ARTICLE

OPEN

Global evidence on the economic effects of disease suppression during COVID-19

Jonathan T. Rothwell [1,2✉], Alexandru Cojocaru[3], Rajesh Srinivasan[1] & Yeon Soo Kim [3]

Governments around the world attempted to suppress the spread of COVID-19 using restrictions on social and economic activity. This study presents the first global analysis of job and income losses associated with those restrictions, using Gallup World Poll data from 321,000 randomly selected adults in 117 countries from July 2020 to March 2021. Nearly half of the world's adult population lost income because of COVID-19, according to our estimates, and this outcome and related measures of economic harm—such as income loss—are strongly associated with lower subjective well-being, financial hardship, and self-reported loss of subjective well-being. Our primary analysis uses a multilevel model with country and month-year levels, so we can simultaneously test for significant associations between both individual demographic predictors of harm and time-varying country-level predictors. We find that an increase of one-standard deviation in policy stringency, averaged up to the time of the survey date, predicts a 0.37 std increase in an index of economic harm (95% CI 0.24–0.51) and a 14.2 percentage point (95% CI 8.3–20.1 ppt) increase in the share of workers experiencing job loss. Similar effect sizes are found comparing stringency levels between top and bottom-quintile countries. Workers with lower-socioeconomic status—measured by within-country income rank or education—were much more likely to report harm linked to the pandemic than those with tertiary education or relatively high incomes. The gradient between harm and stringency is much steeper for workers at the bottom quintiles of the household income distribution than it is for those at the top, which we show with interaction models. Socioeconomic status is unrelated to harm where stringency is low, but highly and negatively associated with harm where it is high. Our detailed policy analysis reveals that school closings, stay-at-home orders, and other economic restrictions were strongly associated with economic harm, but other non-pharmaceutical interventions—such as contact tracing, mass testing, and protections for the elderly were not.

[1] Gallup Inc, Washington, DC, USA. [2] The Brookings Institution, Washington, DC, USA. [3] The World Bank, Washington, DC, USA.
✉email: Jonathan_Rothwell@Gallup.com

Introduction

In 2020, governments around the world started putting in place extraordinary policies to mitigate the spread of the coronavirus, including stay-at-home orders, restrictions on travel and gatherings, and closures of schools and workplaces. These policies have been found to be associated with reduced economic activity, in the form of visits to workplaces, parks, restaurants, and non-grocery retail establishments, both in global analyses (Hale et al. 2021), and based on detailed evidence from single countries or a subset of countries (Deb et al. 2020; Boone and Ladreit 2021; Lozano Rojas et al. 2020; Aminjonov et al. 2021; Carvalho et al. 2021; Coibion et al. 2020; Gathergood et al. 2021, including illegal economic activity (Nivette et al. 2021)). Using U.S. data from early in the pandemic, other scholars have argued that widespread economic disruptions would have happened in the absence of restrictions (Goolsbee and Syverson 2020; Forsythe et al. 2020; Gupta et al. 2020).

Epidemiological theory and empirical evidence suggest that these policies likely reduced the number of deaths (Liu et al. 2021; Chernozhukov et al. 2021; Violato et al. 2021; Qi et al. 2022), although given the methodological challenges, the evidence on the causal links between mobility restrictions and COVID-19 mortality is not entirely consistent (Berry et al. 2021; Herby et al. 2022; Spiegel and Tookes 2022). Similarly, public health policies that do not restrict economic activity—such as contact tracing (Fetzer and Graeber 2021) and surgical mask use (Abaluck et al. 2022)—have been found to reduce infections.

Regardless of the causes, pandemic-era economic distress has been widely documented. Global output contracted by 3.4% in 2020, with output contractions observed in 95% of countries, a scale that rivals that of the Great Depression (World Bank 2022), and global poverty increased (Mahler et al. 2021; Kim et al. 2021). Furthermore, low-income countries faced widescale income losses (Egger et al. 2021; Josephson et al. 2021). Within countries, the economic effects of the pandemic have been worse for households with relatively low socioeconomic status, as measured by income rank or educational attainment (Rothwell and Smith 2021; Narayan et al. 2022; World Bank 2022; Bundervoet et al. 2021; Kugler et al. 2021), which is consistent with past pandemics (Furceri et al. 2022).

This paper contributes to the literature by providing the most comprehensive analysis to-date on several research questions related to the pandemic: How prevalent was the economic harm related to the pandemic, and how did that harm relate to subjective well-being and financial security? What is the association between economic harm and the stringency of regulations on economic and social activity? How did harm vary by socioeconomic status within and across countries? How do estimated effects of stringency compare to alternative non-pharmaceutical interventions in terms of job loss and similar outcomes?

The analysis relies heavily on the Gallup World Poll, which used random samples of individuals in 117 countries representing nearly three-quarters of the global population. From July 2020 to March 2021, the survey collected detailed demographic data on income and education, subjective well-being measures, and information on several economic outcomes in which respondents were explicitly asked if they were caused by the pandemic. The resulting database provides the only harmonized quasi-global database available to study individual and country-level employment and income outcomes. By directly measuring forms of economic harm—rather than proxy measures such as mobility—these data provide insights that would otherwise be lost to history. Data on policy interventions, COVID deaths, and other contextual data were matched to individual responses using cumulative-to-date means, such that a respondent's self-reported degree of harm from COVID-19 could be compared to the policy regime used up to match the month of the interview. The primary analysis uses multilevel mixed modeling to simultaneously estimate associations with individual demographic characteristics—including socioeconomic status—and time-varying country-level variables, aggregated to month-year units. The database needed to replicate our analysis will be released upon publication, allowing other scholars to use the data for their own novel analyses.

Methods

The hypotheses tested in this paper require data on three main components of our empirical models: (i) measures of economic harm or welfare impact; (ii) measures of the stringency of economic restrictions imposed by governments in the wake of the COVID-19 pandemic, and (iii) measures of the disease burden of COVID-19. These are described in turn in this section. Summary statistics from the country-level and individual data are available in Supplementary Table 1. The analysis was conducted using Stata 17.0.

Economic harm measures. The main source of data supporting the analysis in this paper comes from surveys fielded by the Gallup World Poll between July 9, 2020, and March 3, 2021, with 321,386 observations of people aged 15 and older in 117 countries/territories.[1] The survey included demographic information as well as items related to health and well-being that were designed to be nationally representative for each country in the sample.[2] The relevant ethics statement is provided below.

The main focus of this analysis is on five survey items that broadly measure social or economic harm from COVID-19. The first item, fielded to all survey respondents, solicits an answer to the question, "In general, to what extent has your own life been affected by the coronavirus situation?" We recode responses as a binary variable, which takes a value of one if the response is "a lot," and zero if respondents reply "some" or "not at all."

The second item, applicable only to people working at the start of the pandemic, solicits whether respondents have experienced each of the following as a result of the coronavirus situation?

(i)   Temporarily stopped working at your job or business
(ii) Lost your job or business
(iii) Worked less hours at your job or business
(iv) Received LESS money than usual from your employer or business

Respondents are instructed to answer "Yes" or "No", or "Does not apply" if they did not have a job leading up to the pandemic. The World Poll includes a large number of other respondent-level variables, which we only briefly describe below in context. On average, 42% of adults responded that they were affected a lot; 24% reported permanently losing their job; 48% reported a temporary job loss, and 47% reported lost income (see Supplementary Table 9 for full text).

The advantage of the above survey items for the purpose of assessing the welfare impacts of the COVID-19 pandemic, vis-à-vis more traditional measures of economic changes, such as employment status or income, stems from the fact that the respondent is asked to attribute the severity of the overall impact, or the different aspects of job and income losses to the pandemic in a causal sense. This is important because many factors other than COVID-19 could cause people to lose their job, leave the labor force or experience emotional stress. Thus, while the analysis in this paper relies on cross-sectional variation, the framing of the key questions related to impacts help at least partially guard against common omitted variable bias concerns in such settings.

A second advantage of these items is that respondents are well-placed to know if an economic change in their life was caused by the pandemic. The event itself was highly salient and often highly disruptive to daily life, with clear time boundaries, tied to events like international and national emergency declarations and stay-at-home orders. In many cases, employers may have specifically told employees that the cause of their layoff was the pandemic, but even in the absence of that messaging, respondents would be well aware of the timing of their job loss and the circumstances leading up to it, which may include the business being shut down or customers canceling contracts or no longer showing up.

Stringency of economic restrictions measures. The analysis relates the above measures of economic harm to the stringency of restrictions on economic activity imposed by governments. We measure stringency of lockdowns by using data from the Oxford COVID-19 Government Response Tracker (Hale et al. 2021). The database evaluates national and sub-national government policies along various dimensions.

Data are coded on an ordinal scale, such that values are increasing with stringency. For business closures, 0 means no measures, 1 means recommended closing or recommended work from home, 2 requires the closing of some sectors, and 3 requires the closing of all but essential workplaces, such as healthcare offices and grocery stores. The data are collected for every day since the start of 2020, which facilitates our analysis. Since we are studying the cumulative economic effects up until the time of the survey, we want a measure of the cumulative lockdown up to that point, which we measure as the average stringency up until the month of the survey. Hale et al. (2021) constructed a "stringency index" that is the mean value of the stringency of the 8 containment policies and one health-related policy. The latter measures the degree to which public officials urged caution through coordinated efforts to promote social distancing and related behavioral changes across traditional and social media (see Supplementary Table 4 for full description). For the purposes of our analysis, we restrict the sample to data collected through March of 2021, to coincide with our sample collection period. We standardize the stringency index and its components to have mean 0 and a standard deviation of one across all 184 countries in the database.

In our decomposition analysis, we focus on these containment measures, as well as "health systems policies" which include regulations of facial coverings, contact tracing, testing, vaccination policy, and protections of the elderly. We omitted investment related policies because they are highly dependent on country budgets and GDP.

In most cases, each country has a single daily measure for each indicator. Sub-national data is available for only Brazil, Canada, the United Kingdom, and the United States. For these countries, we obtain population data from national statistical offices for the sub-regions and use these as weights, so that the national value is a population-weighted average of the subregional policies.[3]

It should be noted that stringency measures, such as those related to social distancing and restrictions to physical mobility, need not result in economic harm if the degree of compliance is low, either because people are unable or do not want to comply with the measures and the authorities lack the capacity to enforce them. Thus, it is important to ascertain that stringency measures are actually binding and do, in fact, lead to reduced mobility and economic activity. To verify this, we rely on several sources of data that speak to physical mobility and social distancing dynamics (see Supplementary Materials).

One direct way of looking at restrictions to mobility is with the aid of data from Google Community Mobility Reports. Using mobile phone location software, these data show the percentage change in visits to various places from a pre-COVID baseline (January 3 to February 6, 2020). We focus only on visits to retail and restaurants, described by Google LLC (2021) as: "places like restaurants, cafes, shopping centers, theme parks, museums, libraries, and movie theaters." Roughly half of the adult population in the sample reported direct contact with non-household members, and visits to restaurants and similar places were down 25% on the average day through the end of 2020.

One additional proxy measure for the degree of social distancing that is directly relevant to COVID-19 transmission is related to changes in the transmission of a parallel respiratory virus that was common before the pandemic. A large drop in the transmission of this parallel virus would suggest major behavioral changes relevant to disease transmission, whereas the absence of change in its transmission would suggest limited behavioral changes. Influenza is close to ideal in providing this analytic opportunity. The most significant problem is that flu cases are measured only based on testing, and flu testing conditional on symptoms—like COVID testing—is likely to vary by country. At the same time, by comparing pre-COVID flu rates to COVID-era flu rates, we control for unchanging country-level testing infrastructure and practices. We, therefore, believe these data provide a valid measure of changes in social distancing that include both policy-induced and non-policy-induced behaviors. These data are from the World Health Organization's FluNet and include total positive influenza cases per week by country for each year from 2016 to the 45th week of 2021. The use of weekly data allows us to adjust for seasonal effects, which vary by hemisphere and countries within hemispheres. We are interested in the percentage change in weekly cases during flu season before and after the pandemic, ending the analysis on the 12th week of 2021 to coincide with the World Poll data collection period. To identify the flu season for each country, we calculate the weekly share of cases from 2016 to 2019 and classify any week with at least 1% of annual cases as being part of flu season. For the United States, this would include weeks 1–16 and weeks 47 through 52 of every year. For Australia, in the Southern Hemisphere, it would include weeks 20–41. Unfortunately, FluNet data are far from being comprehensive and many countries are either entirely missing or only report a few weeks out of the year. We require 90% reporting coverage during flu-season weeks before and after the pandemic. We find that flu cases in the 2020–2021 flu season were just 18% of the mean number of flu cases measured from 2016 to 2019 flu season in 74 countries.

As part of our robustness checks, we analyze the relationship between stringency and changes in flu rates at the country-month level. For this analysis, we calculate positive flu cases in the current year (2020) relative to the previous year (2019) and the year before that (2018) and take the average of these two rates before aggregating this average to months. This gives us a measure of flu rates relative to previous years that varies by month and country.

We include an additional measure of subjective social distancing. In partnership with Facebook, the University of Maryland fielded a large-scale global daily survey of Facebook users (the COVID-19 World Symptom Survey Data), reweighted to be representative of the national population (Barkay et al. 2020). Using the University of Maryland API (Fan et al. 2020), we were able to get weighted data for the percentage of respondents who have reported having had direct contact (longer than 1 min) with people not staying with them within the past 24 h. Using data aggregated across 103 countries in 2020, 46% of respondents report direct contact with a standard deviation of 10%.

Disease burden measures. The disease burden is measured in terms of deaths per capita, which is preferable to the number of

COVID-19 cases. We do not claim that every national health system is equally likely to capture and correctly identify every COVID-19 death, but nearly every country has formal systems to record the causes of death. By contrast, the probability of seeking testing conditional on the experience of symptoms is highly contingent on factors that vary widely by country, such as cost, guidelines on testing priorities, and the availability of tests. Asymptomatic testing, moreover, also varies widely by country. In short, data on COVID-19 cases per capita are very noisy measures of the disease burden relative to deaths per capita.

To further guard against measurement error—and potential bias stemming from lack of reporting—we include model-based estimates of deaths from COVID-19 in our analysis from the University of Washington's Institute for Health Metrics and Evaluation (IHME 2021). That analysis uses actual all-cause mortality data for 56 countries, subtracts out known increases in deaths, and determines estimates of actual COVID deaths. The research team then models the ratio between reported deaths and actual deaths for every country to arrive at a measure of total COVID deaths. We regard these as credible alternative measures of disease burden, as argued by Wang et al. (2022). IHME is the source for both the official and estimate deaths used in our models.

Details of analysis. Our analysis tests models at the individual and country levels. For individual analysis, we study (1) how experiences of economic harm relate to subjective-well-being outcomes; (2) which demographic variables predict a greater risk of harm; (3) whether the relationship between low-socioeconomic status and harm is higher or lower in countries with high-stringency versus low-stringency.

Predicting well-being. The initial findings test whether our measures of economic harm predict subjective well-being at the individual level. We run linear OLS regression models of the following form, where W is the outcome of interest, θ is a vector of individual i demographic variables, C is an indicator for the country of residence (a country fixed effect) c, and the errors are clustered at the country level to account for within-country-level measurement error. Since respondents answer the survey at different times t, time periods are measured in months, and months fixed effects are included. In this setup, there are no country-level regressors, other than the fixed effect.

$$W_{i,c,t} = \beta_0 + \beta_1 \theta_{i,c,t} + C_c + M_t + \epsilon_{i,c,t} \qquad (1)$$

Predicting harm in a multilevel framework. The primary analysis combines country-level and individual-level data and therefore uses a multilevel model. We estimate the model using the *mixed* program in Stata v17, allowing for one unique variance parameter per random effect and maximum likelihood estimation. The variance-covariance matrix is calculated to allow intragroup correlation at the country level, where the data are structured by countries and by month-years, allowing random intercepts that vary by time and country ($\beta_{0,c,t}$ in Eq. 2). The dependent variable is economic harm H measured at the individual level i in country c during time t (Eq. 2). We include country-level time-varying variables, captured in X in (Eq. 4). These are cumulative-to-date measures of COVID-19 restrictions, economic support policies, and COVID-19 deaths per capita.

When written out formally, Eq. 2 captures the first-level individual specification. Harm varies by individual, country, and time period and so do the errors, intercepts, and predictors. Equation 3 represents level 2 (the time period). The mean outcome for individuals is modeled as a function of the time period and a random component. The time period mean varies by

country, since countries faced different disease and economic trajectories during the pandemic. Level 3 is modeled in Eq. 4. The mean outcome by country and time period is a function of the mean across all groups, a country and time-varying component, and a random country-varying component. Equation 5 combines the multiple levels into our preferred model. The fixed components are the first three terms, whereas the random components are the final three. The estimation procedure in Stata uses maximum likelihood. This exposition follows the discussion from Tascam Giorgio et al. (2009) and Oshchepkov and Shirokanova (2022).

$$H_{i,c,t} = \beta_{0,c,t} + \beta_1 \theta_{i,c,t} + \epsilon_{i,c,t} \qquad (2)$$

$$\beta_{0,c,t} = \delta_{0,c,t} + u_t \qquad (3)$$

$$\delta_{0,c,t} = \gamma_0 + \gamma_1 X_{c,t} + v_c \qquad (4)$$

$$H_{i,c,t} = \gamma_0 + \gamma_1 X_{c,t} + \beta_1 \theta_{i,c,t} + \epsilon_{i,c,t} + u_t + v_c \qquad (5)$$

Country-level variables are cumulative-to-date time-varying for several reasons. A single cumulative measure would include information that occurred after measurement for survey respondents who interviewed in early waves, and this would introduce unnecessary error into the model. A time-varying metric that is not cumulative-to-date would be a problem, because the outcome variable measures cumulative harm-to-date, as in "have you ever lost your job as a result of the coronavirus situation?" Since a measure that ignores the past can hardly be expected to predict the past, this approach would also introduce error.

The individual-level measures are demographic indicators for age, gender, foreign-born status, education, income, and urbanicity. We also include an indicator for whether the respondent is out of the labor force at the time of the survey. Since most of our measures of harm involve job loss, they are not usually applicable to those who were out of the labor force, whose lives were less likely to have been affected. We omit current unemployment status because many people recently harmed through job loss or one of the other measures may still be unemployed at the time of the survey.

The results from our baseline model can be used to assess the appropriateness of our multilevel modeling strategy. Both of the random intercepts are significant at 95% confidence levels (see Table 1). The country level explains approximately 8.8% of the total variation, whereas the time-period effect is just 0.8%. The intraclass correlations are 8% at the country level and 8.8% combined and both significant, confirming our assumption that individual-level errors are correlated with higher-level errors. The model's results are reported in Table 1 using the harm index and job loss as the predicted outcomes; results predicting income loss, whether the respondent was affected a lot by the pandemic, temporary job loss, and loss of hours are reported in the Supplemental Materials (ST5, ST6, ST7, and ST8).

We also report the results of models that interact household income quintile with stringency (Fig. 2). These models are identical to our primary specification except they include additional interaction terms along the lines of

$$\beta_2 X_{c,t} \times \theta_{i,c,t}$$

where β_2 identifies the slope of harm for an income group as stringency increases. Figure 2 plots the mean predicted values from these models for each quintile after collapsing the data to centiles of stringency. Since we are interested in how effects vary by socioeconomic status, we drop educational attainment levels from the model (which is included in our benchmark model), so that the income effects are not conditional on education level. Standard errors are estimated in the plots by regressing the

Table 1 a Multilevel model regressing harm and job loss on country-level restrictions on economic activity and individual-level predictors. b Multilevel model regressing job loss on country-level restrictions on economic activity and individual-level predictors.

a

	Harm index					
	1	2	3	4	5	6
Country-level variables						
Stringency index—cumulative-to-date (standardized)	0.314** (0.053)	0.288** (0.054)	0.316** (0.053)	0.285** (0.054)	0.308** (0.053)	
Economic support index—cumulative-to-date (standardized)	−0.068 (0.040)	−0.058 (0.039)	−0.076 (0.039)	−0.062 (0.038)	−0.085* (0.037)	−0.076* (0.036)
Officially reported COVID-19 deaths per capita	−0.001 (0.001)	−0.001 (0.001)	−0.002 (0.001)			
Visits to restaurants and related establishments			0.007* (0.003)		0.006 (0.003)	0.005 (0.003)
Estimated actual COVID-19 deaths per capita				−0.000 (0.001)	−0.000 (0.001)	−0.000 (0.001)
1st quintile of stringency (ref = 5th)						−0.292** (0.041)
2nd quintile of stringency (ref = 5th)						−0.243** (0.034)
3rd quintile of stringency (ref = 5th)						−0.156** (0.028)
4th quintile of stringency (ref = 5th)						−0.106** (0.024)
Individual-level variables						
Not in labor force		−0.184** (0.017)	−0.187** (0.018)	−0.184** (0.017)	−0.187** (0.018)	−0.185** (0.018)
City (relative to rural)		0.085** (0.010)	0.076** (0.010)	0.085** (0.010)	0.076** (0.010)	0.077** (0.010)
Suburb (relative to rural)		0.074** (0.011)	0.066** (0.010)	0.074** (0.011)	0.066** (0.010)	0.066** (0.010)
Town (relative to rural)		0.036** (0.008)	0.034** (0.008)	0.036** (0.008)	0.034** (0.008)	0.034** (0.008)
Born outside of country		0.069** (0.012)	0.064** (0.012)	0.069** (0.012)	0.064** (0.012)	0.067** (0.012)
Elementary education (relative to tertiary)		0.096** (0.014)	0.100** (0.014)	0.096** (0.014)	0.100** (0.014)	0.100** (0.015)
Secondary education (relative to tertiary)		0.068** (0.010)	0.066** (0.011)	0.068** (0.010)	0.066** (0.011)	0.065** (0.011)
Age 15–19		−0.067* (0.028)	−0.074* (0.030)	−0.067* (0.028)	−0.074* (0.030)	−0.073* (0.030)
Age 20–29		0.055* (0.026)	0.049 (0.027)	0.055* (0.026)	0.049 (0.027)	0.050 (0.027)
Age 30–39		0.110** (0.024)	0.105** (0.025)	0.110** (0.024)	0.105** (0.025)	0.108** (0.025)
Age 40–49		0.091** (0.025)	0.086** (0.025)	0.091** (0.025)	0.086** (0.025)	0.088** (0.026)
Age 50–64		0.045 (0.024)	0.043 (0.025)	0.045 (0.024)	0.043 (0.025)	0.045 (0.025)
Age 65–79		−0.001 (0.020)	−0.000 (0.021)	−0.001 (0.020)	−0.000 (0.021)	0.001 (0.021)
Married or with domestic partner		−0.029** (0.007)	−0.029** (0.007)	−0.029** (0.007)	−0.029** (0.007)	−0.030** (0.007)
Live with children under age 15		0.030** (0.007)	0.026** (0.007)	0.030** (0.007)	0.026** (0.007)	0.025** (0.007)
Female living with children under 15		−0.020* (0.009)	−0.017 (0.009)	−0.020* (0.009)	−0.017 (0.009)	−0.015 (0.009)
Female		0.007 (0.007)	0.007 (0.008)	0.007 (0.007)	0.007 (0.008)	0.007 (0.008)
Bottom-income quintile (relative to top-fifth)		0.200** (0.014)	0.206** (0.015)	0.200** (0.014)	0.206** (0.015)	0.209** (0.015)
Second income quintile (relative to top-fifth)		0.163** (0.012)	0.169** (0.012)	0.163** (0.012)	0.169** (0.012)	0.171** (0.012)
Third income quintile (relative to top-fifth)		0.124**	0.130**	0.124**	0.130**	0.132**

Table 1 (continued)

a

	Harm index					
	1	2	3	4	5	6
		(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Fourth income quintile (relative to top-fifth)		0.064**	0.066**	0.064**	0.066**	0.067**
		(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Constant	0.002	−0.226**	−0.045	−0.228**	−0.061	0.060
	(0.022)	(0.034)	(0.096)	(0.033)	(0.096)	(0.092)
Random effect: county-level variance	−1.523**	−1.558**	−1.602**	−1.557**	−1.600**	−1.573**
	(0.069)	(0.071)	(0.074)	(0.072)	(0.076)	(0.079)
Random effect: month-level variance	−2.681**	−2.710**	−2.711**	−2.710**	−2.710**	−2.729**
	(0.074)	(0.074)	(0.078)	(0.074)	(0.078)	(0.084)
Residual variance	−0.318**	−0.344**	−0.348**	−0.344**	−0.348**	−0.348**
	(0.011)	(0.010)	(0.011)	(0.010)	(0.011)	(0.011)
Log-likelihood	−329,253	−289,861	−265,161	−289,861	−265,162	−260,937
Sample size	298,985	269,725	247,679	269,725	247,679	243,636

b

	Lost job, permanently					
	1	2	3	4	5	6
Country-level variables						
Stringency index—cumulative-to-date (standardized)	0.159**	0.142**	0.156**	0.146**	0.158**	
	(0.031)	(0.030)	(0.029)	(0.030)	(0.030)	
Economic support index—cumulative-to-date (standardized)	−0.059**	−0.054**	−0.064**	−0.051*	−0.063**	−0.061**
	(0.022)	(0.020)	(0.021)	(0.020)	(0.021)	(0.020)
Officially reported COVID-19 deaths per capita	0.000	0.000	−0.000			
	(0.001)	(0.001)	(0.001)			
Visits to restaurants and related establishments			0.005**		0.005**	0.004**
			(0.002)		(0.002)	(0.002)
Estimated actual COVID-19 deaths per capita				0.000	−0.000	0.000
				(0.000)	(0.000)	(0.000)
1st quintile of stringency (ref = 5th)						−0.137**
						(0.029)
2nd quintile of stringency (ref = 5th)						−0.125**
						(0.027)
3rd quintile of stringency (ref = 5th)						−0.088**
						(0.026)
4th quintile of stringency (ref = 5th)						−0.065**
						(0.025)
Individual-level variables						
Not in labor force		−0.049**	−0.051**	−0.049**	−0.051**	−0.052**
		(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
City (relative to rural)		0.019**	0.016*	0.019**	0.016*	0.016*
		(0.007)	(0.006)	(0.007)	(0.006)	(0.006)
Suburb (relative to rural)		0.020**	0.019**	0.020**	0.019**	0.018**
		(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Town (relative to rural)		0.008	0.008	0.008	0.008	0.008
		(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Born outside of country		0.030**	0.025**	0.030**	0.025**	0.026**
		(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Elementary education (relative to tertiary)		0.099**	0.100**	0.099**	0.100**	0.101**
		(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Secondary education (relative to tertiary)		0.051**	0.049**	0.051**	0.049**	0.049**
		(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Age 15–19		−0.026	−0.029	−0.026	−0.029	−0.030
		(0.016)	(0.017)	(0.016)	(0.017)	(0.017)
Age 20−29		0.053**	0.049**	0.053**	0.049**	0.051**
		(0.014)	(0.015)	(0.014)	(0.015)	(0.015)
Age 30−39		0.052**	0.048**	0.052**	0.048**	0.050**
		(0.013)	(0.012)	(0.013)	(0.012)	(0.012)
Age 40−49		0.031*	0.028*	0.031*	0.028*	0.030*
		(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Age 50−64		0.007	0.005	0.007	0.005	0.007
		(0.013)	(0.013)	(0.013)	(0.013)	(0.013)

Table 1 (continued)

b

	Lost job, permanently					
	1	2	3	4	5	6
Age 65−79		−0.027*	−0.026*	−0.027*	−0.026*	−0.025
		(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Married or with domestic partner		−0.024**	−0.023**	−0.024**	−0.023**	−0.024**
		(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Live with children under age 15		−0.005	−0.007	−0.005	−0.007	−0.008
		(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Female living with children under 15		0.010	0.012*	0.010	0.012*	0.012*
		(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Female		−0.001	−0.001	−0.001	−0.001	−0.001
		(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Bottom-income quintile (relative to top-fifth)		0.169**	0.171**	0.169**	0.171**	0.172**
		(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Second income quintile (relative to top-fifth)		0.115**	0.119**	0.115**	0.119**	0.118**
		(0.008)	(0.008)	(0.008)	(0.008)	(0.009)
Third income quintile (relative to top-fifth)		0.082**	0.086**	0.082**	0.086**	0.086**
		(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Fourth income quintile (relative to top-fifth)		0.042**	0.043**	0.042**	0.043**	0.043**
		(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Constant	0.089**	0.089**	0.221**	0.092**	0.223**	0.290**
	(0.019)	(0.019)	(0.053)	(0.018)	(0.052)	(0.056)
Random effect: county-level variance	−2.147**	−2.147**	−2.169**	−2.151**	−2.170**	−2.152**
	(0.088)	(0.088)	(0.092)	(0.087)	(0.092)	(0.102)
Random effect: month-level variance	−3.347**	−3.347**	−3.347**	−3.343**	−3.347**	−3.355**
	(0.082)	(0.082)	(0.089)	(0.082)	(0.088)	(0.094)
Residual variance	−0.920**	−0.920**	−0.924**	−0.920**	−0.924**	−0.927**
	(0.021)	(0.021)	(0.022)	(0.021)	(0.022)	(0.022)
Log-likelihood	−106,704	−106,704	−98,179	−106,704	−98,178	−96,003
Sample size	221,921	221,921	205,125	221,921	205,125	201,699

Standard errors in parentheses, robust to heteroskedasticity and adjusted for autocorrelation within time-period and country. $**p < 0.01$, $*p < 0.05$. Google mobility trends refer to visits to places like restaurants, cafes, shopping centers, theme parks, museums, libraries, and movie theaters relative to pre-pandemic patterns.

group-specific effect sizes on the stringency centile rank. These approximate the standard errors from the larger database.

To test the differences between policies, we replicate the analysis from Eq. 5 and Table 1 using our preferred multilevel model and report the coefficients and standard errors (Fig. 3).

Results

Summary data. Across the 117 countries included in the Gallup World Poll from July 2020 to March 2021, 42% of adults said they were affected a lot by the pandemic, weighting responses by population. Among those who were in the labor force leading up to the pandemic, 51% were laid off temporarily, 50% lost hours, 49% lost income, and 27% lost their job (see ST3).

These outcomes varied widely by country and continental sub-region. In Eastern Asia, Western Europe, and Northern Europe, only 4.3%, 6.4%, and 6.8% permanently lost their job, respectively, but in Southern and South-eastern Asia it was 49.0% and 44.1%, respectively. In Northern America and Western Asia, 50.1% and 57.4% said their lives were affected a lot. In Western Europe, this was just 29.5%. Meanwhile, cumulative deaths per capita were much higher in Europe and North America relative to Africa and South Asia (ST2), suggesting that the disease burden is unlikely to explain these findings.

A general pattern, found in the data, is that low-income countries experienced a relatively low disease burden from COVID-19 but a high economic burden. This mismatch between the health burden of the pandemic and its social burden suggests an important role for policy. GDP per capita measured in 2019 PPP-adjusted dollars is negatively correlated with the share of

population reporting a COVID-related job loss (−0.74), but positively correlated with deaths per capita (0.35) and estimated deaths per capita (0.18), using data from (Wang et al. 2022). GDP per capita is also highly correlated with an economic support index (0.55). Yet, GDP per capita has no correlation with the stringency index for disease suppression policies (0.02), even though stringency predicts greater job loss (0.19) and economic support predicts less job loss (−0.40).

Importantly, these relationships would be missed using Google mobility as economic indicators. Visits to restaurants were negatively correlated with GDP per capita (−0.20) and positively correlated with job loss (0.15). In other words, Google data provides the opposite signal as survey-based data. Other Google mobility measures showed the same pattern, including use of transportation and visits to work. It seems that in rich countries, people were able to withdraw from discretionary in-person economic activity—including work—while preserving their jobs and income to a much greater extent than in low-income countries, likely because of the development of digital service markets.

Validating a novel measure of economic harm. Before describing the primary results, we establish grounds for accepting the validity of our key measures. Further information is provided in the Methods section and Supplementary Text. First, we create a "harm index" as the standardized individual-level mean of responses to five survey items about how respondents' lives have been affected by the coronavirus situation. They are as follows: whether their lives have been affected a lot, whether they lost their

Fig. 1 Estimated mean effect and confidence interval of different forms of economic harm on probability that respondent's living standard is getting worse. Data are from the Gallup World Poll. Analysis includes approximately 222,000 respondents when restricted to the working population, which is used for the economic outcome measures. All models include demographic controls and country effects. The red diamonds show results when all variables, except the harm index, are used in the same model.

job or business temporarily or permanently (two distinct items), whether they worked fewer hours, or whether they received less money. Using the global sample, each item is standardized to have a mean of zero and a standard deviation of one.

The results show that the economic harm index—and its component parts—strongly predict four measures of subjective well-being, covering (1) changes in subjective living standards, (2) current life evaluation, (3) experiences of worry, and (4) lack of money for food (wording is provided in Supplementary Table 3). Essentially, we regress these outcomes on the harm index, controlling for respondent demographics and country fixed effects. Each component of harm is strongly and significantly associated with lower well-being using all four measures. Moreover, when each component of harm is included in the same model, all of them are significant, except for the loss of hours, which is highly correlated with the others. Since each variable adds information, we consider the harm index to be the most comprehensive measure of several dimensions: job loss, income loss, and subjective disruption to life (Fig. 1).

We considered several alternative measures of our harm index, including a factor analysis-based index, one that only uses the four labor market items (excluding whether the respondent was affected more generally), and one that combines temporary and permanent layoffs. Based on empirical investigations discussed in the Supplemental Text and summarized in Supplementary Table 10, our preferred measure is the one used here, though the results

reported in Table 1a—testing the association with stringency—are almost exactly the same, when we replace the harm index with these alternatives.[4]

Next, we check the reliability and validity of World Poll data on employment losses against alternative sources. World Poll data on the job loss rate are broadly aligned with administrative data on changes in the official unemployment rate (correlation is 0.52 in 52 countries). Yet, in addition to broader coverage, the World Poll measures are superior in two respects: harmonization in measurement and a causal link with COVID-19. Note, we are not suggesting that this fact implies that our estimators are causal. The point is that respondents are asked to report on an outcome that they believe is causally linked to the pandemic. This is a different question about whether they believe it is causally linked to stringent policies, which is a much harder question. Nonetheless, this is a large conceptual advance over asking whether someone is employed or not and assuming any change from pre- to post-pandemic is caused by the pandemic.

Consider that in normal times, except in rare cases, "unemployment" requires that adults are out of work but seeking and able to work. If the latter two conditions are unmet, the person is considered out of the labor force, but not unemployed. COVID-19, however, resulted in many people losing their job but temporarily halting efforts to find a new one—for various reasons. Statistical offices around the world took different non-harmonized approaches to classifying such persons, resulting in

different methodological bases for documenting unemployment rate levels and changes. Moreover, COVID-19 was not the only causal factor affecting social and economic conditions around the world, so the World Poll data also improve conceptual validity by asking respondents to attribute their economic harm to the pandemic and allowing them to express it along several dimensions (see Supplementary Text for further discussion of these issues).

Finally, we show that our primary policy measure also meets basic validity criteria, as discussed in Hale et al. (2021). Stringency is weakly and positively related to COVID death rates but more closely related to measures of social distancing, particularly those involving declines in visits to restaurants and small businesses (Supplementary Text and Supplementary Fig. 1). We examine two additional and related outcomes in the supplement. In more stringent countries, self-reported social contact (available from a non-representative alternative survey covering a smaller number of countries) tends to be lower and reported cases of seasonal flu fell further from baseline season-adjusted trends—for the subset of countries with high-quality flu data. This provides further evidence that the behaviors associated with respiratory disease transmission (e.g., social contact) fell further where disease-suppression policies were strongest, but flu case data are likely more informative than COVID-19 case count data, since flu surveillance systems were well-established before 2020, and COVID surveillance relied on novel tests that were neither available uniformly globally nor across regions within countries. Taken together, this evidence suggests a plausible link between stringency and economic outcomes.

Stringency measures and economic harm: main results. We now proceed with the main research question—whether more stringent restrictions are associated with a greater degree of economic harm, and what demographic factors are most strongly associated with harm. The stringency of mitigation policies is measured by the COVID-19 Government Response Tracker (Hale et al. 2021). Harm is aggregated from the World Poll microdata, using sample weights to ensure national representation. The analysis regresses harm on stringency (see Methods) and a vector of demographic variables in a multilevel model, with country and month-year levels.

Column (1) of Table 1a reports the regression-adjusted correlation between policy stringency and an index of economic harm, with no individual-level controls. A one standard deviation increase in the stringency index predicts a 0.31 increase in harm (0.40 for a std deviation unit of harm). Colum (2) adds individual-level demographic controls. The coefficient on harm falls—in absolute value terms—only slightly to 0.29 increase in harm (0.37 std dev). This model includes cumulative-to-date measures of reported COVID-19 deaths per capita and an index of economic support, as well as a rich list of individual-level controls, and country and month effects.

The economic support index is measured by the Oxford database (Hale et al. 2021) and captures the record of the government providing direct cash payments to people who lose their jobs or cannot work, including payments to firms that are linked to payroll/salaries, as well as the record of the government freezing financial obligations for households (e.g., stopping loan repayments, preventing services like water from stopping or banning evictions). Models 1–3 suggest that the degree of economic support is not correlated with the extent of economic harm, but a significant and negative effect is found after adjusting for the observed economic behavior of the population (using Google mobility) and using a measure of COVID deaths per capita that considers measurement error in reporting (columns 5

and 6 of Table 1a). These latter results more closely approximate the country-level bivariate correlation between harm—aggregated using all observations—and economic support averaged through 2020, which is negative ($r = -0.32$).

In our conceptual model, stringency may be confounded with behavioral changes—such as voluntary social distancing—that would have happened even in the absence of government regulations. To account for this, in columns/models 3, 5, and 6, we control for a cumulative-to-date measure of visits to restaurants, cafes, and discretionary (non-grocery) shopping, using cell-phone-based data from Google Mobility. Conditional on government policies, visits to restaurants predict greater harm in model 3, but there is no significant relationship when using the error-corrected measure of deaths.

As another robustness check, in columns 3–6, we replace reported deaths per capita with a model-based measure of actual deaths, based on seroprevalence rates and other observable factors. These estimates are generated in (IHME 2021) as discussed in (Wang et al. 2022). Using either measure, the correlation between COVID deaths per capita and economic harm does not reach significance at 95% confidence levels.

Finally, column six drops the continuous measure of stringency in favor of binary measures of stringency set to equal one for each quintile of severity. The most stringent quintile is the omitted reference group. This setup accounts for potential non-linear effects of stringency. The quintiles are all negative, significant, and decreasing in a roughly linear pattern, such that lower levels of stringency predict less harm at each point in the distribution, though not entirely to the same extent. The coefficient on the first quintile is -0.30, which is very close to the corresponding coefficient in column 5 (-0.31), which uses the continuous stringency measure in an otherwise identical model. Thus, comparing the top-to-bottom-quintile of stringency yields a result that is well-approximated by a unit of the stringency index. The marginal effect sizes seen from comparing one quintile to the next range from -0.11 (comparing the fifth quintile to the fourth or the third to the second) to -0.05 (comparing the fourth to the third or second to the first).

To test whether our results are sensitive to the use of our index—and the assumptions underlying it—we run the same six models using permanent job loss (whether the respondent lost his or her job as a result of the pandemic) as the dependent variable (Table 1b). Coefficients on stringency range from 14.2 to 15.9 ppts (with a margin of error of approximately 5.9 at 95% CI). The job loss rate has a standard deviation of 0.45 (see ST1). Converting the effect size to standardized units yields values ranging from 0.32 to 0.35 ppt (MOE of 0.13 at 95% CI), which are slightly smaller than those found for the harm index: 0.37–0.41 (MOE of 0.14 at 95% CI).

Relative to the most stringent quintile, adults in the least stringent countries were 14 ppt less likely to have experienced job loss, which is slightly less in absolute magnitude than the comparable linear estimate (16 ppt in column 5 of Table 1b). The marginal effects on job loss rates of a one-unit change in the quintile range from 1.2 to 6.5 ppt. Economic support is negatively and significantly related to job loss, but we cannot reject the null hypothesis that the number of deaths per capita has no effect on job loss, and visits to restaurants predict slightly more job loss, surprisingly.

In the supplemental materials, we show similar results using income loss (ST5), whether the respondent was affected a lot (ST6), temporary job loss (ST7), and loss of hours (ST8) as the dependent variables. Stringency is strongly and positively related to each outcome. Economic support predicts less adverse outcomes with respect to loss of hours, temporary layoff, but not whether the respondent was affected a lot. The results of

economic support were mixed for income loss, showing some evidence that these policies may have mitigated income loss.

These models also allow us to see which demographic variables are most closely related to harm, and to test the relationship between socioeconomic status and harm. The ratio of coefficients to standard errors (the t-statistic in OLS models and a z-statistic in random effects models) provides a valid measure of variable importance, as pointed out by Bring (1996). Columns 2 of Table 1a, b are our preferred models, as they allow government restrictions on economic activity to affect visits to restaurants and use reported COVID deaths.

By variable importance, income status was more closely linked to harm than any other demographic variable considered in our model, and lower-income predicted more harm, consistent with the literature. Moreover, even with income in the model, higher educational attainment is strongly associated with less harm. Socioeconomic status—measured through income or education—is more predictive than urbanicity, gender, foreign-born status, age, child-rearing status, or marital status. These results are consistent with the findings of several earlier studies that similarly find greater job losses among the more vulnerable population groups (Bundervoet et al. 2021; Kugler et al. 2021; Narayan et al. 2022; World Bank 2022; Rothwell and Smith 2021), but confirm them across a much wider and more comprehensive number of countries during the same time period.

The z-stat for those in the second-lowest quintile is 14.1, and it is 14.0 for those at the bottom, with coefficients of 16 and 20 ppt, respectively. These are the two largest z-statistics in absolute value terms. Workers in the bottom-quintile experienced 0.26 standard deviations of additional harm relative to those in the top quintile. Workers with an elementary education also saw a large increase in harm (9.6 ppt) compared to those with a tertiary (or college) education. Women did not experience more or less harm than men, but women with children under 15 experienced less harm, as did married couples or those living as domestic partners, relative to those living alone or in other arrangements. Young adults aged 30 to 39 experienced more harm than any other age group. By urbanicity, residents of cities saw the most harm, whereas workers in rural areas experienced the least harm. Foreign-born residents experienced significantly more harm than those born domestically. People out of the labor force at the time of the survey—and possibly throughout the pandemic—reported less harm, likely because they had less to lose.

Looking at job loss (Table 1b) and other outcomes (ST5-ST8), the patterns are broadly similar. The largest estimated effect (16.9 ppt) and largest z-statistic (16.9) is for those at the bottom quintile of the household earnings distribution. The next largest effect is for those in the second-lowest quintile (12 ppt; 13.9 z-stat), followed by those with elementary education (10 ppt; 12.7 z-stat). There are no gender differences in job loss risk, and workers who are married or living with a domestic partner experienced less risk. Again, young adults aged 30 to 39 (5.2 ppt; 4.1 z-stat).

From the perspective of variable importance, the stringency index is less important than household income but more important than age, gender, and some measures of urbanicity in predicting harm (5.3 z-stat). It is comparable to secondary education (6.6 z-stat) and foreign-born status (5.7) In models predicting job loss, stringency is more important than foreign-born status and urbanicity, but less important than education or income (4.8 z-stat).

In the supplemental text, we discuss an out-of-sample test of the relationships between stringency and loss of income, using alternative U.S.-based data sources organized at the state level. Using a linear OLS model, we find that cumulative stringency predicts 0.6 standard deviations of additional COVID-related job loss (3 ppt) across U.S. states, controlling for median household income, COVID deaths per capita, and the percent of households who telework. The results are even stronger using pre-COVID party control as an instrumental variable in a two-stage least squares model. While party control is strongly related to policy, it is unclear if it is truly exogenous to loss of income from COVID-19 through channels other than COVID-related policies (ST11). We regard these results as supportive but still quite limited in establishing a causal relationship.

Socioeconomic status and heterogenous effects of stringency. The previous discussion considered the average country-level effects of policy stringency on economic harm. Here, we consider that public health policies, even when they are implemented uniformly at the national level, may not affect all households and individuals in the same way. To do so, we add interaction effects to our baseline multilevel models, multiplying stringency with binary variables for each income quintile. Using the results of this model, we forecast the harm index and job loss separately for each income group (see methods).

If lower-income households were more likely to be affected by restrictions on economic and social activity, then we should find a steeper slope between harm and stringency for lower-income groups relative to higher-income groups. This is exactly what we find (Fig. 2). The slope on the stringency-bottom-income quintile coefficient is 0.12 ppt higher than the slope on stringency-top income for the harm index and job loss rate; it is 0.08 for whether the respondent was affected a lot and 0.04 for temporary layoffs. Except for temporary layoffs, the results are significant at 99% confidence levels.

In the least restrictive policy regimes by centile rank, there is little or no difference in outcomes between the highest and lowest-income groups. For example, using the job loss rate, the gap is negative (6.5% for highest versus 5.2% for lowest) for the bottom centile of stringency, but it expands rapidly as stringency increases. In the top centile of stringency, the gap is very large (18.8% for highest-income group versus 46.9% for the lowest), with small error bars. Thus, this is strong evidence that policies meant to suppress the spread of the coronavirus were associated with a widening of economic inequality between income groups.

Alternative policy responses and economic harm. We consider that restricting social and economic activity was not the only tool available to public health officials. Widespread testing, meticulous contact tracing, social distancing focused on the elderly, travel restrictions, and use of facial coverings outside of one's home are alternatives to universal social distancing and some do not necessarily limit economic behavior. The Oxford database tracks these and several other policies, and we tested these in our preferred multilevel model that controls for COVID-19 deaths and economic support.

In predicting the harm index (Fig. 3, top panel), we find that some, but not all policy measures are positively associated with economic harm, and many policies have no significant relationship with harm. While no policies were negatively and significantly related to harm, vaccination policy, restrictions on gathering, public information, testing policy, contact tracing, and protection of the elderly were all insignificant, with the latter two having negative coefficients. In order of effect size, the overall index, school closings, internal travel restrictions, the closing of public transportation, stay-at-home-orders, canceling public events, mask orders, workplace closings, and international travel restrictions were each significantly associated with harm. The results and policy rankings are very similar using job loss as the dependent variable (Fig. 3, bottom panel). The only substantive difference is that restrictions on gatherings are significantly associated with job loss but not the harm index.
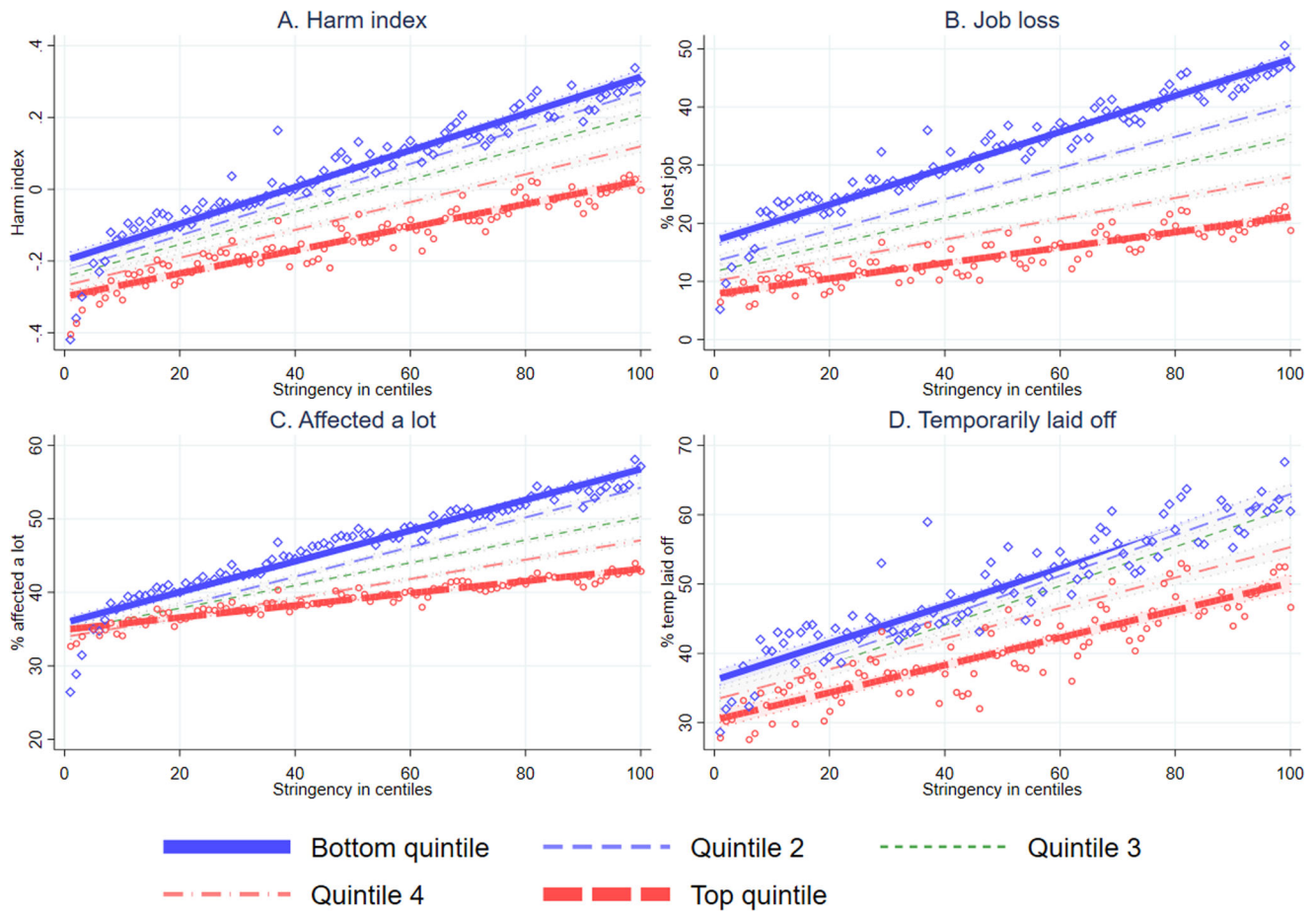
Fig. 2 Testing for heterogenous associations between stringency and socioeconomic status by within-country quintile of household income. Plots model-predicted outcome by income group, where outcome is measured as the Harm index (**A**); whether the respondent experienced permanent job loss (**B**); whether respondent is affected a lot by COVID (**C**); whether respondent was temporarily laid off (**D**). Each model is a weighted mixed multilevel linear regression with country and month-year levels and interactions between cumulative-to-date stringency and within-country household income quintiles. Other control variables include whether the respondent was out of the labor force at time of survey, age group, gender, foreign-born status, marital status, the presence of children, and a category for level of urbanicity. Sample size is 269,725 observations in 117 countries. 95% confidence intervals are shaded around the prediction line. Point-estimates are shown as hollow diamonds or circles for top and bottom-income quintiles, respectively.

Discussion

The restrictions to mobility and economic activity deployed by governments to contain the spread of coronavirus presumably saved many lives, by delaying transmission until effective treatments and vaccines counterhardship for many households around the world, and restrictions may have exacerbated the scale and severity of hardship. Our research uses the best available data to describe these outcomes, how they varied within and across countries, and to what extent they were linked to pandemic restrictions and other public health policies.

Soon after the start of the pandemic and throughout the first year, we document that many people around the world believed they were affected a lot and/or experienced job loss, income loss, or loss of hours, because of the pandemic. These outcomes were associated with negative levels and self-reported changes in subjective well-being and hardship. Moreover, this hardship was greater—around the world—for people with lower socioeconomic status, as measured by educational attainment or within-country income rank.

People were not asked whether or not their government's response to the pandemic was harmful or helpful or how it affected them across various dimensions, and it would be difficult for any individual to know. We, therefore, use variation in the timing and severity of government responses across 117 countries to estimate the association between government policies and self-described economic harm at the individual level.

We find robust evidence that countries adopting more stringent disease suppression policies experience a higher rate of economic harm. This is found in models that simultaneously account for individual-level demographic variables, month-specific disease dynamics, and time-varying country measures of the disease burden and policy environment. These partially confirm the work of previous scholars conducting country-specific or regional analyses and add richer, more comprehensive evidence to support that work. Moreover, the associations between stringency and harm are stronger for individuals with lower levels of baseline socioeconomic status.

In addition, we find that stringent economic and social restrictions predict more severe harm globally, but other public health efforts—such as contact tracing, widespread vaccination, and special protections of the elderly did not predict economic harm.

A limitation of this analysis is that disease-suppression policies —like all government laws—are not randomly assigned, even though many of the policies studied were novel in their application as of 2020. We cannot rule out the possibility that our results are biased by omitted time-varying variables operating at the country level. Few variables successfully predict the disease burden of COVID, making instrumentation difficult, though we attempted such an analysis using U.S. states, as discussed in the supplemental materials. Another limitation—one shared by all
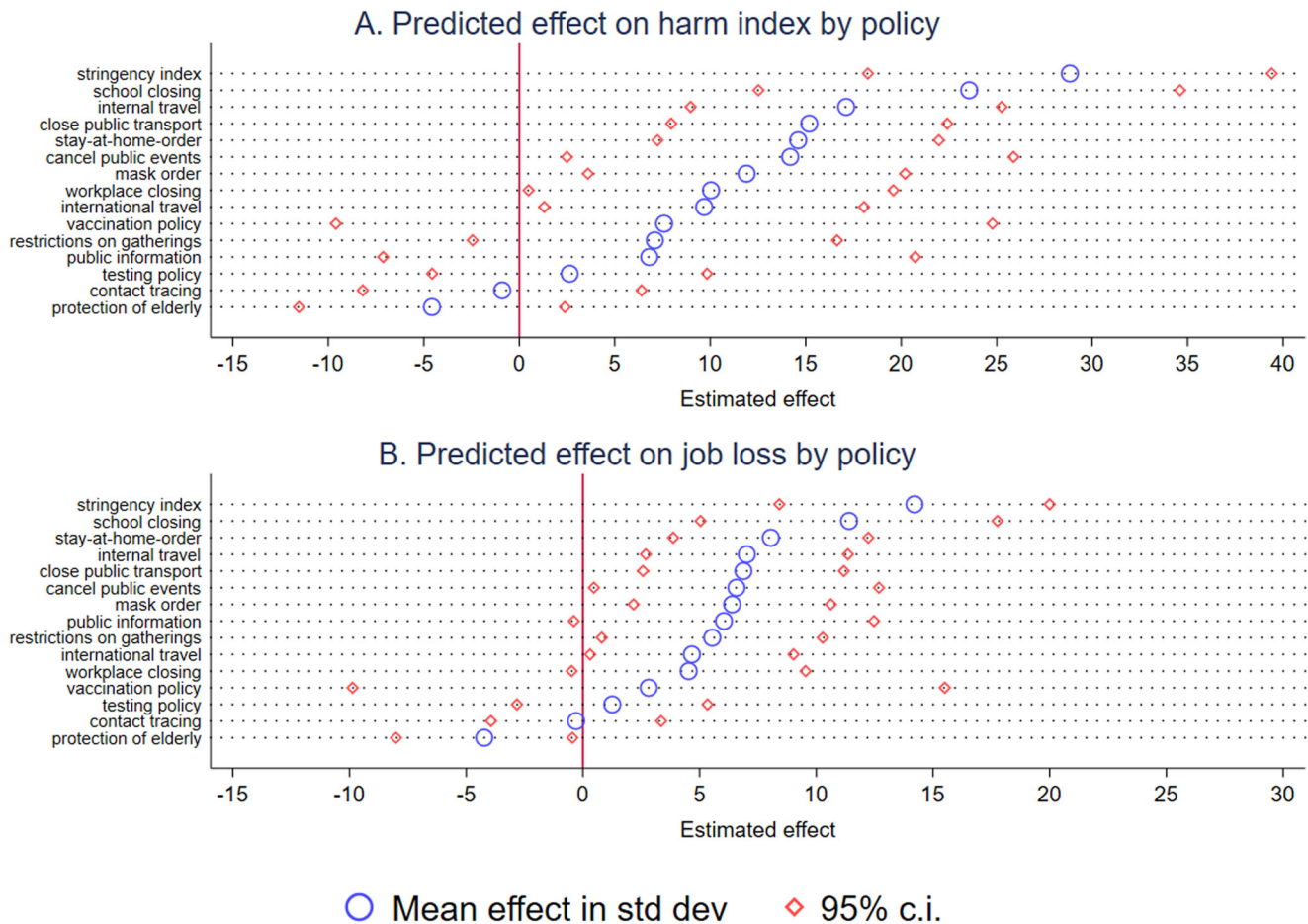
Fig. 3 Alternative policy measures and economic harm. Plots of standardized coefficients of mixed multilevel regression models. Dependent variables are the harm index (**A**, at top) or the job loss rate (**B**, at bottom). The figures plot the coefficients and 95% confidence intervals of 15 policy measures, each separately estimated, standardized, and measured as cumulative-to-date averages at the time the respondent was surveyed. The model includes controls for reported COVID-19 deaths per capita and an index of economic support, as well as individual demographic data with country and month-year fixed effects.

data collection agencies worldwide—is that the pandemic forced Gallup's partners to switch the mode of collection from face-to-face to phone-based collection in most countries during the early periods of the pandemic. The data were mode-adjusted, but there may be residual error from these adjustments affecting the results.

To be clear, these data do not allow for any clear cost-benefit calculation with respect to the pandemic-related policies. It is well beyond the scope of this research to attempt to estimate benefits and how they might vary by country or groups of people within countries. These data, however, could help provide benchmark estimates for costs, though we stress that our estimates should not be interpreted as causal effects, but rather associations between economic outcomes and policies.

A key question for future research is whether alternative policies could have been enacted—or could be enacted during future pandemics—that are able to save lives and mitigate disease without generating widespread economic harm, or disproportionate harm to the poorest households. These are difficult issues involving substantial uncertainty as well as value judgments that are sure to vary across individuals and countries.

Data availability
All replication code (Stata v17), microdata, and aggregate data needed to replicate the analysis are available here

https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/8SET4W. Questions should be directed to the corresponding author.

Notes
1 While China was part of the World Poll, it is not included in the analysis, since these items were not fielded on its survey.
2 A weighting variable was calculated by Gallup statisticians in multiple stages for each respondent to allow researchers to generate nationally representative estimates for each country. Base weights were calculated based on the probability of selection and post-stratification weights were calculated to adjust for non-response bias by gender, age, and, in some countries, education or socioeconomic status.
3 Stringency measures are missing for several countries in our sample, including Guinea, North Macedonia, Montenegro, and Kosovo. We found no relationship between our harm index and whether stringency, deaths per capita, inequality, or flu growth variables were missing. The correlation between the harm index and whether these variables are missing is approximately zero.
4 Formally, we standardize the three outcome variables: our original harm index, the PCA-based harm index, and the labor market harm index (with 4 instead of 5 items) and run the models shown in Table 1 column 4. The coefficient on stringency is 0.38, 0.39, and 0.38, across the respective models. These results are shown in Supplementary Table 10, panel A.

References

Abaluck J, Kwong LH, Styczynski A, Haque A, Kabir MA, Bates-Jefferys E, Mobarak AM (2022) Impact of community masking on COVID-19: a cluster-randomized trial in Bangladesh. Science 375(6577):eabi9069

Aminjonov U, Bargain O, Bernard T (2021) Gimme shelter. Social distancing and income support in times of pandemic. IZA DP No. 14967

Barkay N, Cobb C, Eilat R, Galili T, Haimovich D, LaRocca S, Morris K, Sarig T (2020) Weights and methodology brief for the COVID-19 symptom survey by University of Maryland and Carnegie Mellon University, in Partnership with Facebook Version 3–19. https://covidmap.umd.edu/document/css_methods_brief.pdf

Berry CR, Fowler A, Glazer T, Handel-Meyer S, MacMillen A (2021) Evaluating the effects of shelter-in-place policies during the COVID-19 pandemic. Proc Natl Acad Sci USA 118(15):e2019706118

Boone L, Ladreit C (2021) Fear of COVID and non-pharmaceutical interventions: an analysis of their economic impact among 29 advanced OECD countries COVID. Economics (73)

Bring J (1996) A geometric approach to compare variables in a regression model. Am Stat 50(1):57–62

Bundervoet T, Dávalos ME, Garcia N (2021) The short-term impacts of COVID-19 on households in developing countries: an overview based on a harmonized data set of high-frequency surveys. Policy Research Working Paper No. 9582, The World Bank, Washington, DC

Carvalho VM, Garcia JR, Hansen S, Ortiz Á, Rodrigo T, Rodríguez Mora JV, Ruiz P (2021) Tracking the COVID-19 crisis with high-resolution transaction data. R Soc Open Sci 8(8):210218

Chernozhukov V, Kasahara H, Schrimpf P (2021) Causal impact of masks, policies, behavior on early COVID-19 pandemic in the US. J Econ Pandemic Econ 220(1):23–62. https://doi.org/10.1016/j.jeconom.2020.09.003

Coibion O, Gorodnichenko Y, Weber M (2020) The cost of the COVID-19 crisis: lockdowns, macroeconomic expectations, and consumer spending. NBER Working Paper. p. 27141

Deb P, Furceri D, Ostry JD, Tawk N (2020) The economic effects of COVID-19 containment measures. IMF Working Paper WP/20/158

Egger D, Miguel E, Warren SS, Shenoy A et al (2021) Falling living standards during the COVID-19 crisis: quantitative evidence from nine developing countries Sci Adv 7(6):eabe0997

Fan J, Li Y, Stewart K, Kommareddy AR, Garcia A, Ma J, Liu Z, O'Brien J, Bradford A, Deng X, Chiu S, Kreuter F, Barkay N, Bilinski A, Kim B, Galili T, Haimovich D, LaRocca S, Presser S, Morris K, Salomon JA, Stuart EA, Tibshirani R, Barash TA, Cobb C, Gros A, Isa A, Kaess A, Karim F, Eliat R, Kedosha OE, Matskel S, Melamed R, Patankar A, Rutenberg I, Salmona T, Vannette D (2020). COVID-19 World Symptom Survey Data API. https://covidmap.umd.edu/api.html

Fetzer T, Graeber T (2021) Measuring the scientific effectiveness of contact tracing: evidence from a natural experiment. Proc Natl Acad Sci USA 118(33):e2100814118

Forsythe E, Kahn LB, Lange F, Wiczer D (2020) Labor demand in the time of COVID-19: evidence from vacancy postings and UI claims. J Public Econ 189:104238

Furceri D, Loungani P, Ostry JD, Pizzuto P (2022) Will COVID-19 have long-lasting effects on inequality? Evidence from past pandemics. J Econ Inequal 20:811–839

Gathergood J, Guttman-Kenney B (2021) The English patient: evaluating local lockdowns using real-time COVID-19 & consumption data. Covid Economics 64, Centre for Economic Policy Research

Google LLC (2021) Google COVID-19 Community Mobility Reports. https://www.google.com/covid19/mobility/ Accessed 31 Mar 2021

Goolsbee A, Syverson C (2020) Fear, lockdown, and diversion: comparing drivers of pandemic economic decline 2020. J Public Econ 193:104311

Gupta S, Simon K, Wing C (2020) Mandated and voluntary social distancing during the COVID-19 epidemic. Brookings Papers on Economic Activity. pp 269–316

Hale T, Angrist N, Goldszmidt R, Kira B, Petherick A, Phillips T, Webster S, Cameron-Blake E, Hallas L, Majumdar S, Tatlow H (2021) A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). Nat Hum Behav 5:529–538

Herby J, Jonung L, Hanke S (2022) A literature review and meta-analysis of the effects of lockdowns on COVID-19 mortality. Studies in Applied Economics (200)

Institute for Health Metrics and Evaluation (2021) Estimation of total mortality due to COVID-19. http://www.healthdata.org/special-analysis/estimation-excess-mortality-due-covid-19-and-scalars-reported-covid-19-deaths

Josephson A, Kilic T, Michler JD (2021) Socioeconomic impacts of COVID-19 in low-income countries. Nat Hum Behav 5:557–565

Kim LY, Lugo MA, Mason AD, Uochi I (2021) Inequality under COVID-19: taking stock of high-frequency data for East Asia and the Pacific. Policy Research Working Paper; No. 9859. World Bank, Washington, DC

Kugler M, Viollaz M, Duque D, Gaddis I, Newhouse DL, Palacios-Lopez A, Weber M (2021) How did the COVID-19 crisis affect different types of workers in the developing world? IZA Discussion Paper No. 14519

Liu Y, Morgenstern C, Kelly J, Lowe R; CMMID COVID-19 Working Group and Mark Jit (2021) The impact of non-pharmaceutical interventions on SARS-CoV-2 transmission across 130 countries and territories. BMC Med. 19:40

Lozano Rojas F, Jiang X, Montenovo L, Simon KI, Weinberg B, Wing C (2020). Is the cure worse than the disease? Immediate labor market effects of COVID-19 case rates and school closures in the US. Technical report, National Bureau of Economic Research

Mahler DG, Yonzan N, Lakner C, Castaneda Aguilar RA, Wu H (2021) Updated estimates of the impact of COVID-19 on global poverty: turning the corner on the pandemic in 2021? Data Blog (blog). June 24, 2021. https://blogs.worldbank.org/opendata/updated-estimates-impact-covid-19-global-poverty-turning-corner-pandemic-2021

Narayan A, Cojocaru A, Agrawal S, Bundervoet T, Davalos M, Garcia N, Lakner C, Mahler DG, Montalva Talledo V, Ten A, Yonzan N (2022) COVID-19 and economic inequality: short-term impacts with long-term consequences. Policy Research Working Paper, No. 9902. World Bank, Washington, DC

Nivette AE, Zahnow R, Aguilar R et al. (2021) A global analysis of the impact of COVID-19 stay-at-home restrictions on crime. Nat Hum Behav 5:868–877

Oshchepkov A, Shirokanova A (2022) Bridging the gap between multilevel modeling and economic methods. Soc Sci Res 104:102689

Qi J, Zhang D, Zhang X et al. (2022) Short- and medium-term impacts of strict anti-contagion policies on non-COVID-19 mortality in China. Nat Hum Behav 6:55–63

Rothwell J, Smith E (2021) Socioeconomic status as a risk factor in economic and physical harm from COVID-19: evidence from the United States. Ann Am Acad Pol Soc Sci 698(1):12–38

Spiegel M, Tookes H (2022) All or nothing? Partial business shutdowns and COVID-19 fatality growth. PLOS ONE 17:e0262925

Tascam Giorgio A, Illing V, Joyce AS, Ogrodniczuk JS (2009) Three-level multilevel growth models for nested change data: a guide for group treatment researchers. Psychother Res 19(4–5):453–461

Violato C, Violato EM, Violato EM (2021) Impact of the stringency of lockdown measures on COVID-19: a theoretical model of a pandemic. PLOS ONE 16(10):e0258205

Wang H, Paulson KR, Pease SA, Watson S et al (2022) Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020–21 Lancet 399:1513–1536

World Bank (2022) Global Economic Prospects, January 2022. Washington, DC, World Bank. https://doi.org/10.1596/978-1-4648-1758-8

Author contributions

JTR: conceptualization, formal analysis, methodology, writing of original draft and revised drafts. AC: formal analysis, methodology, and writing and revisions. YSK: writing, review, and editing. RS: data collection, methodology, writing, and reviewing. All authors substantially contributed to the article and approved the submitted version.

Ethical approval

Prior to fielding, the World Poll project was reviewed by Gallup's Institutional Review Board (IRB) and approval was granted in January of 2020. Gallup's IRB is publicly registered with the U.S. Department of Health and Human Services (IRB00004888). Documentation of this registration is publicly available on the agency's website, by typing "Gallup" into the search box: https://ohrp.cit.nih.gov/search/search.aspx?styp=bsc. The names of the IRB committee members and their highest level of education is provided here: Lloyd, Camille (MS); Chattopadhyay, Manas (PhD); Rheault, Magali (MS); Houts, Jonah M (MBA); Marken, Stephanie (MA); Phares Sharon (PhD); Royal, Dawn (MS); Fernandez, Cheryl F (PhD); McMurray, Coleen (BA); Odle, Adam (Esq); Ford, Kenya (AA). The database used in this analysis has no personally identifying information. Responses are linked to individuals using a quasi-random number with no external meaning.

As the scale of the coronavirus pandemic became clear in March 2020, Gallup paused global data collection to thoroughly assess risk and prepare contingency plans for data collection. By May, the continued prevalence of COVID-19 made it clear that there was too much risk of community transmission to conduct face-to-face data collection in 2020 for the World Poll. Gallup prepared a contingency methodology based entirely on telephone interviewing. In designing the new approach, Gallup prioritized the safety of interviewers and respondents, retaining high levels of representativity, and ensuring high-quality data collection. Countries were selected for inclusion based on client interest, their population size, phone penetration rates, and Gallup's local partner's ability to execute Computer Assisted Telephone Interviewing (CATI) aligned with Gallup's quality assurance standards. With more than 100 countries included, the 2020 World Poll represents more than 90% of the global adult population. As a standard practice, Gallup and its partners complied with all government-issued guidance from local authorities and took this guidance into account throughout the interviewing process, including following social distancing measures for telephone interviews. Collection began

in July 2020 with these issues in mind and Gallup's weighting strategy adjusts for the mode of data collection. In July and August 2020, collection was exclusively phone-based, but gradually included face-to-face collection as the year went on.

Informed consent

Formal consent was obtained verbally from all research participants in accordance with Gallup's IRB standards. The database used in this analysis has no personally identifying information. Responses are linked to individuals using a quasi-random number with no external meaning. Gallup and its research partners obtain consent from every respondent surveyed. Both consent requests and the survey itself are conducted in the prominent local languages of each specific region, and the survey is conducted in multiple languages in many countries. The survey is only given to people who are at least 15 years old. The English version of the consent language request used in the World Poll data analyzed in the manuscript, read as follows: "Hello, this is _____ calling on behalf of Gallup. Your household has been selected as one of only 1,000 in the [country of respondent] to participate in a special poll about this country… Your participation in this survey is voluntary. Your information will be kept private and combined with responses of other people so that individuals cannot be identified. Do you agree to participate in the survey?" For countries where 15–17-year-olds are not considered adults, the interviewer also collects parental consent. Are you age 18 or older? If not, respondent is asked, "May I speak to your parent or guardian to obtain permission to interview you? Hello, this is _____ with _____. Your child has been selected as one of only 1,000 people in [country of respondent] to participate in a special poll about this country. May I have your permission to interview your child to get their thoughts and opinions about issues in this country?"