying global term and a constant country-specific term. The global term is a dummy variable that takes the value of 1 for all countries in the years of pandemic outbreaks. The country-term we consider captures the factors affecting the severity of the pandemic. For this purpose, we consider the average temperature. As shown in several recent studies (i.e. Ma et al. 2020b; Ujiie et al. 2020),

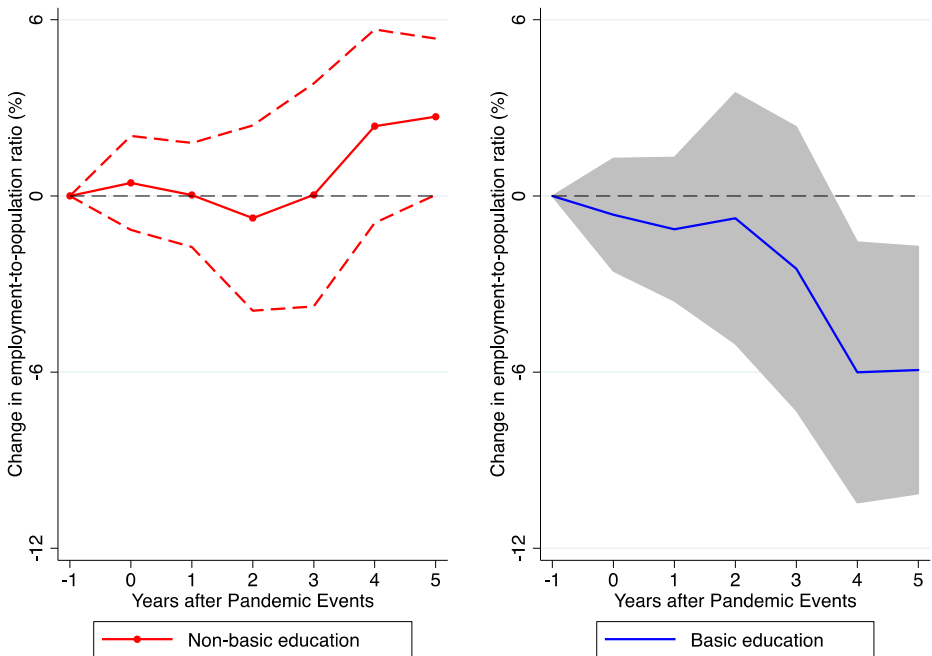


Fig. 13 Impact of pandemics on employment-to-population ratio, by education level. Notes: Impulse response functions are estimated using a sample of 76 countries over the period 1990–2017. The graph shows the response and 90% confidence bands. The x-axis shows years (k) after pandemic events; $t = 0$ is the year of the pandemic event. Estimates are based on $y_{i,t+k} - y_{i,t-1} = \alpha_i^k + \gamma_t^k + \beta^k D_{i,t} + \theta^k X_{i,t} + \varepsilon_{i,t+k}$. $y_{i,t}$ is, in turn, the log of employment-to-population ratio by education level for country i in year t ; α_i are country fixed effects; γ_t are time fixed effects; $D_{i,t}$ is a dummy variable indicating a pandemic event that affects country i in year t . $X_{i,t}$ is a vector that includes two lags of the dependent variable and the pandemic dummy. See Table A1 for the full list of pandemic events

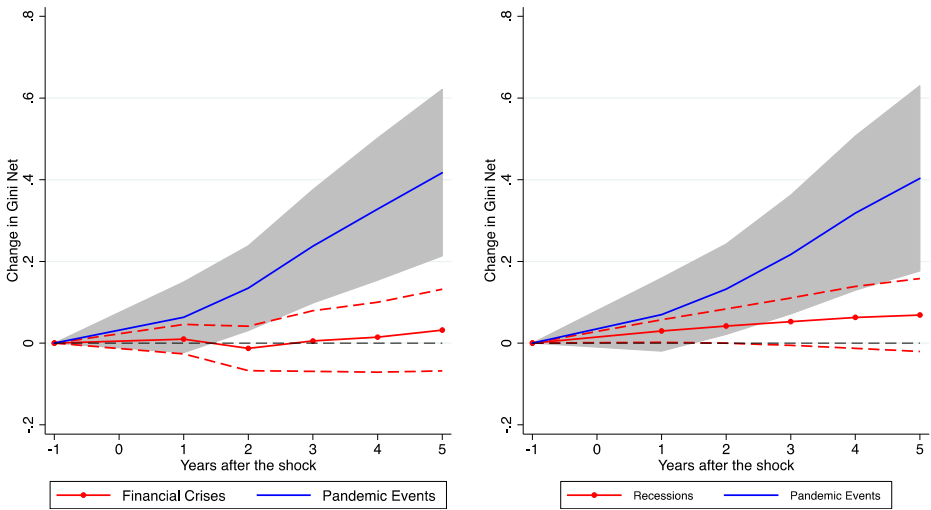


Fig. 14 Impact of Pandemics vs. financial crises and other recessions. Notes: Impulse response functions are estimated using a sample of 177 countries over the period 1960–2019 (1970–2017 in the case of financial crises). The graph shows the response and 90% confidence bands. The x-axis shows years (k) after pandemic events; $t = 0$ is the year of the pandemic event. Estimates based on $y_{i,t+k} = \alpha_i^k + \gamma_t^k + \beta^k D_{i,t} + \psi^k C_{i,t} + \theta^k M_{i,t} + \varepsilon_{i,t+k}$. $y_{i,t}$ is the Gini coefficient for country i in year t ; α_i are country fixed effects; γ_t are time fixed effects; $D_{i,t}$ is a dummy variable indicating a pandemic event that affects country i in year t ; $C_{i,t}$ is a dummy variable denoting, alternatively, the year of occurrence of a financial crisis or a year of negative growth, $M_{i,t}$ is a vector that includes two lags of the dependent variable and the pandemic dummy plus two lags of the financial crisis (recession). See Table A1 for the full list of pandemic events. The F-tests for the difference between the estimations in the case of Pandemics and Financial Crises (Recessions) are shown in Table 6

temperature is an important driver of the evolution of pandemics and it can reasonably be assumed to be exogenous. Our IV estimation reads as follows:

$$y_{i,t+k} = \beta_c^k (\widehat{cases}_{i,t}) + \theta^k X_{i,t} + \alpha_i^k + \gamma_t^k + \epsilon_{i,t}^k \tag{5}$$

with $(\widehat{cases}_{i,t}) = \vartheta^k S_{i,t} + \varphi^k X_{i,t} + \alpha_i^k + \gamma_t^k + v_{i,t}^k$, where S is the instrument and $cases$ is, alternatively, one of the transformations discussed above. The analysis also controls for country and time fixed effects and can therefore be seen as a *differences-in-differences* approach (Nunn and Qian 2014).⁹

In the second approach, we estimate the following equation:

$$y_{i,t+k} - y_{i,t-1} = \alpha_i^k + \gamma_t^k + F(z_{it}) [\beta_L^k D_{i,t} + \theta_L^k X_{i,t}] + (1 - F(z_{it})) [\beta_H^k D_{i,t} + \theta_H^k X_{i,t}] + \varepsilon_{i,t+k}$$

with

$$F(z_{it}) = \frac{\exp^{-\gamma z_{it}}}{(1 + \exp^{-\gamma z_{it}})}, \quad \gamma = 3.5 \tag{6}$$

⁹ The first-stage estimates suggest that the instrument is “strong” and statistically significant. The Kleibergen–Paap rk Wald F-statistic—which is equivalent to the F-effective statistic for non-homoscedastic error in case of one endogenous variable and one instrument (Andrews et al., 2019)—is higher than the associated Stock–Yogo critical value (Table 7).

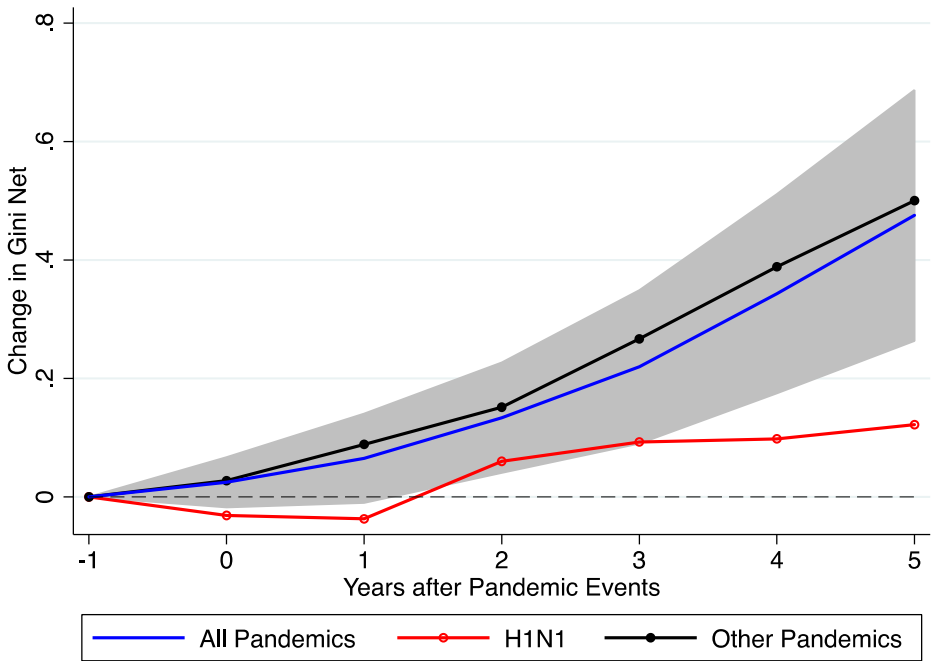


Fig. 15 Impact of pandemics on net Gini, by pandemics – robustness checks. Notes: Impulse response functions are estimated using a sample of 177 countries over the period 1960–2019. The graph shows the response and 90% confidence bands. The x-axis shows years (k) after pandemic events; $t = 0$ is the year of the pandemic event. Estimates based on $y_{i,t+k} = \alpha_i^k + \gamma_t^k + \beta^k D_{i,t} + \theta^k X_{i,t} + \varepsilon_{i,t+k}$. $y_{i,t}$ is the Gini coefficient for country i in year t ; α_i are country fixed effects; γ_t are time fixed effects; $D_{i,t}$ is a dummy variable indicating a pandemic event that affects country i in year t . $X_{i,t}$ is a vector that includes two lags of the dependent variable and two lags of the pandemic dummy. See Table A1 for the full list of pandemic events

where z is an indicator of the severity of the pandemic (which is either the ratio of confirmed cases to population, or GDP growth), normalized to have zero mean and a unit variance. The weights assigned to each regime vary between 0 and 1 according to the weighting function $F(\cdot)$, so that $F(z_{it})$ can be interpreted as the probability of being in a given state of the pandemic. The coefficients β_L^k and β_H^k capture the distributional impact of a pandemic event at each horizon k in cases of mild pandemics in terms of cases-to-population ratio (or alternatively, higher output growth) ($F(z_{it}) \approx 1$ when z goes to minus infinity) and extremely severe pandemic events in terms of cases-to-population ratio (or alternatively, lower output growth) ($1 - F(z_{it}) \approx 1$ when z goes to plus infinity), respectively— $F(z_{it})=0.5$ is the cutoff between severe and weak pandemic event. We choose $\gamma = 3.5$, following Tenreyro and Thwaites (2016).

The results in Fig. 18 show that the distributional effect of pandemic events varies with their severity. Using the continuous variable instead of a (0–1) dummy, the results point to larger effects of pandemics on inequality as case-to-population ratios increase: a 1 % increase in the measure of severity implies a rise in net Gini of about 0.4 (0.15 when using the inverse hyperbolic sine transformation) (Fig. 18 - Panels A and B). In other words, the effect of an *average pandemic*—based on the average infection rate in our dataset (0.80 cases per 1000 inhabitants)—is associated with a medium-term increase in the net Gini of about 0.1. This implies that for a severe pandemic in our sample (i.e. at the 99th percentile, with 15 cases per

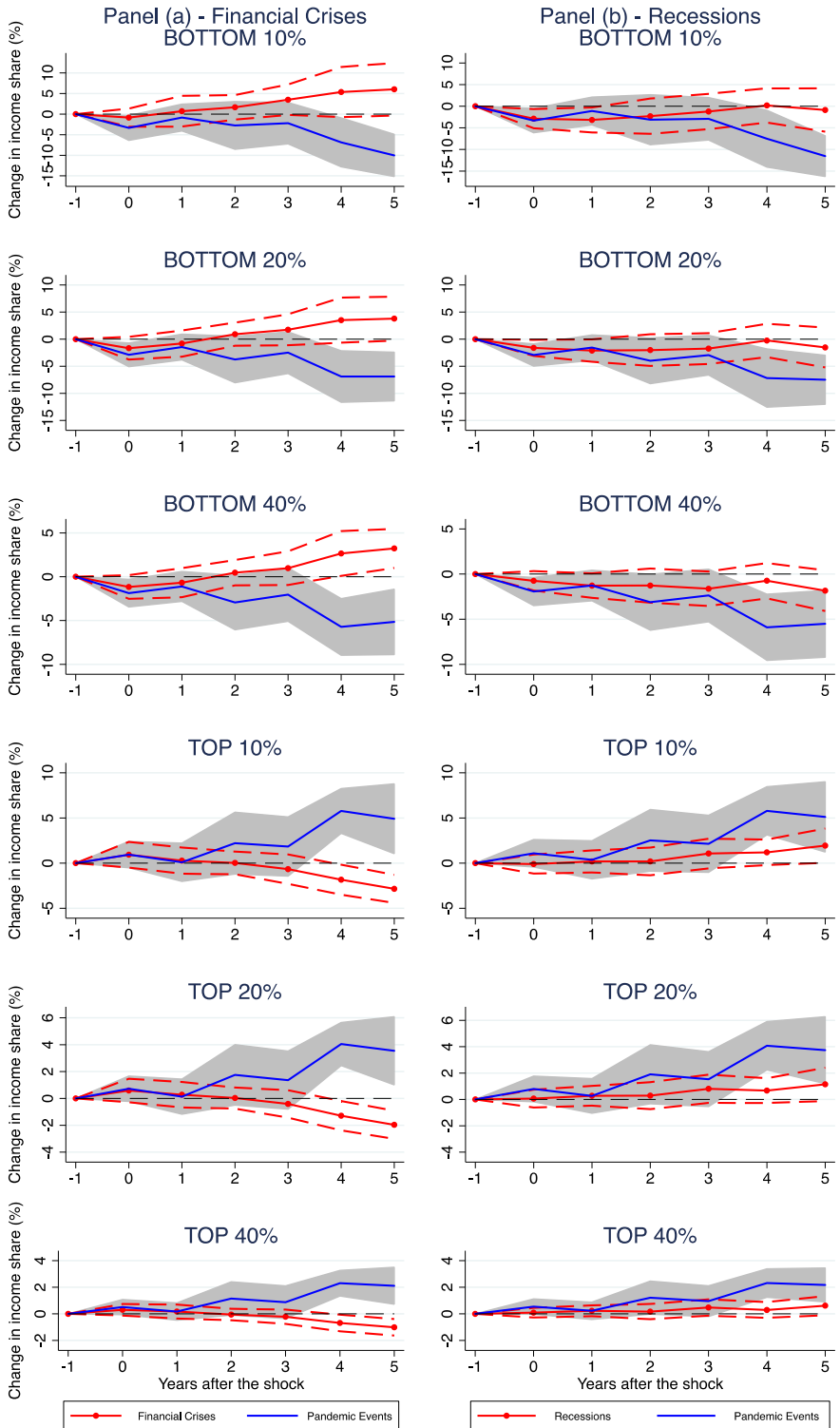


Fig. 16 Impact of pandemics, financial crises and recessions on shares of income, by deciles. Notes: Impulse response functions are estimated using a sample of 64 countries over the period 1981–2017. The graph shows the response and 90% confidence bands. The x-axis shows years (k) after pandemic events (financial crises or recessions); $t = 0$ is the year of the pandemic event (financial crisis or recession). Estimates based on $y_{i,t+k} - y_{i,t-1} = \alpha_i^k + \gamma_t^k + \beta^k D_{i,t} + \vartheta^k C_{i,t} + \theta^k M_{i,t} + \varepsilon_{i,t+k}$. $y_{i,t}$ is, in turn, the log of the income share held by the top (bottom) 10%, 20%, or 40% for country i in year t ; α_i are country fixed effects; γ_t are time fixed effects; $D_{i,t}$ is a dummy variable indicating a pandemic event that affects country i in year t ; $C_{i,t}$ is a dummy variable denoting, alternatively, the year of occurrence of a financial crisis or a year of negative growth, $M_{i,t}$ is a vector that includes two lags of the dependent variable, two lags of the pandemic dummy plus two lags of the financial crisis (recession). See Table A1 for the full list of pandemic events

1000 people), the Gini index would increase, on average, by 0.6 percentage point. Taking this effect at face value and translating it to the current pandemic, it implies that COVID-19 would lead to a medium-term increase in the Gini of at least 0.7/0.9 percentage point (indeed, as of February 15th, 2022, the pandemic counted on average 30 cases per 1000 inhabitants. The IV results confirm the adverse distributional effects of pandemics, with the magnitude of the coefficient significantly larger than the corresponding OLS.

The results obtained from estimating Eq. (6) show that for episodes associated with a larger number of cases relative to population (such as Croatia, H1N1, 2009), the effect is statistically significant and larger than the average effect shown in Fig. 1 (the medium-term effect on Gini increases from 0.4 to about 0.8), while it is not statistically different from zero for episodes

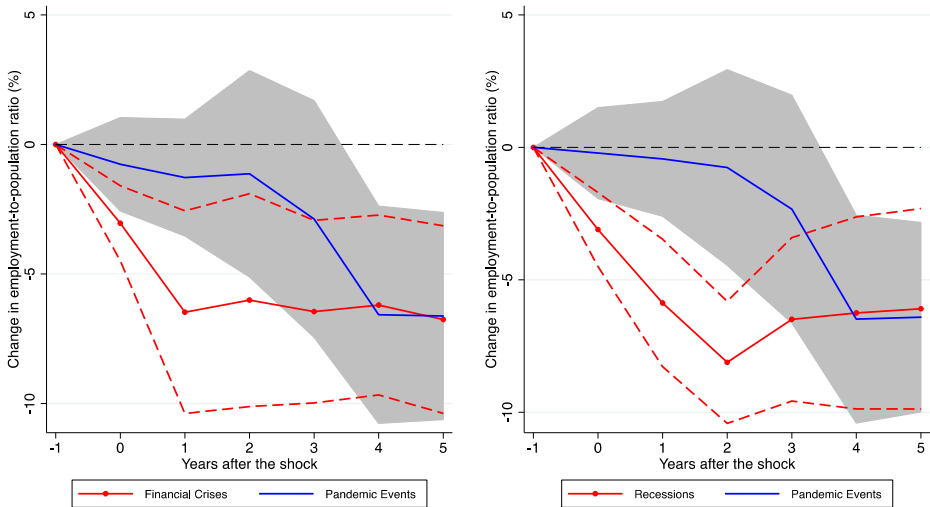


Fig. 17 Impact of pandemics, financial crises and recessions on E/P ratio – basic education. Notes: Impulse response functions are estimated using a sample of 76 countries over the period 1990–2017. The graph shows the response and 90% confidence bands. The x-axis shows years (k) after pandemic events (financial crises or recessions); $t = 0$ is the year of the pandemic event (financial crisis or recession). Estimates based on $y_{i,t+k} - y_{i,t-1} = \alpha_i^k + \gamma_t^k + \beta^k D_{i,t} + \vartheta^k C_{i,t} + \theta^k M_{i,t} + \varepsilon_{i,t+k}$. $y_{i,t}$ is, in turn, the log of employment-to-population ratio by education level for country i in year t ; α_i are country fixed effects; γ_t are time fixed effects; $D_{i,t}$ is a dummy variable indicating a pandemic event that affects country i in year t ; $C_{i,t}$ is a dummy variable denoting, alternatively, the year of occurrence of a financial crisis or a year of negative growth, $M_{i,t}$ is a vector that includes two lags of the dependent variable, two lags of the pandemic dummy plus two lags of the financial crisis (recession). See Table A1 for the full list of pandemic events

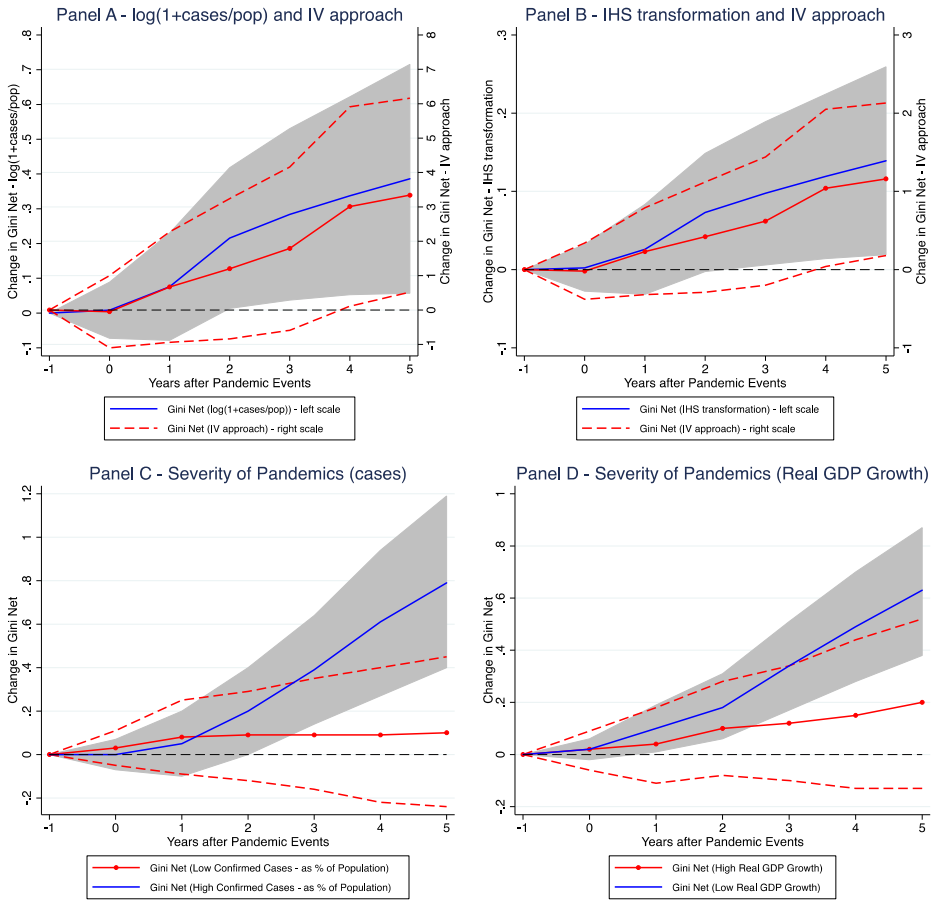


Fig. 18 Impact of pandemics on net Gini—The role of the severity of the pandemic. Notes: Impulse response functions are estimated using a sample of 177 countries over the period 1960–2019. The graph shows the response and 90% confidence bands. The x-axis shows years (k) after pandemic events; $t = 0$ is the year of the pandemic event. For Panel A, estimates based on $y_{i,t+k} = \alpha_i^k + \gamma_t^k + \beta^k \text{cases}_{i,t} + \theta^k X_{i,t} + \varepsilon_{i,t+k}$, where $\text{cases}_{i,t}$ is alternatively $\text{cases} = \log_{10}(1 + x)$ or $\text{cases} = \ln(x + (x^2 + 1)^{1/2})$ with $x = \left(\frac{1000 \cdot \text{confirmed_cases}_{i,t}}{\text{population}_{i,t}}\right)$. The instrumental variable (IV) approach consists of interacting an exogenous time-varying global term—the date of the initial pandemic outbreak in the world—and a country-specific terms exogenous to economic outcomes—such as average temperature (see Table 7 for first-stage estimate). For Panels B and C estimates based on $y_{i,t+k} = \alpha_i^k + \gamma_t^k + F(z_{it}) [\beta_L^k D_{i,t} + \theta_L^k X_{i,t}] + (1 - F(z_{it})) [\beta_H^k D_{i,t} + \theta_H^k X_{i,t}] + \varepsilon_{i,t+k}$. $y_{i,t}$ is the Gini coefficient for country i in year t ; α_i are country fixed effects; γ_t are time fixed effects; $D_{i,t}$ is a dummy variable indicating a pandemic event that affects country i in year t . $X_{i,t}$ is a vector that includes two lags of the dependent variable and the pandemic dummy. $F(z_{it})$ is an indicator function of the severity of the pandemic. The coefficients β_L^k and β_H^k capture the distributional impact of a pandemic event at each horizon k in cases of extremely severe pandemics ($F(z_{it}) \approx 1$ when z goes to minus infinity) and weak pandemics ($1 - F(z_{it}) \approx 1$ when z goes to plus infinity), respectively. The F-tests for the difference between the estimations in the case of low and high regime of the interaction variable with the pandemic dummy are shown in Table 6. See Table A1 for the full list of pandemic events

Table 6 – F-tests difference

	F-test difference					
	k=0	k=1	k=2	k=3	k=4	k=5
Pandemics vs Financial Crises ^a	0.001	0.895	3.976**	5.807**	7.106***	8.473***
Pandemics vs Recessions ^a	0.075	0.472	1.521	2.794*	3.832*	4.545**
Interaction with Cases-Population ratio ^b	0.159	0.0301	0.295	1.425	2.642	3.575*
Interaction with GDP Growth ^b	0.006	0.615	0.539	2.091	3.142*	3.727*

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. ^a The F-test of the difference between the estimations in the case of Pandemics and Financial Crises (Recessions)—(See Fig. 14). ^b The F-test of the difference between the estimations in the case of low and high regime of the interaction variable with the pandemic dummy—(See Fig. 18)

associated with small outbreaks (such as Philippines, SARS, 2003) (Fig. 18 – Panel B).¹⁰ Similarly, the results in Panel C show that the medium-term effect is larger (about 0.7) in episodes associated with low growth (such as Korea, MERS, 2012), while it is not statistically different from zero for episodes associated with high growth (such as China, H1N1, 2009).¹¹

Finally, to further explore how the distributional consequences of pandemics vary with their severity and how they compare with the preliminary effect from COVID-19, we repeated our analysis for a small subset of 17 countries for which Gini data are available for 2020.¹² We add to Eq. (1) a dummy variable which takes the value 1 when there is an outbreak of COVID-19 in the country and 0 otherwise, and re-estimate the model for $k = 0$ (i.e., to obtain the contemporaneous effect):

$$y_{i,t+k} = \alpha_i^k + \gamma_i^k + \beta^k D[COVID-19]_{i,t} + \beta^k D[Past Pandemics]_{i,t} + \theta^k X_{i,t} + \varepsilon_{i,t+k} \quad (7)$$

The results suggest a contemporaneous increase in the net Gini of 0.24 point following the COVID-19 outbreak– which is statistically significant at the 1% level (Table 8). The effect is roughly 6 times larger than that of past pandemics in this restricted sample and suggests that the distributional consequences of COVID-19 may be larger than those following past pandemics in our sample.

4 Conclusion

The COVID-19 crisis is already showing how the more vulnerable socio-economic groups suffer from a greater risk of financial exposure, greater health risks, and worse housing conditions during the lockdown period. These factors may exacerbate inequalities.

Our paper explores this possibility by providing evidence on the impact of pandemics and major epidemics from the past two decades on income distribution. Our results justify the concern that, in the absence of long-lasting supportive policies to protect the vulnerable, the

¹⁰ We obtain very similar results even when considering the mortality rates as a proxy of the severity of the pandemics (see Fig. A1 in the appendix).

¹¹ The F-test of the difference between the estimations in the case of low and high regime of the interaction variable with the pandemic dummy are shown in Table 6.

¹² The countries are: Bolivia, China, Colombia, Costa Rica, Denmark, Estonia, Georgia, Ireland, Kazakhstan, Kyrgyzstan, Paraguay, Poland, Singapore, Taiwan, United Kingdom, United States, Uruguay. Table A3 in the appendix provides a list of pandemic events for this restricted sample of countries.

Table 7 Pandemics and inequality - IV first stage

log₁₀(1+x) transformation						
	k=0	k=1	k=2	k=3	k=4	k=5
<i>Instrument</i>	-0.002*** (-2.770)	-0.002*** (-3.570)	-0.003*** (-3.530)	-0.003*** (-3.840)	-0.004*** (-3.850)	-0.004*** (-4.340)
Observations	4888	4725	4561	4398	4235	4075
Centered R-squared	0.012	0.019	0.020	0.026	0.028	0.040
Kleibergen-Paap_rk_Wald_F_statistic	7.676	12.77	12.45	14.72	14.82	18.86
Inverse Hyperbolic Sine (IHS) transformation						
	k=0	k=1	k=2	k=3	k=4	k=5
<i>Instrument</i>	-0.006*** (-2.780)	-0.007*** (-3.570)	-0.008*** (-3.530)	-0.009*** (-3.830)	-0.010*** (-3.840)	-0.013*** (-4.360)
Observations	4888	4725	4561	4398	4235	4075
Centered R-squared	0.013	0.019	0.020	0.026	0.029	0.038
Kleibergen-Paap_rk_Wald_F_statistic	7.720	12.80	12.46	14.69	14.78	19.00

Note: k = 0 is the year of the pandemic. k = 1,2,3,4,5 are the years after the pandemic event. IV first stage estimates based on Eq. (1) in the main text. Robust t-statistics in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. Control variables included but not reported. The Kleibergen-Paap rk Wald F-statistic tests for weak identification

Table 8 COVID-19 vs past pandemics – preliminary evidence

	k=0
COVID-19	0.244*** (0.063)
Past Pandemics	0.038 (0.111)
Observations	691
R ²	0.998

Note: Estimates are obtained using a sample of 17 countries over the period 1960–2020, and based on

$y_{i,t+k} = \alpha_i^k + \gamma_t^k + \beta^k D[\text{COVID-19}]_{i,t} + \theta^k X_{i,t} + \varepsilon_{i,t+k}$. $y_{i,t}$ is the Gini coefficient for country i in year t ; α_i are country fixed effects; γ_t are time fixed effects; $D[\text{Past Pandemics}]_{i,t}$ is a dummy variable indicating a pandemic event that affects country i in year t ; $D[\text{COVID-19}]_{i,t}$ is a dummy variable indicating COVID-19 that affects country i in year t . $X_{i,t}$ is a vector that includes two lags of the dependent variable. Standard errors in parentheses are clustered at the country level. Country and time fixed effects included but not reported. Table A3 in the appendix provides a list of pandemic events for this restricted sample of countries

pandemic could end up exerting a significant impact on inequality: past events of this kind, even though much smaller in scale, have led to increases in the Gini coefficient, raised the income shares accruing to the higher deciles of the income distribution, and lowered the employment-to-population ratio for those with basic education compared to those with higher education.

Will the past be prologue? Our results indicating that the impact of past pandemics on inequality has been greater in the more severe episodes (either in terms of number of cases or output effects) and the preliminary estimates on a small set of countries for the current pandemic, suggest that the distributional consequences of COVID-19 may be larger than those following earlier pandemics in our sample. Our estimates using data on past pandemics are likely to be a lower bound since COVID-19 is more widespread than the average health crisis in our sample, with the median country affected roughly corresponding to twice the pandemic episode at the 99th percentile of the distribution in our sample. However, evidence from previous pandemics shows that the response of income inequality to pandemics also depends on fiscal policy (Furceri et al. 2021). Our results are consistent with the notion that austerity breeds K-shaped recoveries: the rise in inequality is higher when fiscal policy is tighter, while when the fiscal response is supportive, inequality barely increases (Furceri et al. 2021). The short-term fiscal support provided during the current crisis has been unprecedented and, if not withdrawn prematurely, will help to mitigate the regressive effects of the pandemic.

Our results leave several questions for future research. First, the distributional effects of pandemic events are likely to vary considerably across countries, depending on country-specific characteristics, initial income distribution, the stringency of containment measures as well as the policy response. Second, there is growing evidence that the economic effects of COVID-19 may also vary between different segments of the population including by race, age, and gender. Third, the human cost of pandemics is also sadly higher in low-income groups, which are more prone to diseases and have often more limited access to health services. These issues need attention.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10888-022-09540-y>.

Acknowledgements The authors did not receive support from any organization for the submitted work.

Funding The data that support the findings of this study are available from the corresponding author upon request.

Declarations

Competing Interests The authors have no competing interests to declare that are relevant to the content of this article.

References

- Aspachs, O., Durante, R., Montalvo, J.G., Graziano, A., Mestres, J., Reynal-Querol, M.: Real-Time Inequality and the Welfare State in Motion: Evidence from COVID-19 in Spain. CEPR Discussion Paper n. 15118 (2020)
- Atkeson, A.: What Will Be the Economic Impact of COVID-19 in the US? Rough Estimates of Disease Scenarios. NBER Working Paper No. 26867 (2020)
- Aucejo, E.M., French, J., Araya, M.P.U., Zafar, B.: The impact of COVID-19 on student experiences and expectations: evidence from a survey. *J. Public Econ.* **191**, 104271 (2020)
- Barro, R.J., Ursua, J.F., Weng, J.: The Coronavirus and the Great Influenza Pandemic: Lessons from the “Spanish Flu” for the Coronavirus’s Potential Effects on Mortality and Economic Activity. NBER Working Paper No. 26866 (2020)
- Blundell, R., Costa Dias, M., Joyce, R., Xu, X.: COVID-19 and inequalities. *Fisc. Stud.* **41**(2), 291–319 (2020)
- Camacho, M., Palmieri, G.: Do economic recessions cause inequality to rise? *J. Appl. Econ.* **22**(1), 304–320 (2019)
- Coibion, O.Y., Weber, G.M.: Labour markets during the COVID-19 crisis: a preliminary view. NBER working paper no. 27017 (2020)
- Coibion, O., Gorodnichenko, Y., Kueng, L., Silvia, J.: Innocent bystanders? Monetary policy and inequality. *J. Monet. Econ.* **88**, 70–89 (2017)
- Crossley, T.F., Fisher, P., Low, H.: The heterogeneous and regressive consequences of COVID-19: evidence from high quality panel data. *J. Public Econ.* **2020**, 104334 (2020)
- De Haan, J., Sturm, J.-E.: Finance and income inequality: a review and new evidence. *Eur. J. Polit. Econ.* **50**, 171–195 (2017)
- Eichenbaum, M.S., Rebelo, S., Trabandt, M.: The Macroeconomics of Epidemics. NBER Working Paper No. 26882 (2020)
- Emmerling, J., Furceri, D., Monteiro, F.L., Loungani, P., Ostry, J.D., Pizzuto, P., Tavoni M.: Will the economic impact of COVID-19 persist? Prognosis from 21st century pandemics, vol. 119. IMF Working Papers, Washington (2021)
- Furceri, D., Loungani, P., Ostry, J.D.: The aggregate and distributional effects of financial globalization: evidence from macro and sectoral data. *J. Money Credit Bank.* **51**, 163–198 (2019). <https://doi.org/10.1111/jmcb.12668>
- Furceri, D., Loungani, P., Ostry, J.D., Pizzuto, P.: The rise in inequality after pandemics: can fiscal support play a mitigating role? In: Dosi, Stiglitz (eds.) *Industrial and Corporate Change*. OUP (2021)
- Hacioglu, S., Känzig, D., Surico, P.: The distributional impact of the pandemic. CEPR discussion paper, DP15101 (2020)
- Helsinki, G.S.E.: Situation room report 2.12.2021 – latest developments in the labor market, households and firms. Helsinki graduate School of Economics (2021) <https://www.helsinkigse.fi/corona/situation-room-report-2-12-2021-latest-developments-in-the-labor-market-households-and-firms/>. Accessed on 15 February 2022
- ILO.: ILO Monitor (7nd ed): COVID-19 and the world of work (2021)
- Jenkins, S.P.: World income inequality databases: an assessment of WIID and SWIID. *J. Econ. Inequal.* **13**(4), 629–671 (2015)
- Jordà, O.: Estimation and inference of impulse responses by local projections. *Am. Econ. Rev.* **95**, 161–182 (2005)
- Jordà, Ò., Taylor, A.M.: The time for austerity: estimating the average treatment effect of fiscal policy. *Econ. J.* **126**(590), 219–255 (2016)
- Jordà, O., Singh, S.R., Taylor, A.M.: Pandemics: long-run effects. *COVID Econ.* **1**, 1–15 (2020)

- Laeven, L., Valencia, F.: Systemic banking crises database II. *IMF Econ. Rev.* **68**, 307–361 (2020)
- Ma, C., Rogers, J., Zhou, S.: Global financial effects. *COVID Econ.* **5**, 56–78 (2020a)
- Ma, Y., Zhao, Y., Liu, J., He, X., Wang, B., Fu, S., Luo, B.: Effects of temperature variation and humidity on the death of COVID-19 in Wuhan, China. *Sci. Total Environ.* **724**, 138226 (2020b)
- Numn, N., Qian, N.: US food aid and civil conflict. *Am. Econ. Rev.* **104**(6), 1630–1666 (2014)
- Ostry, J.D., Loungani, P., Berg, A.: *Confronting Inequality: how Societies Can Choose Inclusive Growth*. Columbia University Press (2019)
- Romer, C.D., Romer, D.H.: The macroeconomic effects of tax changes: estimates based on a new measure of fiscal shocks. *Am. Econ. Rev.* **100**(3), 763–801 (2010)
- Schmitt-Grohé, S., Teoh, K., Uribe, M.: COVID-19: testing inequality in new York City. *COVID Econ.* **8**, 27–43 (2020)
- Solt, F.: Standardizing the world income inequality database. *Soc. Sci. Q.* **90**(2), 231–242 (2009)
- Stantcheva: *Inequalities in the Times of a Pandemic*. *Economic Policy* (2021)
- Surico, P., Känzig, D., Hacıoglu, S.: Consumption in the time of Covid-19: Evidence from UK transaction data. CEPR Discussion Paper, DP14733 (2020)
- Tenreiro, S., Thwaites, G.: Pushing on a string: US monetary policy is less powerful in recessions. *Am. Econ. J. Macroecon.* **8**(4), 43–74 (2016)
- Teulings, C.N., Zubanov, N.: Is economic recovery a myth? Robust estimation of impulse responses. *J. Appl. Econ.* **29**(3), 497–514 (2014)
- Ujii, M., Tsuzuki, S., Ohmagari, N.: Effect of temperature on the infectivity of COVID-19. *Int. J. Infect. Dis.* **95**, 301–303 (2020)
- World Bank: *Poverty and Distributional Impacts of COVID-19: Potential Channels of Impact and Mitigating Policies BRIEF APRIL 16, 2020*. World Bank (2020)

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